OPEN

Assessing reliability of regional climate projections: the case of Indian monsoon

K. V. Ramesh & Prashant Goswami

CSIR Centre for Mathematical Modelling and Computer Simulation, Wind Tunnel Road, Bangalore-560037, Karnataka, India.

Projections of climate change are emerging to play major roles in many applications. However, assessing reliability of climate change projections, especially at regional scales, remains a major challenge. An important question is the degree of progress made since the earlier IPCC simulations (CMIP3) to the latest, recently completed CMIP5. We consider the continental Indian monsoon as an example and apply a hierarchical approach for assessing reliability, using the accuracy in simulating the historical trend as the primary criterion. While the scope has increased in CMIP5, there is essentially no improvement in skill in projections since CMIP3 in terms of reliability (confidence). Thus, it may be necessary to consider acceptable models for specific assessment rather than simple ensemble. Analysis of climate indices shows that in both CMIP5 and CMIP3 certain common processes at large and regional scales as well as slow timescales are associated with successful simulation of trend and mean.

Climate projections are emerging as critical inputs for many applications and decision support. However, assessment of reliability of climate projections remains a major challenge; the challenge is greater at regional scales. At the same time, accurate projections of regional climate systems, like the Continental Indian Monsoon (CIM), are critical for assessing the sustainability of a large section of the world's population and to determine the future of the global climate system. However, significant uncertainties still exist regarding the reliability of the projections from the dynamical climate models, especially for regional systems like the south Asian summer monson[1–3]. Conceptually, any assessment of reliability of projections has to be based on their present-day (current-era) performance. The climate models today generally possess superior skill in seasonal forecasting[4] than statistical models, and can be evaluated in hindcast mode; however, evaluation of reliability of climate projection will remain a challenge, especially over areas with contribution from natural climate variability[5], such as the Indian monsoon[6]. In spite of their common basic mechanism, monsoons over different regions are subject to diverse forcings[7]; thus the responses of the various regional monsoons to a changing climate are expected to vary[2,3]. Estimates of impacts from anthropogenic climate change rely on projections from climate models. Uncertainties in the climate projections are a strong limiting factor in estimation of impacts and hence policy design, especially at regional scales.

An important issue in our approach is to assess climate projections based on the accuracy of trends for the current climate. Many climate-modelling groups around the world have participated in the Coupled Model Inter-comparison Project phase 5 (CMIP5). Under CMIP5, a series of simulations including the twentieth century historical simulation and the twenty-first century climate projections with four different representative concentration pathway (RCP) scenarios were performed[8]; most CMIP5 models include both direct and indirect effects of aerosols. In the earlier phase (CMIP3) also twentieth century simulations and twenty-first century climate projections were carried out with three different climate scenarios[9]. Several works report that the CMIP5 multi-model mean is more skillful than that from the CMIP3 models[10]. The CMIP5 models simulated mean precipitation varies from 500 mm to 900 mm and coefficient of variation from 3 to 13%[11]. Studies have reported an increase of global monsoon area and precipitation intensity under the RCP4.5 scenario of CMIP5[12,13].

An evaluation of the Indian summer monsoon rainfall for the period 1850 to 2100 in 20 CMIP5 models showed[14] a consistent increase in summer monsoon mean rainfall; essentially all models were reported to simulate stronger seasonal mean rainfall in the future compared to the historic period under different RCP scenarios, with the highest increase under strongest warming scenario RCP8.5. However, a consensus among models does not necessarily imply reliability of the projections. Similarly, analysis of regional monsoonal rainfall and their changes in the 21st century under RCP4.5 and RCP8.5 scenarios from 29 CMIP5 climate models showed[7] that the global monsoon precipitation intensity and the global monsoon total precipitation are also projected to increase. However, as noted[7], the limited ability of the models to reproduce the current monsoon climate, along with the large scatter among the simulations, imply only low confidence in the projections.

1

An evaluation of simulations of Asian summer monsoon from 25 CMIP5 and 22 CMIP3 simulations of the late twentieth Century was reported in terms of time-mean, climatological annual cycle, inter-annual variability, and intraseasonal variability[10]. However, no comparative evaluation in terms of trends was reported. Besides, as noted earlier, the characteristics for CIM can be very different from those for a larger system like the Asian monsoon. Certain improvements, such as in simulation of seasonality, have been reported in CMIP5 over CMIP3[15], with most CMIP5 models correctly simulating very low rainfall rates outside of the monsoon season.

