

Large scale features and assessment of spatial scale correspondence between TMPA and IMD rainfall datasets over Indian landmass

R UMA¹, T V LAKSHMI KUMAR¹, M S NARAYANAN^{1,*}, M RAJEEVAN²,
JYOTI BHATE³ and K NIRANJAN KUMAR⁴

¹*Department of Physics, SRM University, Kattankulathur, Chennai 603 203, India.*

²*Ministry of Earth Sciences, Prithvi Bhavan, Lodi Road, New Delhi 110 003, India.*

³*National Atmospheric Research Laboratory, Gadanki 517 112, India.*

⁴*Divecha Centre for Climate Change, Indian Institute of Science, Bangalore 560 012, India.*

**Corresponding author. e-mail: u.m.s.narayanan@gmail.com*

Daily rainfall datasets of 10 years (1998–2007) of Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) version 6 and India Meteorological Department (IMD) gridded rain gauge have been compared over the Indian landmass, both in large and small spatial scales. On the larger spatial scale, the pattern correlation between the two datasets on daily scales during individual years of the study period is ranging from 0.4 to 0.7. The correlation improved significantly (~ 0.9) when the study was confined to specific wet and dry spells each of about 5–8 days. Wavelet analysis of intra-seasonal oscillations (ISO) of the southwest monsoon rainfall show the percentage contribution of the major two modes (30–50 days and 10–20 days), to be ranging respectively between ~ 30 –40% and 5–10% for the various years. Analysis of inter-annual variability shows the satellite data to be underestimating seasonal rainfall by ~ 110 mm during southwest monsoon and overestimating by ~ 150 mm during northeast monsoon season.

At high spatio-temporal scales, viz., $1^\circ \times 1^\circ$ grid, TMPA data do not correspond to ground truth. We have proposed here a new analysis procedure to assess the minimum spatial scale at which the two datasets are compatible with each other. This has been done by studying the contribution to total seasonal rainfall from different rainfall rate windows (at 1 mm intervals) on different spatial scales (at daily time scale). The compatibility spatial scale is seen to be beyond $5^\circ \times 5^\circ$ average spatial scale over the Indian landmass. This will help to decide the usability of TMPA products, if averaged at appropriate spatial scales, for specific process studies, e.g., cloud scale, meso scale or synoptic scale.

1. Introduction

Rainfall estimation from meteorological satellites have been pursued ever since the launch of the first meteorological satellite, TIROS-1 (April 1, 1960). Satellites, particularly of the geostationary class, with visible and infrared sensors, despite

their limitation of spatio-temporal scales in comparison with ground measurements, have continued to provide global rainfall on a continuous basis for more than 3 decades (Barrett and Martin 1981; Rao *et al.* 1990; Kidder and Vonder Haar 1995). With the availability of microwave measurements (albeit with poorer horizontal resolution, but with

Keywords. Southwest monsoon rainfall; TMPA; IMD gridded data; spatial scales; ISO.

a capability to retrieve rainfall through direct interpretation of radiances in terms of rain), there have been steady improvements in the precipitation measurements from space (Mishra *et al.* 2010). TRMM, with its radar, looking directly at raindrops at high spatial resolution (~ 4 km) compared to passive microwave radiometers (~ 20 km), has made a very significant contribution in this direction for estimating instantaneous rainfall from space. The TRMM Multi-satellite Precipitation Analysis (TMPA) product available from NASA site <http://disc.sci.gsfc.nasa.gov/>, has provided an excellent global rainfall database at high spatio-temporal resolution ($0.25^\circ \times 0.25^\circ$ and 3 hourly) since 1998. However, its accuracy *vis a vis* ground measured rainfall has to be assessed over individual rainfall regimes, both over land and ocean.

The study by Narayanan *et al.* (2005), on validation of TRMM 3B42 V5 data with gridded IMD rain gauge data over Indian landmass revealed that 3B42 V5 does not pick up small (< 1 mm) and very high (> 80 mm/day) daily average rainfall. Thus, the daily variance estimated by 3B42 V5 was found to be poor compared to that provided by gauge-analysed data. This was explained as due to predominance of indirect IR estimates (over direct estimates by microwave methods) of rainfall in the 3B42 dataset. Studies of Rahman *et al.* (2009) suggest that although satellite data are reasonably good on large spatial scales, at individual grid points the agreement with gauge data is far from perfect. They have also mentioned that interannual variability in satellite rain has the right phase but the amplitude is generally underestimated, over orographic regions in particular.

The study of Islam *et al.* (2005) on comparison of TRMM 3B42 products with surface rainfall measured at 31 stations over Bangladesh for five years (1998–2002) infers that TRMM overestimates the rainfall during pre-monsoon (March–May) and underestimates during monsoon (June to September) and also it underestimates over heavy-rainfall region on the whole. Validation of 3B42 datasets over Thailand with more than hundreds of rain gauge data by Chokengamwong and Chiu (2005) show that 5-year (1998–2002) daily average rainfall for gauge, 3B42-V5 and 3B42-V6 are 4.7, 5.6 and 4.6 mm/day, respectively. The TRMM merged analysis product on monthly scale has the closest agreement with rain gauge data. Comparative study of TRMM data with Global Precipitation Climatology Centre (GPCC) product (Adeyewa and Nakamura 2003) has shown that TRMM PR data overestimates rain in the tropical rain forest region of Africa.

Xie *et al.* (2007) studied the performance of five satellite-based high-resolution precipitation outputs by comparing them with gauge analyzed

data: TRMM (3B42, 3B42RT) National Weather Service, USA (NWS) product, Climate Prediction Centre (CPC), Morphing Techniques (CMORPH) and Naval Research Laboratory (NRL). They found that all satellite datasets performed better in depicting precipitation for wet climate regions and wet seasons, but were having limited skills in estimating precipitation over central Asian arid and semi-arid regions.

The merged TRMM product (TMPA) is being extensively used by a large number of researchers, primarily to validate results of model forecast/hindcast experiments and new satellite observations. While there is a reasonably adequate network of conventional observations for rainfall measurements over many areas of land, those over the ocean are very sparse. For validation over oceans, TMPA data have to be compared only with another independent satellite observation. Over the coastal oceans, however, we have coastal rain radars.

The TMPA satellite based global rain products are available at very high spatio-temporal resolution ($0.25^\circ \times 0.25^\circ$ spatial grid and 3 hourly temporal). These TMPA products are being used unreservedly in many studies to compare with other rainfall results (including model outputs), both over land and oceans – assuming these products to be representative of the situation at such fine resolutions.

It has been attempted here to arrive through a detailed analysis, the minimum spatial scale (at daily temporal scale) at which the TMPA data can be considered a proxy/representative of ground truth over the Indian landmass. It could, however, be different over different climate regimes of the globe, and over the oceans. Such an assessment of TMPA spatial scale is expected to help us study meteorological processes of different scales (e.g., cloud, meso or synoptic scale) using appropriate grid averaged TMPA data.

The present study has been presented in two parts. In part 1 (section 3), results of comparison of TMPA 3B42 V6 data with India Meteorological Department (IMD) gridded ground gauge data over Indian landmass (Rajeevan *et al.* 2006), for the 10-year period 1998–2007 over large spatial scale (all India scale) are presented. The comparison has been carried out both for southwest (SW) and northeast (NE) monsoon seasons. In addition, the dominant modes of intraseasonal oscillation (ISO) have been examined to understand the contribution of dominant periodicities to the seasonal rainfall.

In part 2 (section 4), we present a new analysis procedure to assess the minimum spatial scale at which the two datasets compare satisfactorily. For arriving at this spatial scale, the two rainfall datasets have been analysed, besides the basic

1°×1° scale, in average spatial scales of 2°×2°, 5°×5° and in all India scale. Further, the contribution to total seasonal rainfall from different rainfall rate windows (at 1 mm intervals) have been studied on these different spatial scales (all at daily time scale). From this analysis, the compatibility spatial scale has been arrived at and this is suggested to be beyond 5° × 5° average spatial scale.

2. Data

The present study makes use of 10 years (1998–2007) of southwest monsoon data over Indian landmass from the following two sources:

- TRMM Multi-satellite Precipitation Analysis (TMPA) 3B42 V6 global rainfall product, which are available at the NASA website <http://disc.sci.gsfc.nasa.gov> at 0.25°×0.25° lat.–long. spatial scale and at 3-hr time intervals (Adler *et al.* 2007).
- IMD gridded rainfall product (Rajeevan *et al.* 2006) over Indian landmass at 0.5°×0.5° lat.–long. spatial and daily time scales.

For the present analysis, both the above two datasets were brought to an identical 1°×1° spatial and daily (24 hr cumulative) time scale over the Indian landmass. At this spatial scale there were 333 grids for the Indian landmass in both the datasets for any single day.

3. Large scale features

Figure 1(a, b) shows the spatial distribution over India of 10-year seasonal mean rainfall pattern over India during the southwest monsoon season from IMD and TMPA data. High rainfall regions over Western Ghats and northeast India, and scanty rainfall over northwest and southern parts of India are brought out clearly from both datasets, albeit with quantitative differences – by even as much as 100% over the Western Ghats. The overall pattern

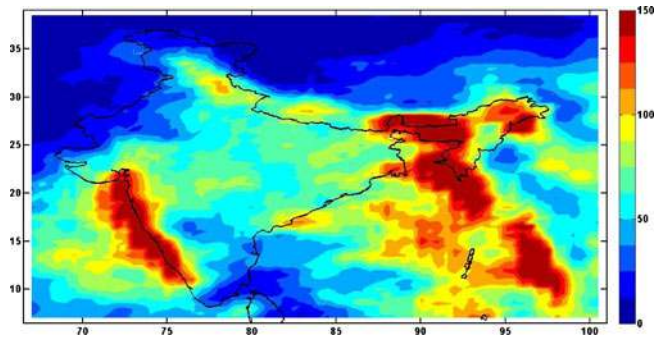


Figure 2. Spatial variability of 10-year (1998–2007) mean monthly southwest monsoon rainfall from TMPA over India and the adjoining ocean.

correlation between the two datasets averaged over the Indian landmass is 0.78. From these figures, the activity of SWM covering major portion in India can be seen. Parts of Eastern Ghats and leeward side of Western Ghats represent the rain shadow regions over India. The confluence of SWM can be seen at the head Bay where highest rainfall ranging from 150 to 200 mm is experienced.