An analysis of expected future pulse of the Indian monsoon climate based on observational and CMIP3 projections was used to claim[16] that the Indian monsoon rainfall in the latter half of the 21th century may be very similar to the current monsoon in terms of amount. However, the analysis was based on only CMIP3 simulations; nor were trends, nor any other measure, considered to ascertain reliability (or confidence). It has been already noted that projected global temperature change in CMIP5 is remarkably similar to that from those from CMIP3[17]; in spite of substantial effort in model development and improved computational capacity, there has been no significant change in the local model spread.

It is known that the trends in rainfall over the continental India are quite different (often opposite) to those over the neighboring oceans[18,19]; thus, only consideration of the continental rainfall can provide meaningful inputs to assessment and planning of sustainability[20]. Most of the models now simulate the annual cycle of rainfall over India fairly accurately; the spread in the simulations[20–26] is not unacceptably high. However, these results do not necessarily guarantee accurate simulation of the trends for the recent past, and hence do not imply reliability of the projections.

As noted earlier, quantification and assessment of reliability of any projection is a challenge as the projection cannot be immediately verified like a short-term forecast. This challenge is particularly greater for climate projections due to the presence of (non-linear) trends. In particular, variation of trends in different epochs due to low-frequency variability introduces additional uncertainties. However, it is reasonable and logical to expect that the model simulations reproduce current trends with sufficient accuracy; this argument can be further strengthened through examination of consistency between cross-epochal trends. For a quantitative measure of reliability of projection, we adopt accuracy in the simulation of the historical trends. We have adopted a hierarchical approach to identify reliability of trends, based on increasingly demanding metrics, than on a single parameter like absolute error. An objective here is to assess the reliability of the projections at regional scale relevant for application, specifically to CIM. To avoid any ambiguity due to the selection of geographical coverage as well as to ensure relevance of the results for application, we consider only continental Indian monsoon (CIM) rainfall (June-September); this also allows a robust analysis with multiple sets of observation (Table 1). Indian Institute of tropical Meteorology[20] (IITM) and Indian Meteorological Department[21] (IMD) provide the all India summer monsoon rainfall index. The gridded data averaged over continental India for rainfall from different source like Indian Meteorological Department[22] (IMDG), Climate Research Unit[23] (CRU), Global Precipitation Climatology Project[24] (GPCP), Asian Precipitation - Highly-Resolved Observational Data[25] (APHRO) and National Centers for Environmental Prediction[26] (NCEP) are used to analyze differences between observations.

Historical precipitation records over the monsoon regions around the globe reveal a decreasing trend in the global land monsoon precipitation[27,28] over the last half of the century (1948–2003), with primary contributions from a weakening of the summer monsoon systems in the Northern Hemisphere (NH). When the oceanic monsoon rainfall is combined with the land monsoon, the global monsoon precipitation is found to have increased for the 1979–2008 epoch, mainly due to an increase in the NH summer monsoon precipitation[13,28]. Thus the future trajectory for a given monsoon cannot be in general inferred from the studies of the other monsoon systems.

One of the possible sources of error (difference) in simulations is the model's ability to simulate various regional and large-scale climate processes. Both CMIP3 and CMIP5 simulations exhibit large spreads in simulations of average monsoon rainfall and their inter-annual variability[29], although the multi model ensemble mean monsoon rainfall is found within the observational uncertainty. In terms of the seasonal cycle of rainfall, the CMIP5 models were reported[29] to generate relatively more realistic features than CMIP3. However, these analyses did not involve analysis of trends, and thus cannot be used to assess reliability of projections in our framework.

An analysis of the Indian summer monsoon-ENSO relationship showed over persistence of ENSO events in many CMIP models, while the relation between Indian Ocean Dipole (IOD) and the Indian monsoon-IOD was not found to be significant in the simulations[29]. These analyses were used as a basis for a methodology for selecting (12) "best" models to analyze projections in the RCP8.5 scenario[29]. However, the objective criteria did not include the quality of the simulation of the trends.