Though the present study is confined only over Indian landmass, it is well known that monsoon activity has an equally important oceanic component. Figure 2 shows the TMPA 10-year climatology, including the adjoining oceans and landmass of neighbouring countries. Heavy rainfall over Bay of Bengal, Myanmar coast and on the windward side of the Western Ghats in the Arabian Sea, etc., have been well picked up by TMPA data. Low rainfall amounts on the leeward side of Eastern Ghats over southwest Bay of Bengal is also well marked. However, there is no adequate sea-truth data at this spatio-temporal scale for comparison with TMPA data for a detailed analysis.

It is interesting to see the resemblance of drought/flood year rainfall anomalies from these two datasets. Since, drought/flood contrast can be seen only over specific regions, we have addressed the spatial anomalies for the years 2002 (bad

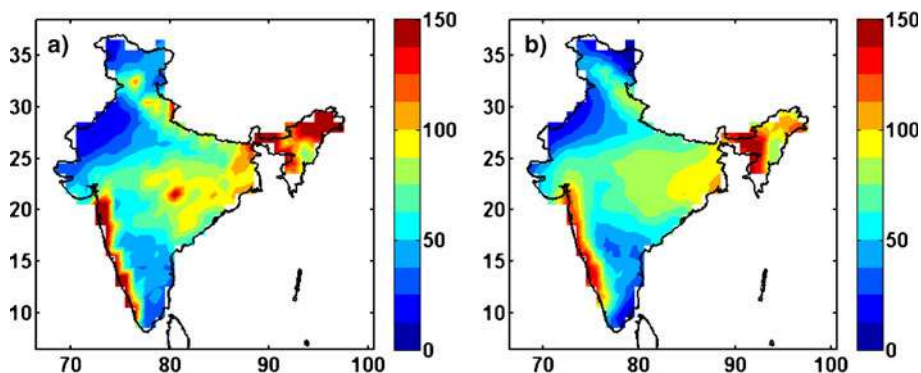


Figure 1. Spatial variability of 10-year (1998–2007) mean of southwest monsoon rainfall: (a) IMD and (b) TMPA.

monsoon) and 2007 (good monsoon) (figure 3a–d). It is observed from the figure that in the year 2002 the southern peninsular region, including parts of Western Ghats, is found to have recorded deficient rainfall ranging from -40% to -20% from the normal by both datasets. Also, similarity has been observed in a few parts of central India with normal monsoon activity by the two datasets. It is reported by IMD that the subdivisional regions, viz., west and east Rajasthan experienced severe drought ($>50\%$ below normal) and Haryana and Punjab experienced moderate drought (26–50% below normal) (IMD Season Report 2002). These features are also observed by and large from TMPA datasets. There were differences between the two datasets over Himalayan region, Indo Gangetic plains, a few portions of northeast India and Kerala. These differences could be because of the inability of satellite to capture the orographic rainfall (Rahman *et al.* 2009). The deficient rainfall activity over Western Ghats in the year 2002 is reported by Gadgil *et al.* (2003, 2005), Kalsi *et al.* (2004) and has been explained as due to a moderate El Niño condition during late June which extended to August.

The year 2007 displayed high variability with active monsoon on many days in June, September and first week of July and August as reported by IMD (Annual Summary of Climate 2007) which

has been captured well by both the datasets (figure 3c, d). The TMPA anomaly for this year has also shown similar features in peninsular India but not to the same extent in rest of India. The region along the east coast from 12.5° to 16.5°N has recorded excess rainfall as observed from the TMPA and IMD datasets. Deficient rainfall activity over parts of Indo Gangetic plains is depicted by both datasets (figure 3c, d).

The time series of 10-year climatology of southwest monsoon all India average daily rainfall for the two datasets are shown in figure 4. IMD 30-year climatology is also shown. High rainfall (>8 – 10 mm/day) is observed during the months of July and August and lower values (~ 4 – 7 mm) during the months of June and September. Correlation coefficient between the mean daily rainfall of 3B42 dataset and IMD data is 0.8. However, the total rainfall recorded by TMPA 3B42 and IMD are 832 and 939 mm (IMD 30-year mean is 940 mm), respectively. The satellite data is seen to be showing a negative bias of about 110 mm for the season as a whole. The seasonal total rainfall for each of the 10 years period (of 122 days), their standard deviation and correlation coefficient of the IMD and TMPA daily rain are given in table 1. The year-wise seasonal daily correlation is varying from 0.4 to 0.7. In the above computations, the contribution from orographic grids (with heights >500 m)

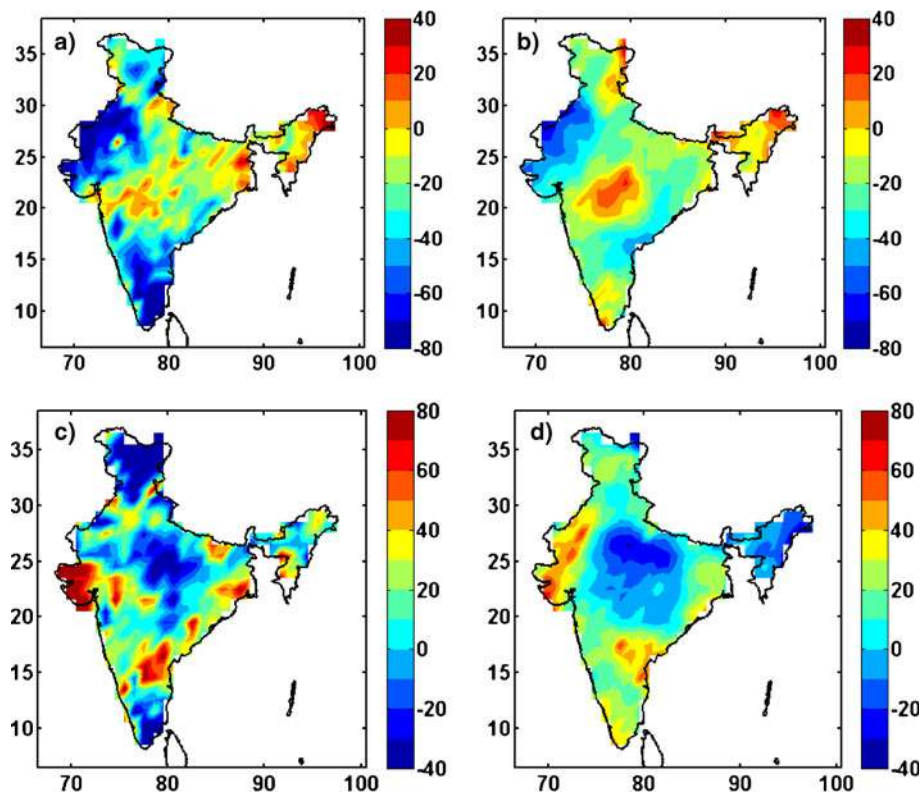


Figure 3. Spatial anomalies of southwest monsoon rainfall for: (i) poor monsoon of 2002 (a) IMD, (b) TMPA, and (ii) good monsoon of 2007 (c) IMD and (d) TMPA.

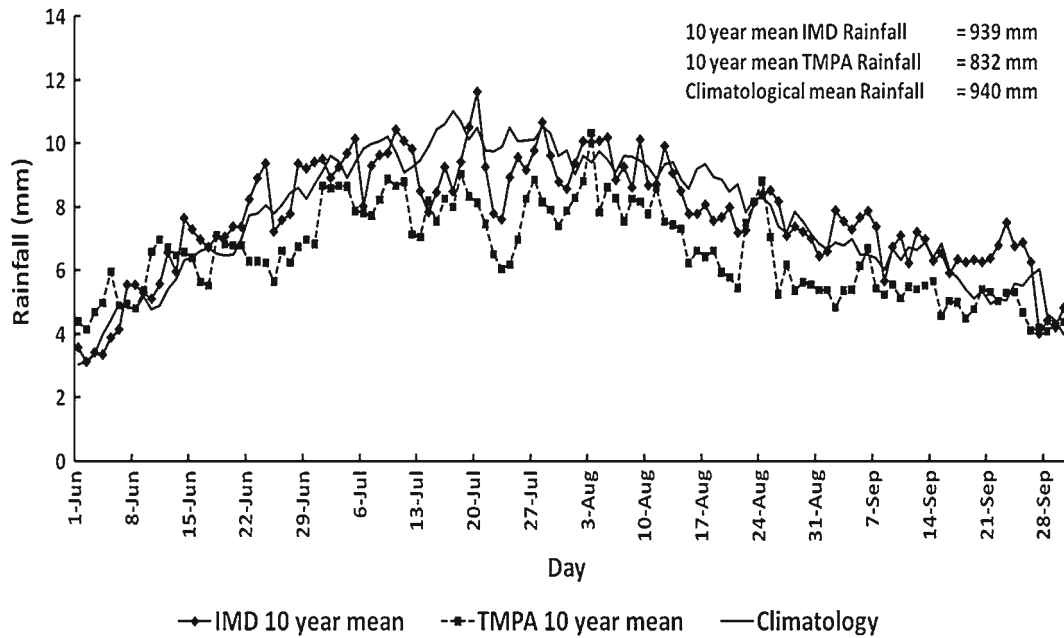


Figure 4. Time series of daily 10-year mean all India rainfall for TMPA, IMD and IMD 30 years climatology (1961–1990).

Table 1. Seasonal total, daily mean, standard deviation and correlation.

Year	Seasonal total (mm)		Daily mean and standard deviation (mm)		Correlation of daily rainfall of IMD and TMPA (r)
	IMD	TMPA	IMD	TMPA	
1998	987	856	8.1 ± 2.9	7.0 ± 3.3	0.49*
1999	872	780	7.1 ± 3.1	6.4 ± 3.2	0.45*
2000	843	774	6.9 ± 3.4	6.3 ± 3.1	0.38*
2001	870	807	7.1 ± 3.0	6.6 ± 3.3	0.54*
2002	764	695	6.3 ± 3.2	5.7 ± 2.9	0.65*
2003	978	855	8.0 ± 3.4	7.0 ± 2.9	0.67*
2004	865	823	7.1 ± 3.7	6.7 ± 3.2	0.57*
2005	1144	839	9.4 ± 5.6	6.9 ± 3.5	0.69*
2006	1007	892	8.3 ± 3.7	7.3 ± 3.4	0.20**
2007	1064	997	8.7 ± 3.4	8.2 ± 3.1	0.50*
Mean of 10 years	939	832	7.7 ± 3.5	6.8 ± 3.2	0.80*

*: 0.01 level of significance **: 0.05 level of significance.

amounting to 98 numbers have also been included. In general, their contribution to the total all India rainfall is about 30% and 45% in the two datasets. The correlations between the rainfall values of the two datasets over the orographic region during all the years of the study period are observed to be low.

The pie diagram (figure 5) shows the climatological monthly contribution to total all India rainfall. Here again, though as percentage contribution the two datasets are in good agreement with each other, the absolute values of rainfall are at variance.

The daily rainfall time series for one of the monsoon years (2002) is shown in figure 6. The TMPA data has picked the active and break spells of the Indian monsoon as observed in IMD data. Many of the significant peaks and troughs of IMD are also picked up in phase by TMPA dataset; however, notable differences at some epochs can also be seen. The correlation between the two time series is 0.65 for 2002. There is a significant inter-annual variation in the correlation between the two datasets. For the 10 years considered in this study, it varied between 0.4 and 0.7, except for the year 2006, when it was very poor, being just

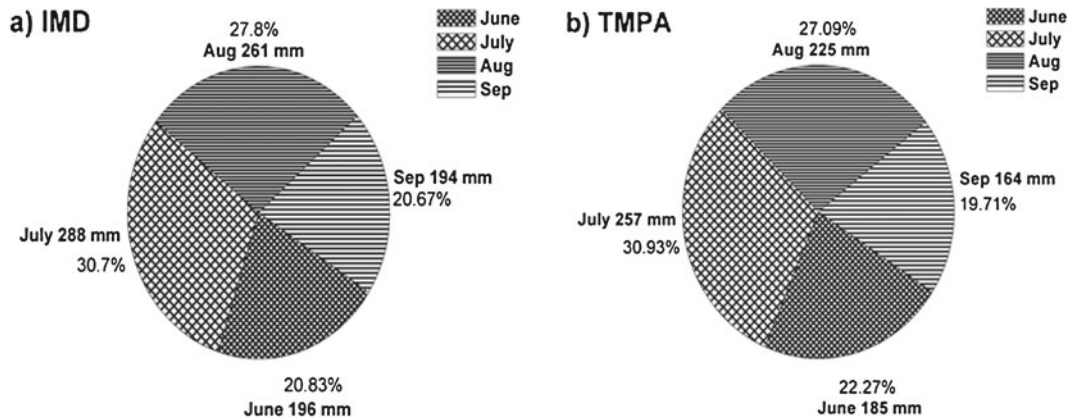


Figure 5. Pie diagram showing percentage contribution of mean monthly all India rainfall: (a) IMD and (b) TMPA.

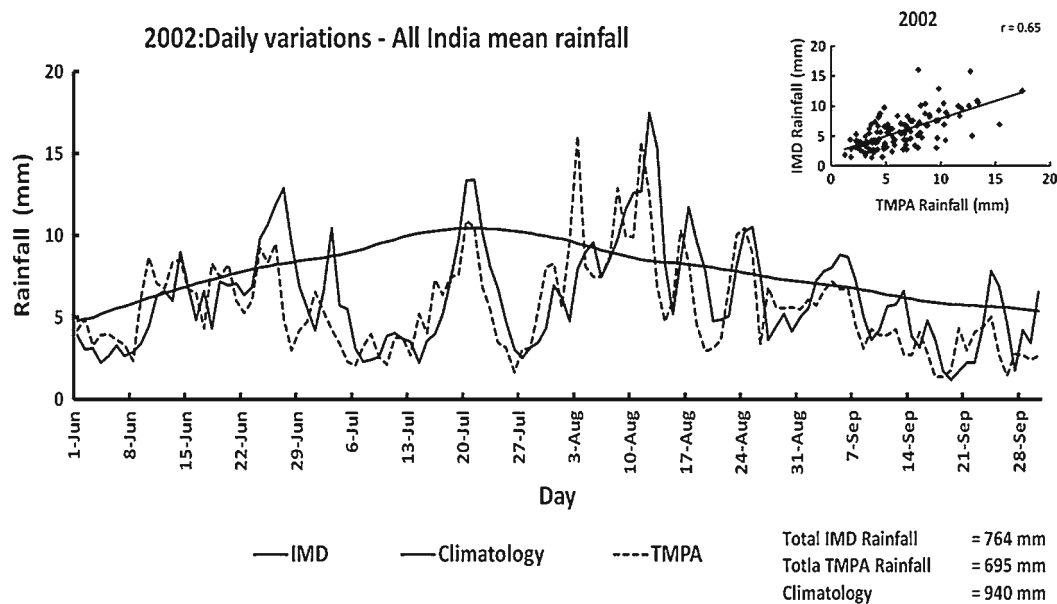


Figure 6. Time series of daily all India average rainfall for the year 2002: IMD, TMPA and IMD 30 years climatology. Inset: scatter plot of TMPA and IMD rainfall.

0.2. This poor correlation in the year 2006 may be because of:

- prolonged dry spell of about 15 days over India when the difference in TMPA and IMD rainfall was considerably high,
- a major difference of 60 mm rainfall for the spell during 24 August–14 September. This was recorded as wet spell by IMD data but not by TMPA. It is also noteworthy to mention that Rahman *et al.* (2009) observed satellite rainfall to have spurious differences in recording the actual rainfall on some occasions. They have concluded that the inadequate spatial coverage of IMD gauges and inability of TMPA to capture actual rainfall over orographic region could be reasons for the same.

The correlation substantially improved to ~ 0.9 (figure 7) when the analysis was confined to individual wet and dry epochs (each of around 5–10 days) during the various years. There were a total of 37 and 32 cases of such wet and dry spells respectively in the 10-year period. In the present study, the criteria for wet and dry spells have been adopted from Rao *et al.* (2009). A wet/dry spell is three or more continuous days for which all India average rainfall is greater/less than one standard deviation from the climatological mean (1961–1990). The outlier corresponding to 120 mm rainfall was observed from TMPA data in a dry spell where the dry spell has been obtained based on IMD data (figure 7b). During this period, i.e., from 8 to 23 June 2006, IMD has recorded 62 mm of rainfall. As we have mentioned earlier in the

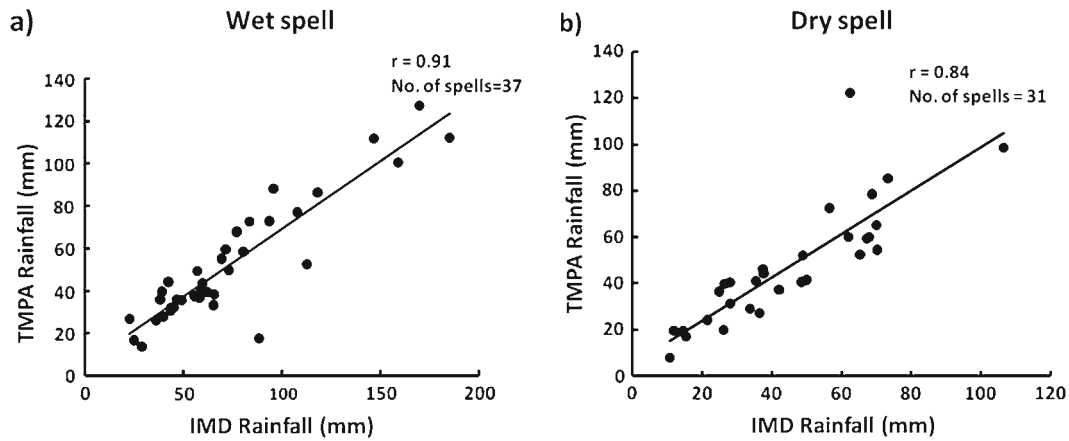


Figure 7. Scatter plot of cumulative rainfall during (a) wet spells and (b) dry spells.

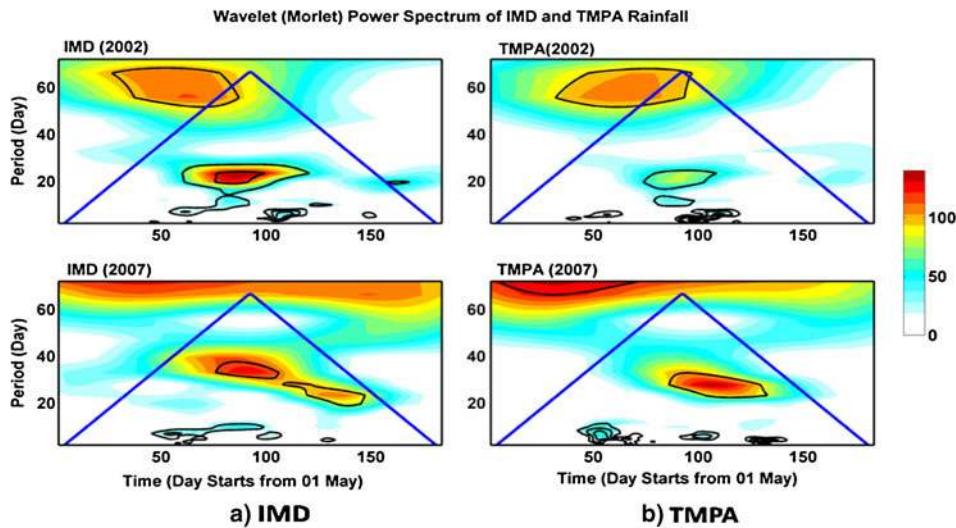


Figure 8. Wavelet power spectrum, with cone of influence, of monsoon rainfall showing dominant periodicities during poor monsoon (2002) and good monsoon (2007). (a) IMD and (b) TMPA.

introduction, TMPA comparison with gridded ground gauge data is far from perfect during the heavy rainfall events. This could be one of the possible reasons for the outlier of 120 mm.