There have been several works on evaluation of CMIP3 projections in different contexts[30], and sometimes with limited number of models[31]. It is generally noted that the confidence in the projections of rainfall is lower than that in temperature[30,31]. The similarity in the projections of global temperature changes from CMIP3 and CMIP5 models was also noted[16]. However, rarely the question of reliability of the projections (in terms of accuracy in simulating current trends) has been addressed; at the same time, as noted earlier quality of simulation of mean and variability does not imply accurate simulation of trends. We address this issue with a hierarchical evaluation. An equally important question addressed here is whether, and how much, progress has been achieved in CMIP5 in projecting regional systems like CIM; this question has not been unambiguously

Table 1 | List of observed datasets with symbols used, along with the climatological (1951–2005) mean, standard deviation, linear trends and their significance levels for seasonal (June-September) rainfall over continental India (land only: 70-85E, 5-30N). Cases with insignificance trends (<90%) are highlighted

Symbol	Source/Name	Mean (mm)			Standard deviation (% of mean)			Trend (mm/year)		
		1951–2005	1951–1975	1975–2005	1951–2005	1951–1975	1975–2005	1951–2005	1951–1975	1975–2005
O_1	IITM	892	906	884	10.6	11.4	9.9	−1.1	−0.9	−1.9
O_2	CRU	773	797	797	12.7	13.3	11.8	−1.8	−2.3	−3.7
O_3	GPCP	839	846	846	12.0	13.1	11.3	**−0.8**	−2.0	−3.2
O_4	IMD	892	906	884	10.6	11.4	9.9	−1.1	−0.9	−1.9
O_5	IMDG	851	867	867	12.8	13.6	12.1	−1.3	−2.6	−3.8
O_6	APHRO	726	736	736	12.9	14.0	12.1	−1.2	−2.5	−3.8
O_7	NCEP	772	795	757	12.5	13.4	11.3	−2.3	−4.5	−5.6
Composite (all)		821	836	824	12.0	12.9	11.2	−1.4	−2.2	−3.4

addressed due to arbitrary choice of monsoon domains in many studies. The results are expected to identify methodology and directions for more reliable, and applicable[5] climate projections.

Results

The assessment of the simulated trends is made further complex by the fact that the observed trend itself has a degree of uncertainty; different observed data sets show appreciable differences among them (Table 1). The differences in the basic statistical quantities like the mean and the standard deviation in different observations are quite high (Table 1). To avoid any bias in the evaluation, we have considered six data sets from different sources (Table 1). Further, a multi-epochal analysis (Table 1) was carried out to examine variation in trends due to low-frequency variability for three separate epochs: 1951–2005, 1951–1975 and 1976–2005. We have also considered an ensemble formed through an equal-weight averaging for both observations and simulations. The epochal trends for 1951–2005 and 1976–2005 as % of standard deviation of the respective period are of the same signs (Table 1); this suggests that success in simulation of trend in an earlier epoch is indicative of skill for future projection. It is worth noting that all the seven observations show similar (negative) trends (Table 1). Although the trend for 1951–2005 in GPCP is negative but significant only at 86% level; all other trends are negative and significant at more than 90%. In what follows, unless otherwise mentioned, we shall consider the composite observation for evaluation of the simulations. For gaining insight, we consider three (equal-weight average) ensembles based on the levels of the statistical significance of the trends: High significance (Probability, P < 0.01), Low significance (P < 0.05) and all-average observations. For model simulations, such as trends and correlation coefficients, we shall adopt a value P < 0.2 as the acceptable level of significance in view of inherent uncertainties in simulations. The trends for three epochs had been used to examine and establish the consistency of the analysis; however, only long-term (1951–2005) trends are used to evaluate the CIM simulations.

The results and the conclusions regarding monsoon can also change depending on the selection of the domain[14,15]. To ensure robustness of our results, we have considered five domains that encompass relatively small case of CIM as well as larger domain that

also includes parts of Indian Ocean (Fig. 1). As expected, there are small differences among the seasonal mean rainfall and the trends for different domains even in observations (Fig. 1); however, the values are generally consistent. We note that for D5 (continental India plus ocean) the trends are opposite for 1951–2005 and 1951–1975; however, this could be a manifestation of low-frequency variability. For our discussion, it is important to note that both CMIP3 and CMIP5 (all) ensemble reproduce the current (1976–2005) trends within the margin of acceptability (Fig. 1).

It has been argued that an ensemble of all simulations may be the best option as performance varies with metric[30]. In contrast, we look for the best model(s) for a given metric (application). A highlight of our methodology is a hierarchical evaluation in which certain criteria (metrics) are considered more important than the others in evaluating model performances. We organize our hierarchy of criteria (in terms of increasing constraint) as follow:

(a) Higher/lower (more/less than 2σ) mean & negative trend (not necessarily significant)

(b) Comparable (in between ± 2σ) mean & negative trend (not necessarily significant)

(c) Higher/lower (more/less than 2σ) mean & acceptable (significant within ± 10% observed) trend

(d) Comparable (in between ± 2σ) mean & acceptable (significant within ± 10% observed) trend