3.1 Intra-Seasonal Oscillations (ISO) of southwest monsoon rainfall

The southwest monsoon rainfall exhibits a substantial variability on the intra-seasonal scale between active spells with good rainfall and weak spells or breaks with little rainfall. The ISO of southwest monsoon rainfall essentially have time scales between 10 and 90 days, but have two preferred bands, one between 10 and 20 days and the other between 30 and 60 days (Krishnamurti and Bhalme 1976; Yansunari 1980; Goswami and Ajaya Mohan 2001). Here we show some similarities and differences between the IMD and TMPA rainfall

datasets in terms of intra-seasonal variability. We show the dominant modes associated with the monsoon rainfall using the wavelet (Morlet) analysis (Torrence and Compo 1998).

The Morlet wavelet is applied to the daily rainfall time series, for both the datasets, during May–October in each year from 1998 to 2007. For example, figure 8 shows the wavelet spectrum of daily rainfall time series for the years 2002 and 2007 respectively for IMD (left panel) and TMPA (right panel). The contours indicate the 95% confidence based on the Chi-square statistics. The figure indicates the existence of two period ranges, between 10 and 20 days and 30 and 50 days. The existence of 10–20 day oscillation has been shown previously by several studies and found that the oscillation is westward propagating mode closely related to monsoon active/break conditions (e.g., Krishnamurti and Bhalme 1976; Krishnamurti and Arduway

1980; Chen and Chen 1993). In addition to the 10–20 day oscillation, the 30–50 day oscillation is also seen in monsoon circulation (Dakshinamurthy and Keshavamurthy 1976). The 30–50 day and 10–20 day oscillations are seen in both IMD and TMPA datasets. However, the amplitude of the ISO in 10–20 day period band is underestimated by the TMPA.

We see a clear inter-annual variability in these oscillations during the monsoon months. Hence we made the comparison between the two datasets in terms of percentage contribution of both these important oscillations to total monsoon rainfall variability. Figure 9(a, b) shows the percentage contribution of the total power in the 10–20 day and 30–50 day modes. The percentage contribution is estimated by dividing the power in the 10–20 day and 30–50 day period band by total power in all the periodicities. It indicates that the 10–20 day oscillation merely contributes an average of ~ 5 – 10% during the years 1998–2007. There are some discrepancies in 10–20 day mode ISO between IMD and TMPA datasets in terms of percentage contribution. However, the 30–50 day oscillation in TMPA dataset is comparing well with that of IMD (figure 9b). The 30–50 day oscillation contributes an average of $\sim 40\%$ of total spectrum for both IMD and TMPA datasets during 1998–2007. Thus, it is observed that the low frequency variability of southwest monsoon rainfall is comparing well, whereas the high frequency variability have differences between the two datasets. It can be observed that the 30–50 days periodicity has the maximum variance explained over all the periodicities whereas the 10–20 day band has relatively less percentage variance. Also we found that the variance in the 10–20 day is much sensitive than the power in 30–50 day to other high frequencies. Hence there could be difference in the variance at the high frequencies spectra which may affect the

percentage at 10–20 day band. This effect is relatively less significant on 30–50 day band. However, the average percentage contribution of ISOs (10–20 day and 30–50 day) to the total monsoon rainfall variability between 1998 and 2007 is almost similar in both IMD and TMPA datasets during summer monsoon months.

3.2 Northeast monsoon

The comparison of IMD and TMPA datasets during northeast monsoon (NEM) over peninsular India has also been performed for the same 10-year period (details not presented here). The correlation coefficient of the 10-year mean daily rainfall of the two datasets is 0.7 (0.01 level of significance) and for the individual year's the correlation is ranging from 0.4 to 0.6 with an exception for the years 2000 and 2005 (0.1 and 0.06, respectively). There were two cyclonic events observed during the NEM of 2000 and 2005 (cyclone atlas database of IMD). The effect of the cyclones were seen as heavy rain over the latitudes ranging from 8.5° to 15.5°N . This may have affected in getting significant correlations between the two datasets during 2000 and 2005. The monthly rainfall contribution of each month to the seasonal rainfall recorded by the two datasets shows a good match with an exception for the month of December.

The comparisons show a positive bias of about 150 mm in the satellite derived TMPA rainfall. Islam *et al.* (2005) have also observed biases of different signs during pre-monsoon and monsoon over Bangladesh.

The large difference in the bias (including the sign of bias) of the satellite derived rainfall during the two monsoons (negative bias of 110 mm during southwest *vs.* positive bias of 150 mm during northeast) over India suggests some infrared retrieval related differences for the two types of monsoon

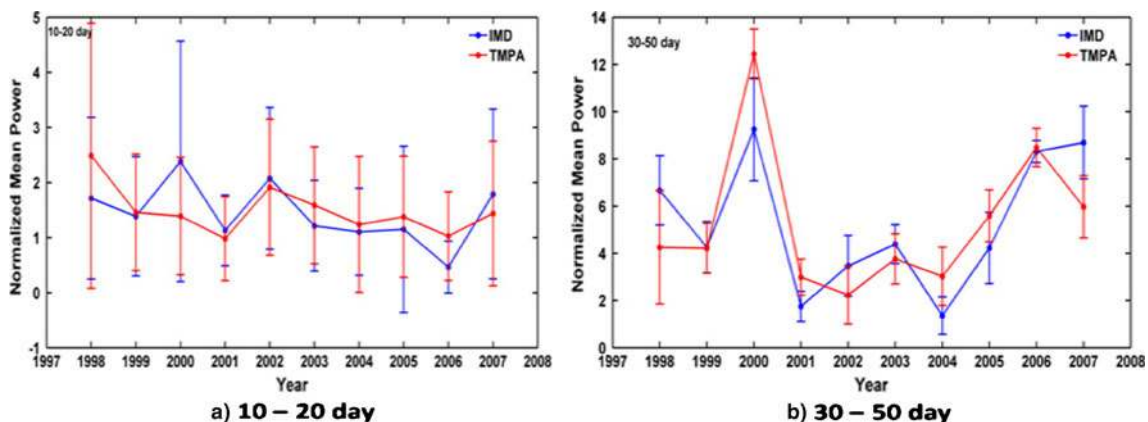


Figure 9. Time series of normalized mean power in (a) 10–20-day and (b) 30–50-day band for IMD and TMPA datasets during 1998–2007.

rainfall (thus, a need for applying different thresholds of brightness temperature of clouds in the empirical relationship on empirical estimation of daily rainfall).

4. Assessment of spatial scale

The merged TRMM products (TMPA) are available at high spatio-temporal resolution. These are being extensively used by a large number of researchers to compare with the results of model forecast experiments and/or other satellite observations. This could lead to erroneous conclusions, as the spatial scales of match between TMPA and ground truth (IMD gridded data) appear to be much larger. In this section, the contributions of different rainfall rate windows (at 1 mm intervals) towards total rainfall have been studied at various spatial scale averages (at daily time scale) for the two datasets. The error at different spatial averages has been estimated from this analysis, from which it has been possible to arrive at the compatibility spatial scale of the two datasets. The applicability of TMPA data to study specific meteorological phenomena (with specific spatio-temporal scale) depends on what type of spatial average is performed on this data.

4.1 Analysis procedure

The following analysis procedure was adopted for arriving at the minimum spatial scale compatibility:

- The number of grids (and corresponding rainfall amounts) in the two datasets at $1^\circ \times 1^\circ$ for different rainfall rates were first reconfirmed to be in fair agreement (log-normal distribution).
- Correlation between the two datasets at different spatial averages, viz., at $1^\circ \times 1^\circ$, $2^\circ \times 2^\circ$, $5^\circ \times 5^\circ$ and at all India scale for each year were performed. This yielded poor results even up to $5^\circ \times 5^\circ$ spatial averages. Though at individual pockets and during specific epochs, the correlations were better.
- Next, the distribution of number of TMPA grids (and associated rain values) with respect to IMD rain rates were examined. For this purpose, we created a large database of daily rain rates (in both datasets) in windows of 1 mm each from 0 mm up to 20 mm (normally encountered very high rainfall rates at $1^\circ \times 1^\circ$ average scale in IMD gridded rainfall data). This procedure is further elaborated below.

Three years of southwest monsoon data of 1998–2000 were clubbed together for the analysis

(yielding a total of 122 monsoon days/year \times 3 years = 366 days of data each over 333 grids of $1^\circ \times 1^\circ$ size over the Indian landmass). From this set, the grids over the Indian plains (i.e., grids with mean topography < 500 m) were selected, to exclude orographic rainfall. There were, thus, 235 such non-orographic grids over Indian plains. The dataset now consisted of 235 grids \times 122 days \times 3 years $\sim 90,000$ grids of $1^\circ \times 1^\circ$ size with rainfall rates of varying amounts. It was assumed that this dataset provides a fairly representative number of grids for all rain rates at least up to 10 mm/day (arrived from the daily IMD climatological averages).

The next analysis steps adopted by us are explained and illustrated below by taking as an example, the IMD rain rate window between 2 and 3 mm/day (central value 2.5 mm). The lat.–long. location, day and year of such rainfall rate grids were first identified in the IMD dataset of all three years. In the present example of 2.5 mm rain rate window, the number of IMD grids was 3625. Next, exactly for those grids (location, date and year) the rain rate values of TMPA grids were searched – most of which were surprisingly outside the 2–3 mm range.

4.2 Discussions

In an ideal case, a normal distribution of TMPA grid numbers should be centred around (or with a shift in case of a retrieval bias between the two datasets) the IMD value of 2.5 mm, as shown in figure 10(a). However, our results showed a large departure from a normal distribution (figure 10b). This distribution had a peak number (940) of TMPA grids having rainfall rates between 0 and 1 mm (central value 0.5 mm) and an equally large number, around 700 grids, with 0 mm rainfall. The remaining grids were having rainfall rates of higher values, many of them even beyond 10 mm (this number itself being more than 600).

This led us to carry out a more elaborate analysis for other rain rate windows of IMD rainfall. The distribution of number of TMPA points for each of IMD rain rate windows (each of 1 mm) had some similar pattern (i.e., the peak number occurring between 0 and 1 mm, a large number of grids with 0 mm rainfall and a good number of grids at other rain rates, even beyond 10 mm). Another point of concern was that there was hardly any change in trend in TMPA grid number distribution for increasing IMD rain rates. The peak in the number of TMPA grid number distribution should have ideally been shifting to higher rain rates, along with the shift of IMD windows to higher values. This was not found to be so.

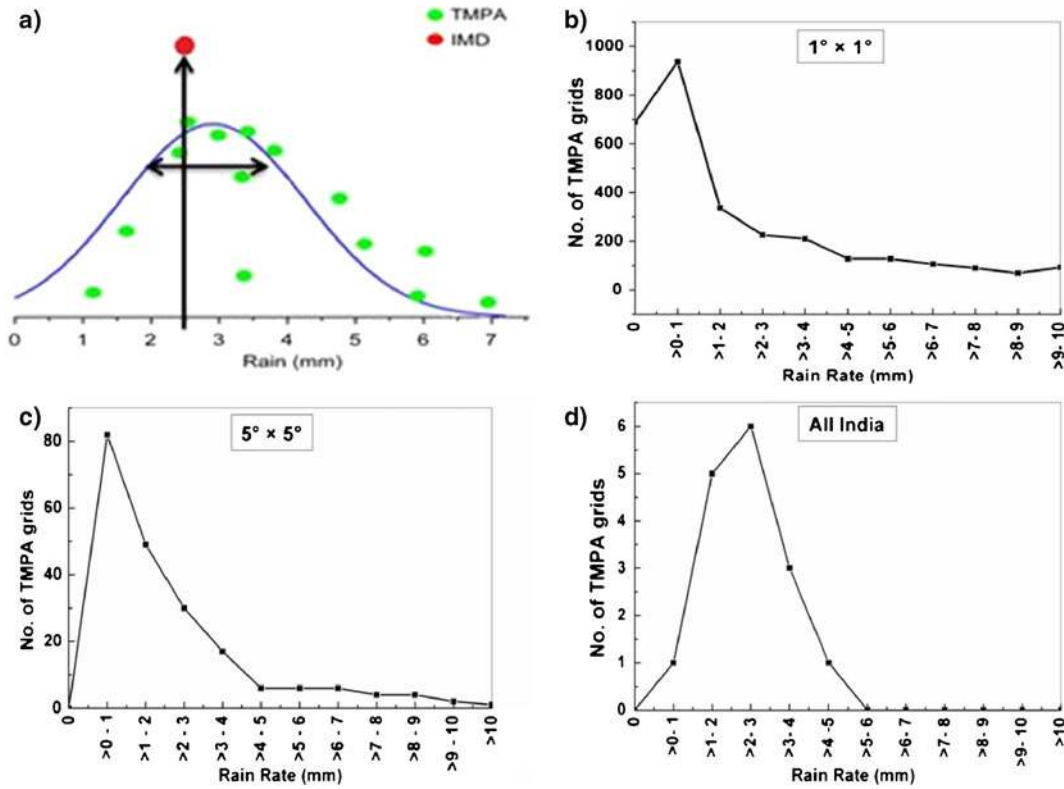


Figure 10. Distribution of number of TMPA grids for IMD rain rate of 2.5 mm/day. (a) Ideal case, (b) 1° x 1°, (c) 5° x 5° average and (d) all India average.

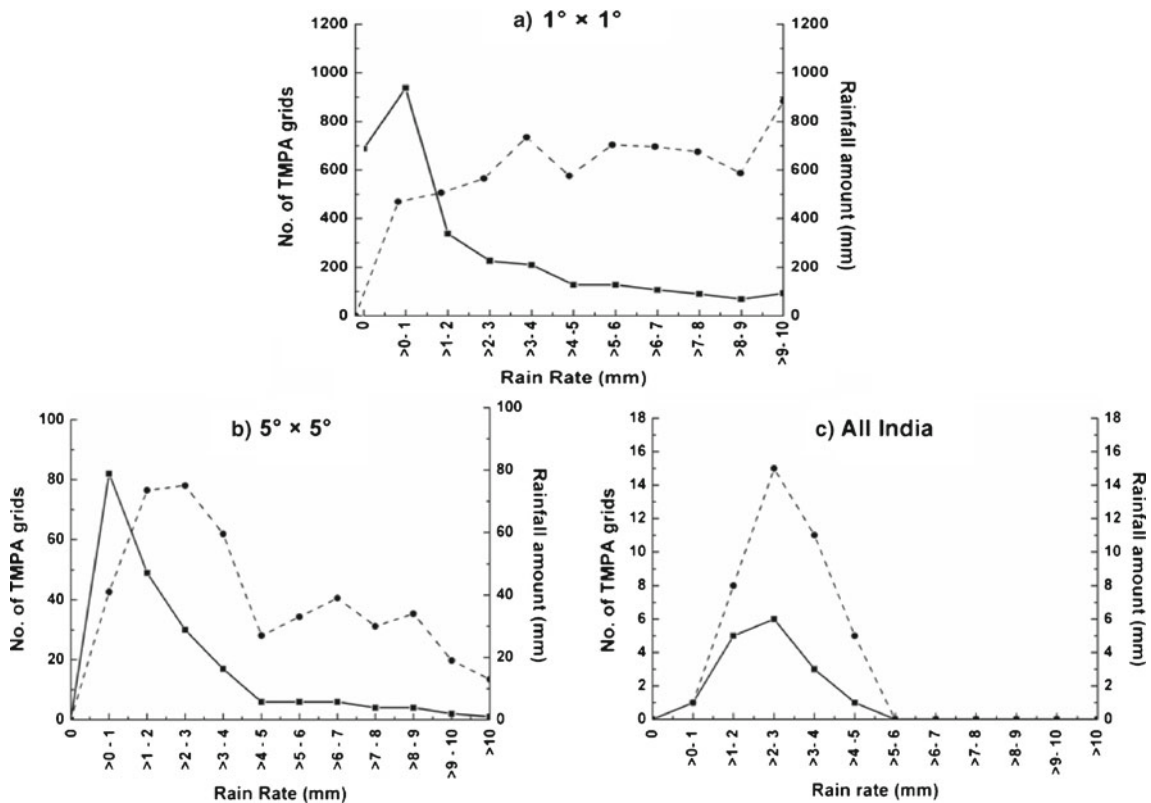


Figure 11. Distribution of number of TMPA grids (continuous curve) and associated rainfall (dashed curve) for IMD rain rate of 2.5 mm/day: (a) 1° x 1°, (b) 5° x 5° average and (c) all India average.

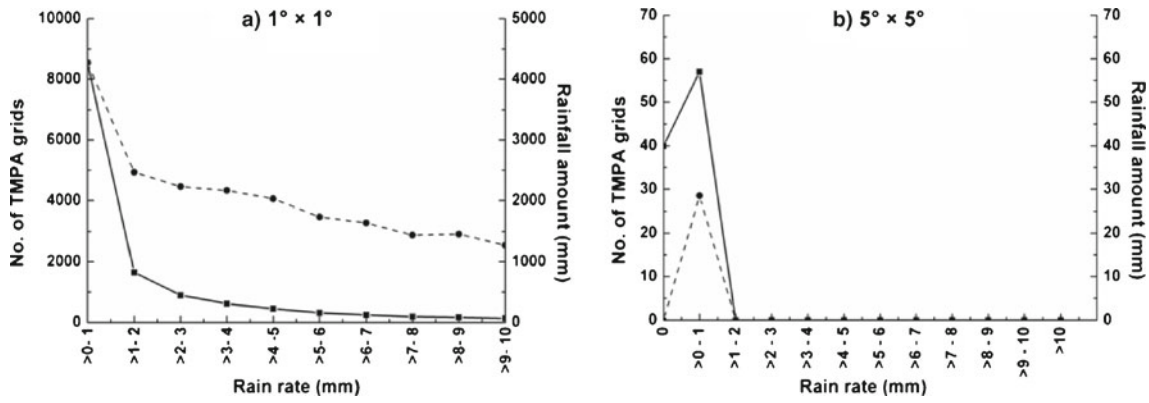


Figure 12. Distribution of number of TMPA grids (continuous curve) and associated rainfall (dashed curve) for IMD rain rate of 0 mm/day. (a) $1^\circ \times 1^\circ$, (b) $5^\circ \times 5^\circ$ average and (c) all India average.