Where σ is defined as the dispersion in observation is defined as the difference between maximum and minimum value divided by 2. A comparison of the trends of the continental India (CIM: D3, 70-85E, 5-30N, Fig. 2a) and CIM plus ocean (D5:60-94E, 10S-30N, Fig. 2b) for the period 1951–2005 shows that all the seven observations show significant negative trends for CIM seasonal rainfall (Fig. 2a, middle panels) but positive trends for the larger domains (Fig. 2b, middle panels). Thus the decreasing trend in the seasonal rainfall is highly regional effect, as also earlier noted[15]. These negative trends are well captured by some of the simulations in both the CMIP5 (Fig. 2a, left panel) and CMIP3 (Fig. 2a, right panels). The symbols A-U/a-x used to represent the individual CMIP3/CMIP5 climate model simulations as described in Table S1. However, while all-member ensemble shows show insignificant negative trends as
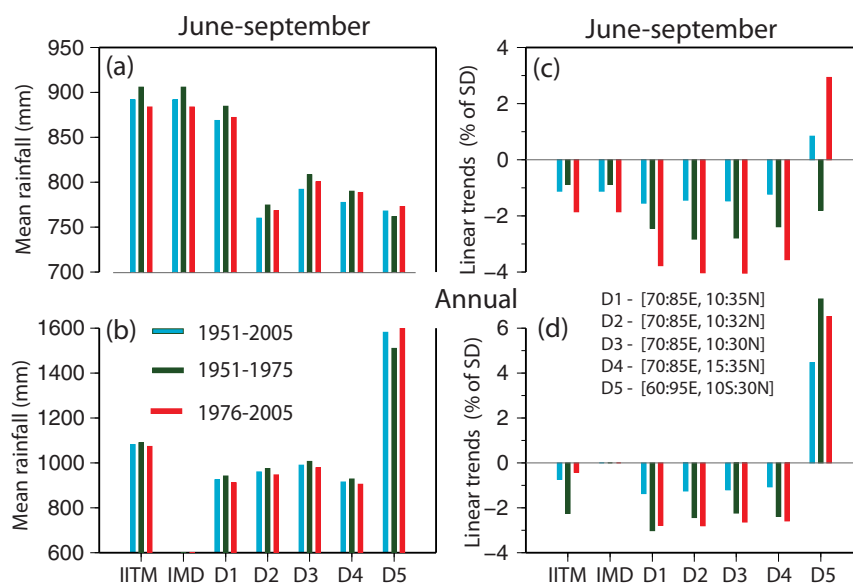