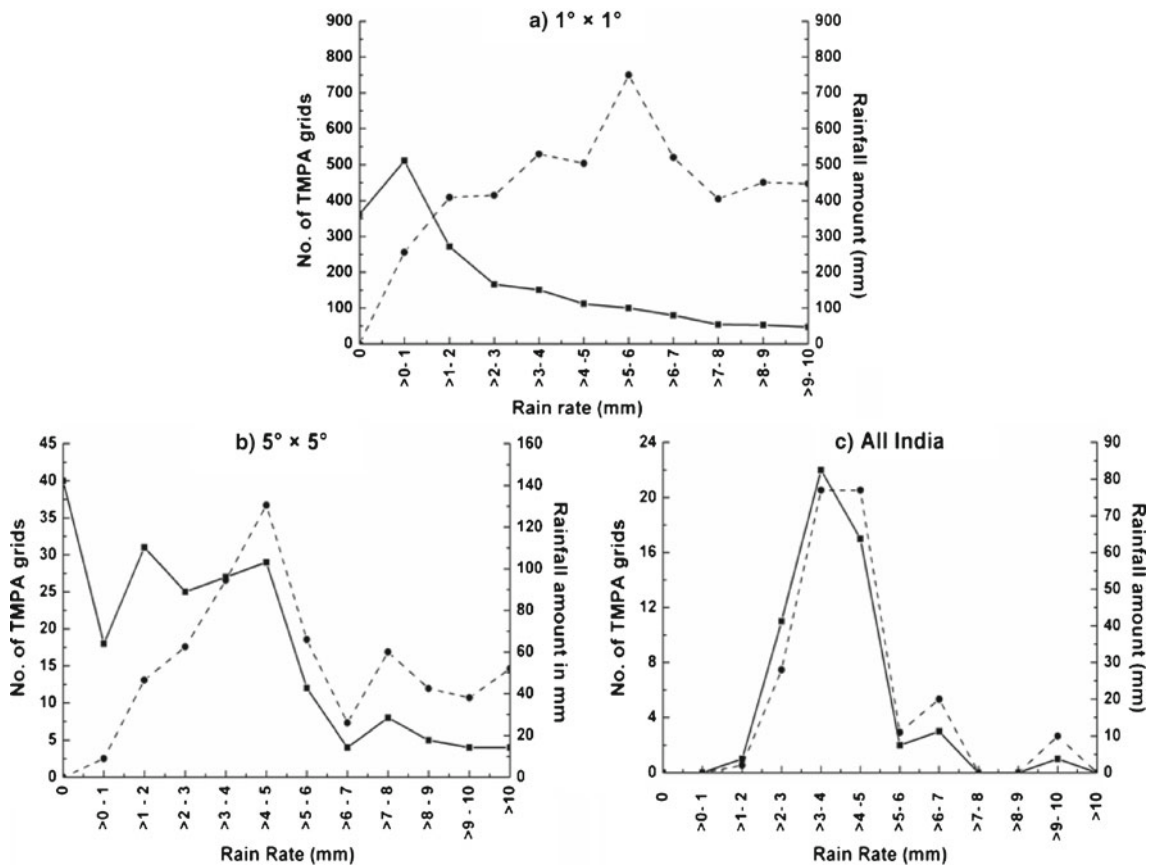


Figure 13. Distribution of number of TMPA grids (continuous curve) and associated rainfall (dashed curve) for IMD rain rate of 4.5 mm/day. (a) $1^\circ \times 1^\circ$, (b) $5^\circ \times 5^\circ$ average and (c) all India average.

To understand this behaviour, next, we considered spatial averaging of rain rates at $2^\circ \times 2^\circ$ grid size, and repeating the above procedure. This also did not result in anything significantly different. At $5^\circ \times 5^\circ$ spatial averaging, only a maximum of 7 grid boxes encompassing Indian plains could be constructed. This spatial scale averaging provided 2562 data points for 3 years (in contrast to $\sim 90,000$ grids at $1^\circ \times 1^\circ$ grid size). At the $5^\circ \times 5^\circ$ spatial average, we began seeing some

trends developing shifting peak (figure 10c), but still not quite satisfactory. We could not go for spatial scales beyond this scale, as more than one grid of size greater than $5^\circ \times 5^\circ$, representative of whole of Indian landmass could be constructed. Thus, the next and ultimate step was to go for all India spatial average. In the case of all India average, we had 366 data points for the statistical analysis. We started getting a peak in number of TMPA grid points around the IMD rainfall

window (figure 10d). Much to our satisfaction, the peak shifted for increasing/decreasing IMD rain rates and in the right direction (as illustrated further in figure 11–14 for other rain rate windows).

As already mentioned, the number of TMPA grids in other ranges of rain rates were very large for any specific IMD rain rate in the $1^\circ \times 1^\circ$ grids, and were large enough even for higher spatial averages. Even in cases of higher rain rates, where the number of grids did not appear to be significant, the contribution to total rainfall from them was substantial.

This led us then to evaluate the contribution of rainfall amounts from these grids (to the overall rainfall) outside the desired rain rate window. The amount of rainfall contributed by TMPA grids in each window was estimated by taking the product of number of grids in a particular window and their rainfall range. We expect the distribution of rainfall amount to be also Gaussian, with a peak around the corresponding IMD rainfall range.

Figure 11 also shows the rainfall contribution of TMPA grids in different windows corresponding to IMD rain rate of 2–3 mm. The total IMD grids in this window (figure 11a) as already mentioned, were 3625 and the corresponding rainfall amount was 9063 mm (taking a central value of

2.5 mm for the window). The TMPA rainfall contribution shows high values even beyond 10 mm rain rate window. Out of the 3625 IMD grids, only 226 TMPA grids estimated 2.5 mm of rainfall. Summing up the rainfall contribution of all the 3625 TMPA grids, gives a rainfall total of 14,333 mm, compared to IMD's 9063 mm (an overestimate by $\sim 55\%$).

At the spatial average of $5^\circ \times 5^\circ$, we are able to see a broad peak in the rainfall contribution around the IMD rainfall window of 2.5 mm along with a few smaller peaks at higher values of rainfall window (figure 11b). It may be noted that the peak in number of grids and peak in rainfall contribution are not lying at the same window at this spatial average. Only at the all India average spatial scale (figure 11c), a clear peak is observed at the rain rate window of 2.5 mm both for number of grids and their corresponding rainfall contribution. Also, there are very few grids from other windows and with minimal rainfall contribution.

In a similar way, the distribution of number of TMPA grids and their contribution to rainfall was constructed for all IMD rain rates from 0 to 10 mm in steps of 1 mm windows. Figures 12–14 show the results for some select IMD rain rate windows at 0, 4.5, and 8.5 mm (2.5 mm already shown in figure 9,

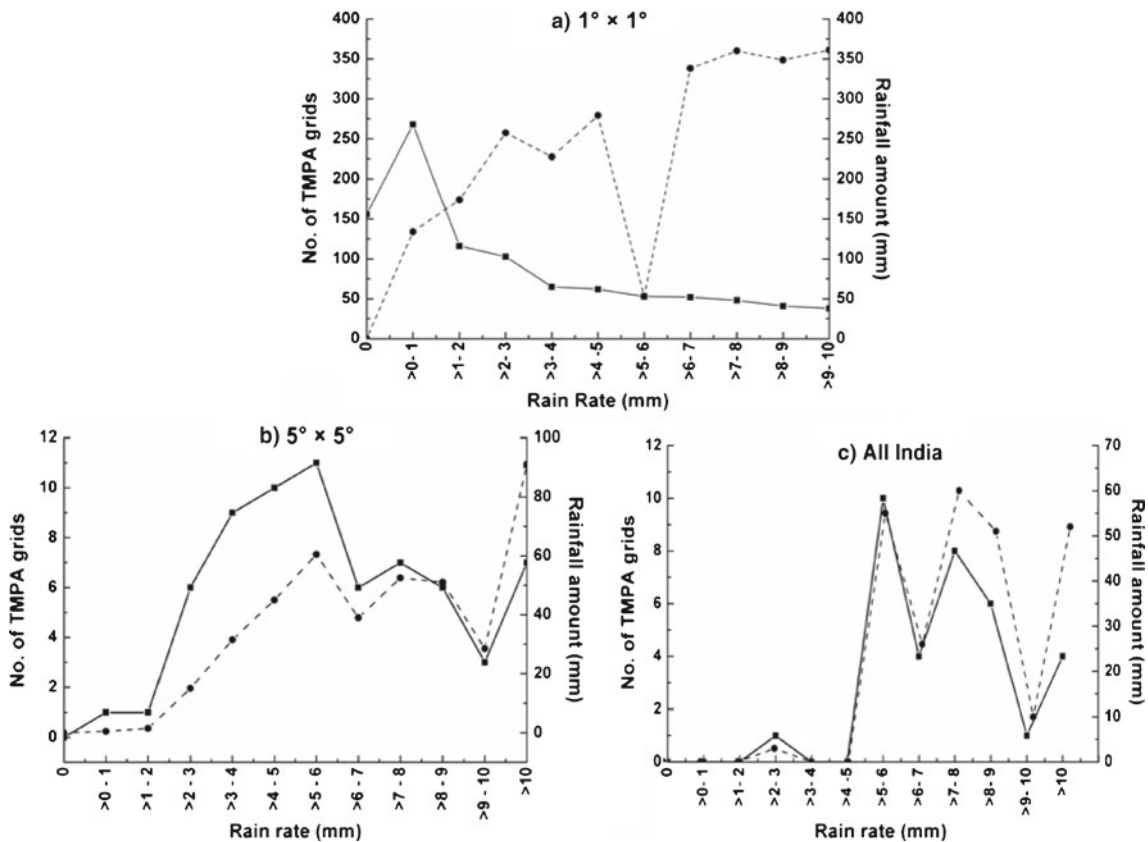


Figure 14. Distribution of number of TMPA grids (continuous curve) and associated rainfall (dashed curve) for IMD rain rate of 8.5 mm/day. (a) $1^\circ \times 1^\circ$, (b) $5^\circ \times 5^\circ$ average and (c) all India average.