Figure 1 | Seasonal (June-September) observed (composite of multiple observations) mean rainfall and the trends over different regions in the summer monsoon region are compared with the simulated mean and trends from CMIP3 and CMIP5 models. D1–D4 is the rainfall average over continental India and D5 shows the rainfall over the continental India plus ocean. The region D2–D4 clearly shows the observed mean rainfall and the trends over continental India don't show much variation, as they indicate the core monsoon region.
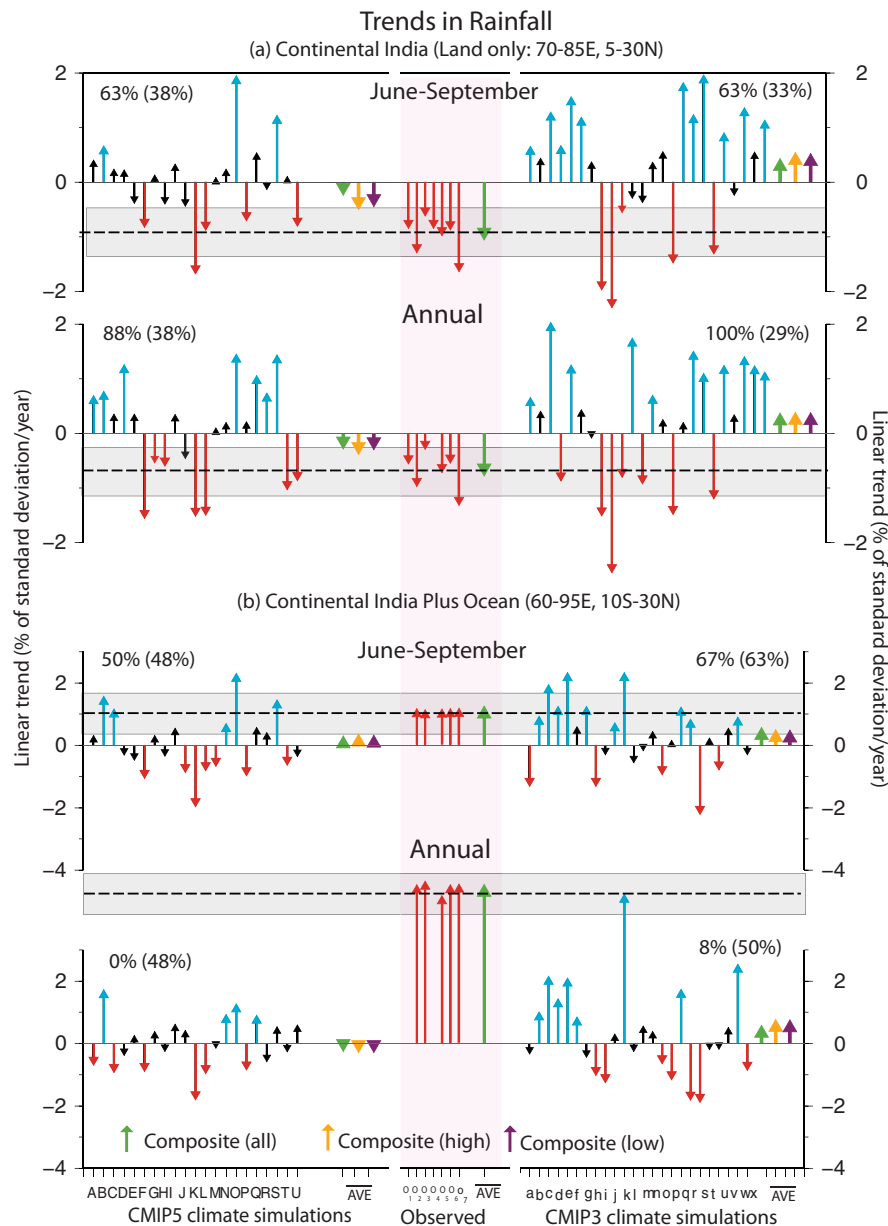
Figure 2 | Trends in the seasonal (June-September) and the annual rainfall over continental India (CIM: D2: 70-85E, 10-30N) and continental India plus ocean (D5: 60-95E, 10S-30N) from CMIP5 (left panels), CMIP3 (right panels) climate model simulations compared with the trends in the multiple observations (middle panel) for the period (1951–2005); the trends are expressed as % of respective standard deviation for the period. For the observations the trend in the different composites (highlighted, yellow) represent average of all (green), high significance (P < 0.01, orange) and low significance (P < 0.05, purple) ensembles. The percentage of models that simulated significant negative trend as against the total number of significant trends (positive and negative) is shown in each panel; the numbers in the brackets show the % of the total simulations with negative trend (as observed) in the respective case. The blue, red and black lines indicate, respectively, significant (P < 0.2) positive trend, significant (P < 0.2) negative trend and insignificant (P > 0.2) trends. The dash line indicates the observed composite trend and the grey band shows the dispersion (1 σ) in the observed trends.

observed, the CMIP3 ensembles show opposite trends (Fig. 2a right panels), essentially due to a few simulations with large positive trends. In terms of the larger domain, both CMIP5 and CMIP3 ensembles show positive trends as observed, but not significant as the observed trends. For the CMIP5 models, however, Fig. 2a (top left) shows that, out of the 21 models considered, only nine reproduce the negative trend as observed; only five of these trends (Fig. 2a, top left panel) are of statistical significance (P < 0.2) and comparable to the observed trend (Fig. 2a, top middle, P < 0.05). In particular, none of the CMIP5 ensembles (thick arrows, Fig. 2a) satisfies the criterion of even negative trend. The result is no different for the annual

rainfall; thus, the lack of skill for CIM rainfall cannot be attributed to shifts in the seasonal rainfall in the simulations (Fig. 2a, bottom left).

In the next step of hierarchical assessment, we have considered the simulations with significant (P < 0.2) negative trends in CIM rainfall (Fig. 3a) that fall in an acceptable error band in the seasonal and annual rainfall (within the dispersion in the observations). While all the ensembles are within or near the box of acceptability (with the observed ensemble at the center by definition), only a few simulations from either CMIP3 or CMIP5 are within the box; many simulations show as much as 60% difference with the observed values

Comparision of Simulations: Differences from observations
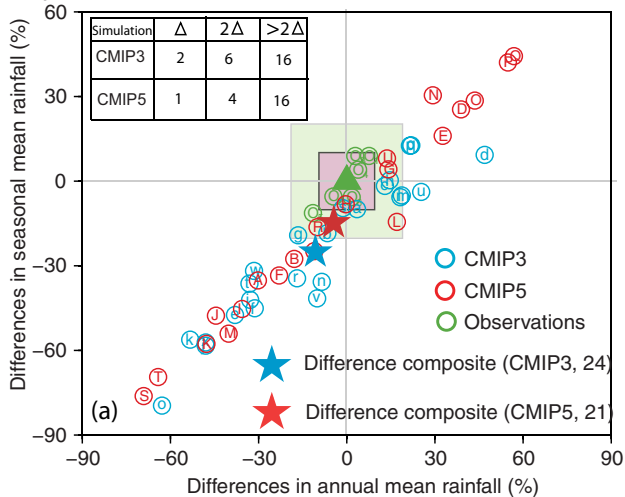
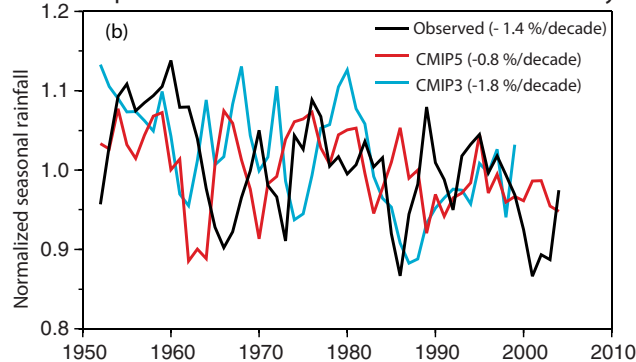## Comparision of Simulations: Interannual variability

Figure 3 | (a) Distribution of historical (1951–2005) simulations in seasonal and annual rainfall over continental India (D2) for CMIP5 (red, uppercase) and CMIP3 (blue, lower case) in terms of the difference (simulation – observed composite) as percentage of the corresponding observed composite. The dispersion in the observations (green) is shown in terms of the difference (observation – observed composite) as percentage of the observed composite. The adopted acceptable uncertainty is defined by the difference (Δ) between the maximum and the minimum in the observed values, centered at the observed composite; the inner shaded box (pink) is defined by 1Δ, while the outer square (green) is defined by 2Δ. The inset table shows the number of simulations that fall in each category. (b) Inter-annual variability in Continental Indian Monsoon (CIM) rainfall from the composite observations and ensemble of simulations (normalized to the respective 1951–2005 mean). The linear trend for the respective case is given in the bracket.

(Fig. 3a), although, the ensemble average of CMIP5 is closer to the observed composite than the CMIP3 composite (Fig. 3a). In terms of the spread around the observed values, the CMIP5 models show wider spread than the CMIP3 simulations. As per our hierarchical evaluation, nine (E,F,H,J,K,**L**,P,R,**U**) CMIP5 simulations and eight (h,i,j,k,l,o,s,u) CMIP3 simulations satisfy weaker twin criteria **(a)** and **(b)** however, there are only two CMIP5 simulation (L, U) that satisfy the twin criteria of simulating the acceptable range of observed trend and the seasonal/annual rainfall over the period 1951–2005; there are two CMIP3 simulations (h, s) that satisfy these criteria. In what follows, we shall focus on seasonal (CIM) rainfall. The area-averaged seasonal rainfall from the selected CMIP5 and CMIP3 ensembles shows decreasing trend as observed (Fig. 3b); however, CMIP5 indicates a weaker trend while CMIP3 produces a stronger trend. It is interesting to note that in terms of absolute error in trends, CMIP5 and CMIP3 are comparable.

To further examine the robustness and the consistency of the results, we have analyzed trends in different epochs. Multi-epochal analysis of trends for CMIP ensembles as well as the selected models shows (Table 2) that both CMIP composites and the selected models provide epochal trends consistent with the observed trends. Both CMIP ensembles (except for CMIP5 ensemble for 1951–1975) show only weak and often positive trends for all the three epochs against significant observed negative trends (Table 2). In contrast, the ensembles of selected models show significant negative trends for each epoch (Table 2). Similarly, four selected models show generally significant negative trends for each epoch consistent with the corresponding observed trends. In particular, epoch-wise trends are consistent with the long-period (1951–2005) trends (Table 2). This strengthens our argument that epochal trends provide a measure of reliability of simulation of future trends; however, epochal trends cannot be used readily as evaluation criteria due to various uncertainties.