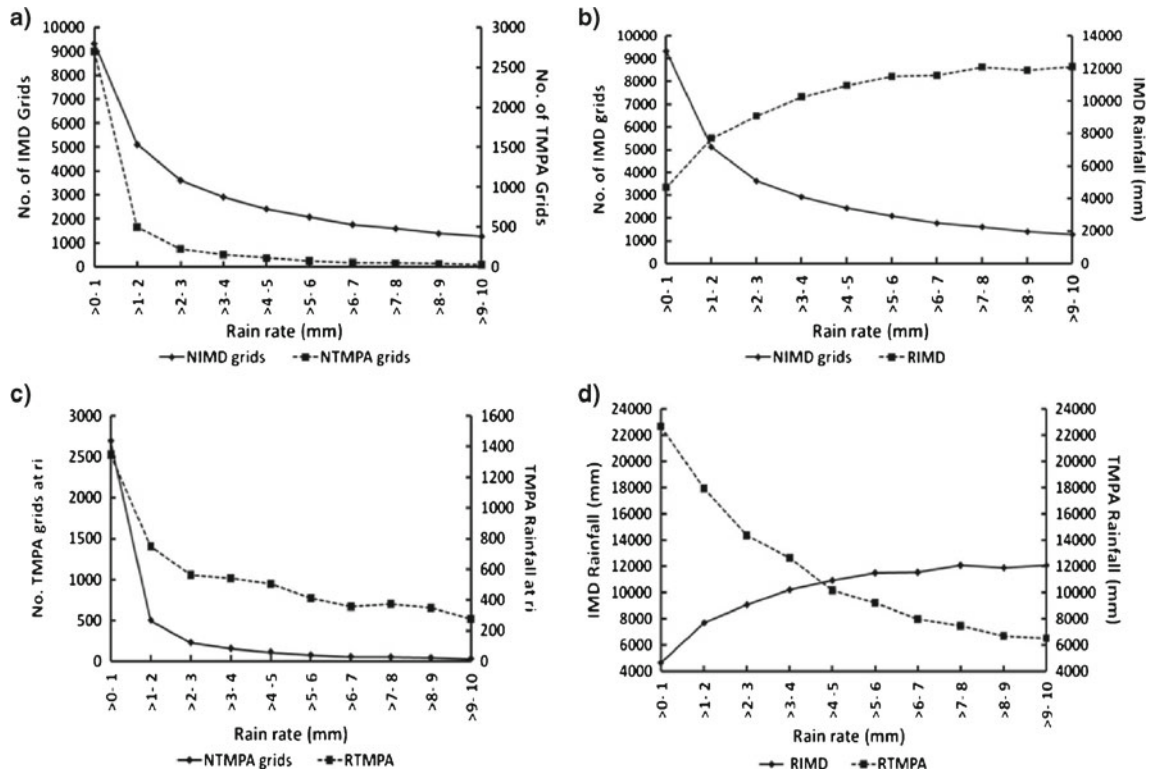


Figure 15. (a) Number of IMD (continuous curve) and TMPA (dashed curve) grids for different rain rate windows. (b) Number of IMD grids (continuous curve) and associated rainfall (dashed) for various rain rate windows. (c) Number of TMPA grids (continuous curve, subset of IMD grids) and associated rainfall (dashed curve) for various rain rate windows. (d) IMD (continuous curve) and TMPA rainfall (dashed curve, subset of IMD grids) for various rain rate windows.

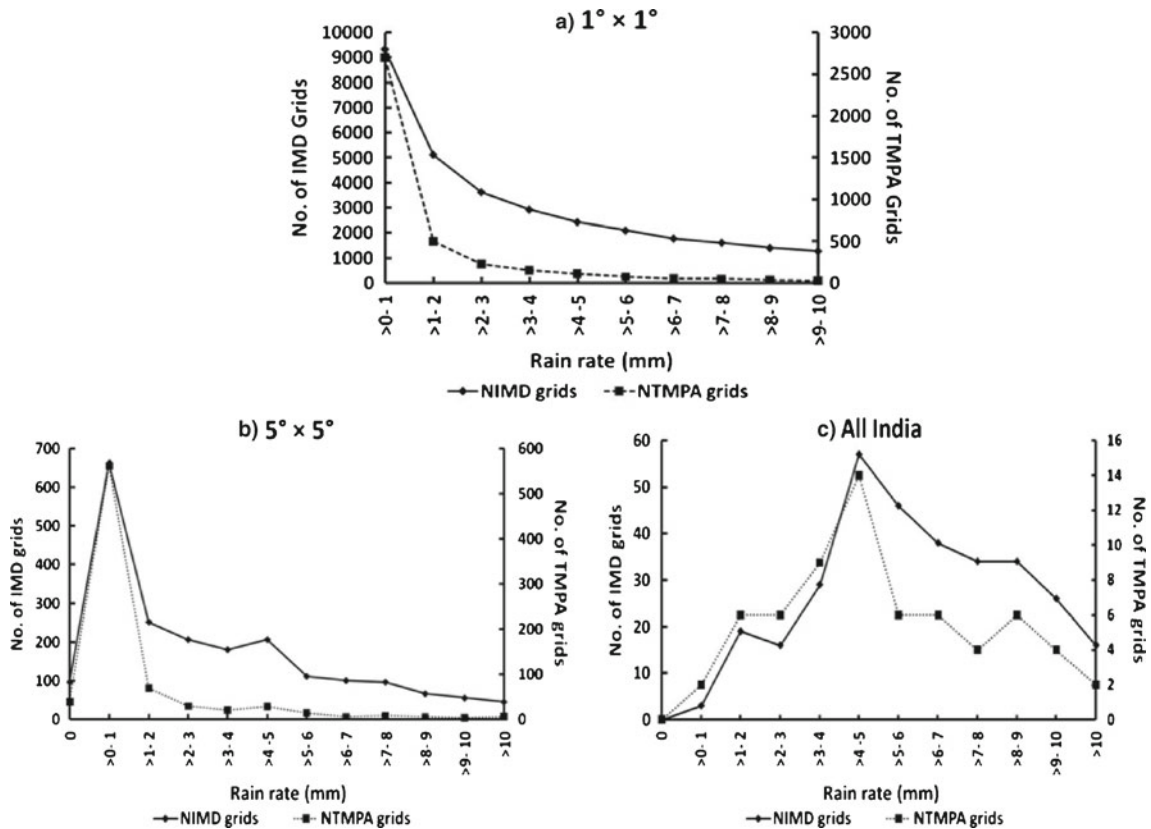


Figure 16. Number of IMD (continuous curve) and TMPA (dashed curve) grids for different rain rate windows. (a) $1^\circ \times 1^\circ$, (b) $5^\circ \times 5^\circ$ average and (c) all India average.

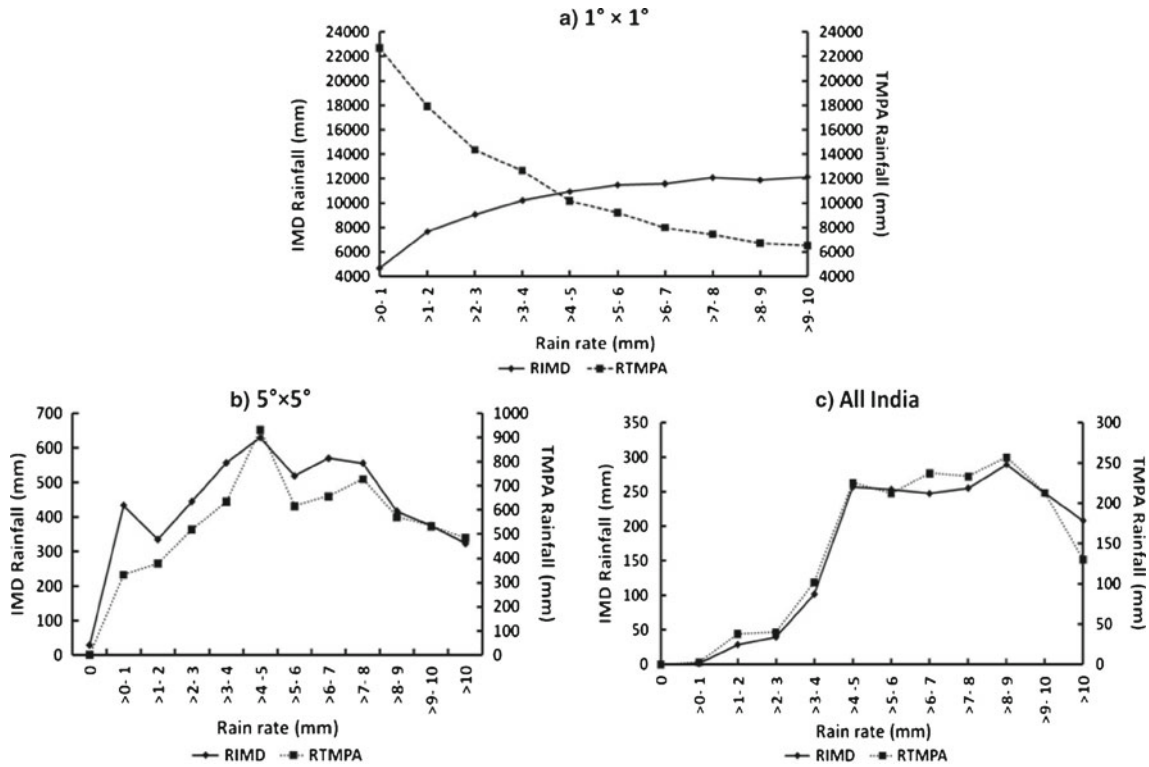


Figure 17. IMD (continuous curve) and TMPA (dashed curve) rainfall (subset of IMD grids) for various rain rate windows. (a) $1^\circ \times 1^\circ$, (b) $5^\circ \times 5^\circ$ average and (c) all India average.

results for other 7 windows not shown here). The results are shown for the three grid sizes $1^\circ \times 1^\circ$, $5^\circ \times 5^\circ$ and all India average. Results observed for these windows are much similar to what has been seen for the case of 2.5 mm window.