However, a large number of physical processes, both regional and larger scales, are known to play important role in the dynamics of CIM[1,32–34]; it is thus a difficult task to assign exact reasons for the failure in simulating satisfying the twin criteria. While aerosol and other anthropogenic forcings[35] may be major drivers of the trends, the roles of the dynamical processes need to be investigated first. We have next considered the following indices (Fig. S1, Table 3) in terms of correlation coefficients (CC) between CIM rainfall and (a) sea surface temperature (SST) anomalies over NINO3.4 region (ENSO-CIM index), (b) Land equator thermal gradient over South Asia (LETG-SA index), (c) deep tropospheric thermal difference (DTTD index), (d) Land equator thermal gradient over India (LETG-IND index) and (e) difference in east-west SST difference (IOD index). The +ve/−ve index means the relationship between interannual variability of CIM rainfall and regional/large-scale climate indicators.

Comparison of the large-scale indices (ENSO index and the LETG-SA) from the simulations (Fig. 4a, brown and green bars, respectively) with the corresponding observed values (Fig. 4a, middle panel) shows that while many (16) of the CMIP3 simulations reproduce observed negative CC (Fig. 3a, top panel) at 99% significant level, only a few (10) CMIP5 simulations show this observed characteristic (Fig. 4a, bottom panel). In particular, the CMIP3 ensemble shows a negative ENSO-CIM index at more than 99% significance level; for the CMIP5 ensemble this significance is ~95% (Fig. 4a,

Table 2 | Epochal trends in observation are compared with climate simulation composite in both CMIP3 and CMIP5; the significant negative trends are in bold. The ''ALL'' represent all models composite and ''SEL'' refers to the composite based on the selected CMIP3/CMIP5 climate model simulations. The selection criteria are based on the simulation of annual and seasonal mean within +− 2 observed standard deviation along with the simulation of significant negative trend in seasonal rainfall for the period 1951–2005

	Epochal trends in CIM rainfall (mm/year)		
Data composite	1951–2005	1951–1975	1976–2005
OBS	**−1.4**	**−2.2**	**−3.4**
CMIP5 ALL	−0.07	**−1.52**	−0.07
CMIP3 ALL	+0.27	−0.01	+0.43
CMIP5 SEL	**−0.82**	**−0.66**	**−1.81**
CMIP3 SEL	**−1.80**	**−4.00**	**−2.59**
INMCM4 (L)	**−0.74**	**−2.23**	**−1.10**
NORESM1-M (U)	**−0.90**	+0.91	**−2.51**
GFDL_CM2_1 (h)	**−2.35**	**−5.36**	**−1.41**
MPI_ECHAM5 (s)	**−1.26**	**−2.64**	**−3.76**

5

Table 3 | Models from CMIP3 and CMIP5 that satisfy the condition of significant negative trend in CIM seasonal rainfall and seasonal (CIM) and annual rainfall within the acceptability band, are shown along with correlation coefficients between CIM rainfall and large-scale and regional scale climate indices. The models that agree with the observed correlations (same sign and significant) are highlighted; the observed values of the CC are given in the parenthesis. The acceptable ranges (±20% of the observed composite; i.e., the observed dispersion) in the ratio of seasonal and annual rainfall to the observed are highlighted. CIM rainfall projection for next 25 years (2006–2030): Linear trend (% of SD/yr) in CIM seasonal rainfall. The numbers highlighted represent significant (P < 0.1). The climate models simulations within acceptable limits are highlighted

| Model Name (symbol) | (Simulated/ observed) seasonal rainfall (%) | Correlation Coefficients (CC) between CIM rainfall and | | | | CIM rainfall projection for 2006–2030 | | |
		ENSO P < 0.05 (−0.54)	LETG-SA P < 0.05 (0.27)	DTTD P < 0.05 (−0.33)	LETG-IND P < 0.05 (−0.23)	RCP2.6/A1B	RCP4.5/B1	RCP6.0/A2
INMCM4 (L)	86	−0.30	+0.17	−0.39	−0.71	-	−3.4	−3.2
NORESM1-M (U)	108	−0.67	−0.28	−0.56	−0.61	0.0	1.35	−1.4
GFDL_CM2_1 (h)	100	−0.64	−0.49	−0.72	−0.69	−1.02	−0.73	−1.82
MPI_ECHAM5 (s)	91	−0.61	−0.32	−0.53	−0.74	1.0	3.37	3.56
Ensemble: CMIP3	75	−0.44	−0.05	−0.56	−0.69	+0.94	+0.28	+1.03
Ensemble: CMIP5	85	−0.25	−0.11	−0.42	−0.45	+0.04	+0.52	+0.59

thick brown bars). With respect to LETG-SA, the CMIP3 ensemble produces a positive CC as observed, but with low significance; for CMIP5 ensemble, this CC is zero (Fig. 4a, thick green bars). Further, only CMIP3 ensemble shows LETG-SA index of the same sign and significance as that of the observed (Fig. 4a). Thus the CMIP5 simulations have in general poorer quality than the CMIP3 simulations in reproducing the observed association between large-scale processes and CIM. Similar conclusions also hold for distribution of simulations for the regional process (Fig. 4b). In terms of the IOD index, eight of the CMIP3 models simulate values similar to the observed

and the negative trend in the seasonal rainfall; five CMIP5 simulations satisfy these twin criteria (Table S2 and S3).