At rainfall window of 0 mm (figure 12a), there are around 40,000 IMD grids ($\sim 45\%$ of the total number of grids) contributing to 'zero' rainfall. On the other hand, there are only around 25,000 TMPA grids out of the above IMD grids with zero rainfall. Other TMPA grids ($\sim 15,000$) show varying amounts of rainfall, their total rainfall contribution working out to be a huge 40,000 mm. In comparison, the total rainfall of all the 90,000 TMPA grids is $\sim 287,000$ mm. This would mean that we will be making an error of $\sim 14\%$ at the 0 mm window itself using the TMPA rainfall values.

The discrepancies related to contributions from other TMPA grids for a specific IMD rain rate have been further illustrated in figure 15. Figure 15(a) shows the number of IMD grids for different ranges of rain rates. The number of TMPA grids (which are a subset of the above number of IMD grids) are shown by the dashed curve.

Figure 15(b) shows the number of IMD grids and their contribution to the total rainfall in different windows which shows an increasing behaviour with increasing rain rate, reaching a plateau around 7–8 mm. Figure 15(c) shows the number of TMPA

grids (subset of IMD grids) and their associated rainfall. This shows a decreasing rainfall trend with increasing rain rate in contrast to IMD rain behaviour.

Figure 15(d) shows the comparison of the contribution of IMD and TMPA grids at various ranges of rain rates (i.e., those TMPA grids which are a subset of IMD grids). The total rainfall from these TMPA grids are 100,700 mm (out of the total TMPA rainfall of $\sim 270,000$ mm). This would mean that only about 30% of grids are contributing to the appropriate IMD rain rate windows.

The above analysis (for $1^\circ \times 1^\circ$ grid) has been carried out for other spatial averages also. Figures 16 and 17(a, b and c) summarise the results at the three grid averages. At the $5^\circ \times 5^\circ$ scale (figures 16b and 17b), there are some indications of emerging match between the two datasets. However, we see almost a perfect match of the two rainfall amounts at all rain rate ranges only at all India average scale (figure 17c). However, there are some differences in the number of grids even at this scale (figure 16c).

5. Conclusions

The TMPA satellite based global rain products are available in public domain at very high spatio-temporal resolution ($0.25^\circ \times 0.25^\circ$ spatial

grid and 3 hourly temporal). The TMPA products are being used at these high resolutions to compare with other rainfall results and model forecasts, both over land and oceans – taking these products to be ground truth at such fine resolutions. At this juncture, it is possible to make a detailed comparison between TMPA data and conventional observations only over land (in the Indian monsoonal region), the gauge observations over oceans being very sparse. This has also become possible only since 2006 with the availability of IMD gridded data products at daily $1^\circ \times 1^\circ$ scale from the large rain gauge net work.

TMPA and IMD gridded data at daily $1^\circ \times 1^\circ$ spatial scale have been found to be mutually in much disagreement from the analysis of 10 years datasets. However, on much larger spatial scales (viz., at all India average scale) the match seems satisfactory, albeit with quantitative differences. The spatial distribution of 10 years climatology exhibits prominent known features of southwest monsoon, as also some flood/drought features. The correlation between the two datasets is higher during cumulative wet and dry spells compared to those during individual years. The wavelet analysis exhibit clearly major ISO modes of 10–20 and 30–50 days, with the power being about 40% in the low frequency component. TMPA data show a negative bias in southwest monsoon and a positive bias during northeast monsoon on all India scale *vis-a-vis* IMD gridded data. This may have implications to setting different thresholds for the empirical rain retrieval from satellite infrared data over the Indian landmass for the two monsoons.

It has been attempted here to arrive through a detailed analysis, the minimum spatial scale (at daily temporal scale) at which the TMPA data can be considered a proxy/representative ground truth over the Indian landmass. It could be different over different climate regimes of the globe, and over the oceans.

For arriving at this minimum scale of compatibility of TMPA and IMD gridded data, we used a large database of three years over the Indian landmass. Both datasets were arranged according to location, date and rainfall rate in 1 mm windows. By systematically increasing the grid area from $1^\circ \times 1^\circ$ in steps of $1^\circ \times 1^\circ$, appropriate match was arrived at by observing the distribution of TMPA grids (and associated total rainfall amounts) with respect to IMD rainfall rate.

Some semblance of match is seen at $5^\circ \times 5^\circ$ grid averages, and a good match can be seen at all-India average scale. More than one grid of size larger than $5^\circ \times 5^\circ$ representative of India (encompassing the Indian landmass) could not be constructed because of the geometry of the Indian land boundary. It is interesting, however, to note that Arkin *et al.*

(1989) have shown good correspondence between INSAT infrared derived rainfall and ground truth at $2.5^\circ \times 2.5^\circ$ scale, but only on monthly time scale, on the basis of which the rainfall from INSAT are derived on an operational basis.

Thus, it may be concluded that only at all India spatial scale the TMPA comparisons with IMD rainfall are valid with good reliability, except for quantitative differences.

Acknowledgements

The authors are thankful to Dr B V Krishnamurthy and Prof. D Narayana Rao for useful discussions during the course of this work. The financial support received from ISRO RESPOND Programme is gratefully acknowledged. TMPA data were obtained from NASA website.

References

- Adeyewa Z D and Nakamura K 2003 Validation of TRMM radar rainfall data over major climatic regions in major climatic regions in Africa; *J. Appl. Meteorol.* **42**(2) 331–347.
- Adler R F, Huffman G J, Gu G, Curtis S and Hong Y 2007 The TRMM multisatellite analysis (TMPA): Quasi-global, multiyear combined-sensor precipitation estimates at fine scales; *J. Hydrometeorol.* **8** 38–55.
- Annual Climate Summary 2007, India Meteorological Department, Pune.
- Arkin P A, Kishna Rao A V R and Kelkar R R 1989 Large scale precipitation and outgoing longwave radiation from INSAT-1B during the 1986 southwest monsoon season; *J. Climate* **2** 619–628.
- Barrett E C and Martin D W 1981 *The Use of Satellites in Rainfall Monitoring*; Academic Press, London.
- Chen T C and Chen J M 1993 The 10–20 day mode of 1979 Indian monsoon: Its relation with the time variation of monsoon rainfall; *Mon. Weather Rev.* **121** 2465–2482.
- Chokengamwong and Chiu 2005 TRMM and Thailand Daily Gauge Rainfall Comparison; Centre for Earth Observing and Space Research, George Mason University.
- Dakshinamurthy J and Keshavamurthy R N 1976 On the oscillations of the period one month in the Indian summer monsoon; *Indian J. Meteor. Geophys.* **27** 201–203.
- Gadgil S, Rajeevan M and Nanjundiah R 2005 Monsoon prediction – Why yet another failure? *Curr. Sci.* **88** 1389–1400.
- Gadgil S P, Vinayachandran N and Francis P A 2003 Droughts of Indian summer monsoon: Role of clouds over the Indian Ocean; *Curr. Sci.* **85** 1713–1719.
- Goswami B N and Ajaya Mohan R S 2001 Intraseasonal oscillations and interannual variability of the Indian summer monsoon; *J. Climate* **14** 1180–1198.
- Islam M N, Terao T, Uyeda H, Hayashi T and Kikuchi K 2005 Spatial and temporal variations of precipitation in and around Bangladesh; *J. Meteorol. Soc. Japan* **83**(1) 23–41.
- Kalsi S R, Hatwar H R, Jayanthi N, Subramanian S K, Shyamala B, Rajeeman M and Jenamani R K 2004 Various aspects of unusual behaviour of monsoon 2002 IMD; *Met. Monograph Syn. Met.* **2** 105.

- Kidder S Q and Vonder Haar T H 1995 *Satellite Meteorology: An Introduction*; Academic Press, 466p.
- Krishnamurti T N and Bhalme H N 1976 Oscillations of monsoon system. Part I: Observational aspects; *J. Atmos. Sci.* **45** 1937–1954.
- Krishnamurti T N and Ardunay P 1980 The 10–20 day westward propagating mode and breaks in the monsoons; *Tellus* **32** 15–26.
- Mishra A, Gairola R M, Varma A K and Agarwal V K 2010 Remote sensing of precipitation over Indian land and oceanic regions by synergistic use of multisatellite sensors; *J. Geophys. Res.* **115** D08106, doi: [10.1029/2009JD012157](https://doi.org/10.1029/2009JD012157).
- Narayanan M S, Shah S, Kishtawal C M, Sathiyamoorthy V, Rajeevan M and Kriplani R H 2005 Validation of TRMM merge daily rainfall with IMD raingauge analysis over Indian landmass; Technical Report, Space Application Centre, Ahmedabad, India.
- Rahman S H, Sengupta D and Ravichandran M 2009 Variability of Indian summer monsoon rainfall in daily data from gauge and satellite; *J. Geophys. Res.* **114** D17113, doi: [10.1029/2008JD011694](https://doi.org/10.1029/2008JD011694).
- Rajeevan M, Jyoti Bhate, Kale J D and Lal B 2006 High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells; *Curr. Sci.* **91(3)** 296–306.
- Rao P K, Holmes S J, Anderson R K, Winston J S and Lehr P E (eds) 1990 *Weather satellites: Systems, data, and environmental applications*; American Meteorological Society, Boston.
- Rao T N, Uma K N, Mohan Satyanarayana T and Narayana Rao D 2009 Differences in draft core statistics from the wet to dry spell over Gadanki, India (13.58N, 79.28E); *Am. Meteorol. Soc.* **137** 4293–4206.
- Southwest Monsoon 2002 *End of Season Report*; India Meteorological Department, Pune.
- Torrence C and Compo G P 1998 A practical guide to wavelet analysis; *Bull. Am. Meteorol. Soc.* **79** 61–78.
- Xie P and Liu C *et al.* 2007 A gauge-based analysis of daily precipitation over East Asia; *J. Hydrometeorol.* **8** 607–626.
- Yansunari T 1980 A quasi-stationary appearance of 30–40 day period in the cloudiness fluctuation during summer monsoon over India; *J. Meteorol. Soc. Japan* **58** 225–229.

MS received 28 September 2012; revised 10 January 2013; accepted 17 January 2013