As noted earlier, the acceptability of the simulation in terms of mean or variability and that with respect to trend needs to be considered separately. We next consider the difference between the observed and the simulated seasonal rainfall and difference between observed and simulated correlation coefficients (Fig. 5). Distribution of the simulations in this plane shows most of the CMIP5 simulations to have either positive or negative bias; however, the average bias is quite small (10% of the observed composite). In contrast, CMIP3
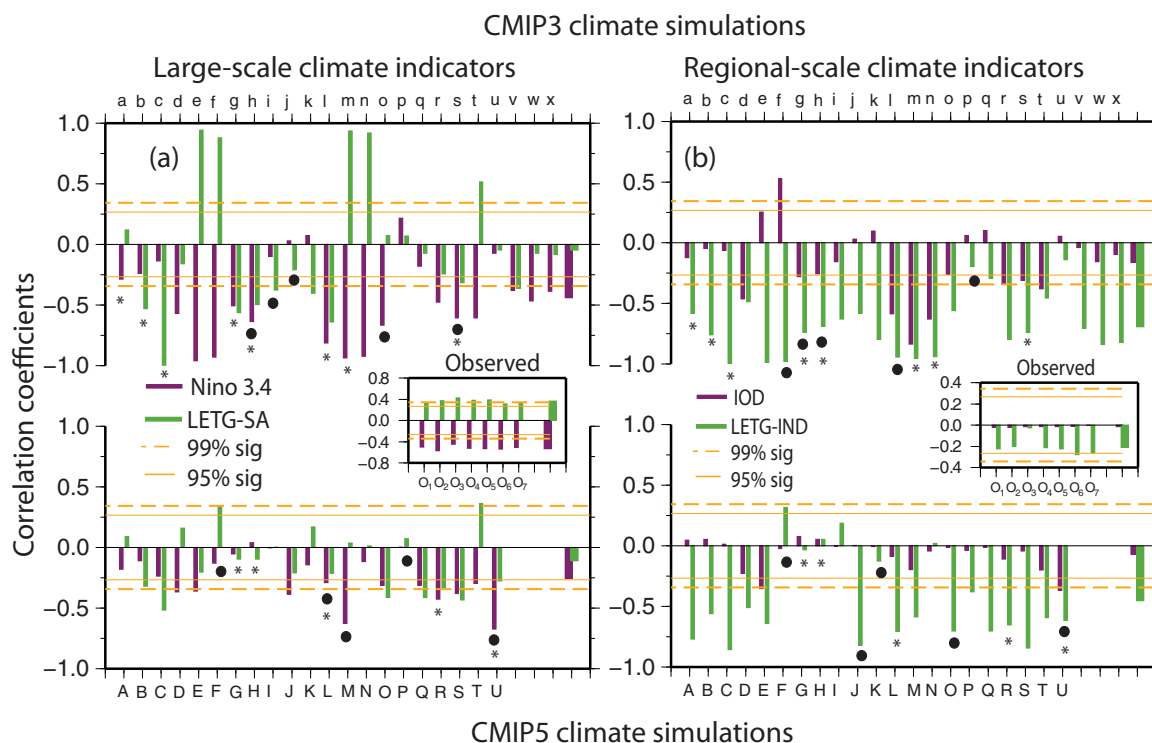

Figure 4 | Correlation coefficients between CIM rainfall and (a) large scale climate indicators (ENSO – Purple, land-ocean thermal gradient: LETG-SA - green) and (b) regional scale indicators (IOD – Purple and land-ocean thermal gradient (LETG-IND, green) for CMIP3 (top panels) and CMIP5 (bottom panels); the corresponding observed correlation coefficients from multiple observations for the period (1951–2005) are shown in the middle panel. The symbol• indicates significant (P < 0.2) negative trend, while * indicates simulations that reproduce current mean annual and seasonal rainfall in the acceptable uncertainty band.
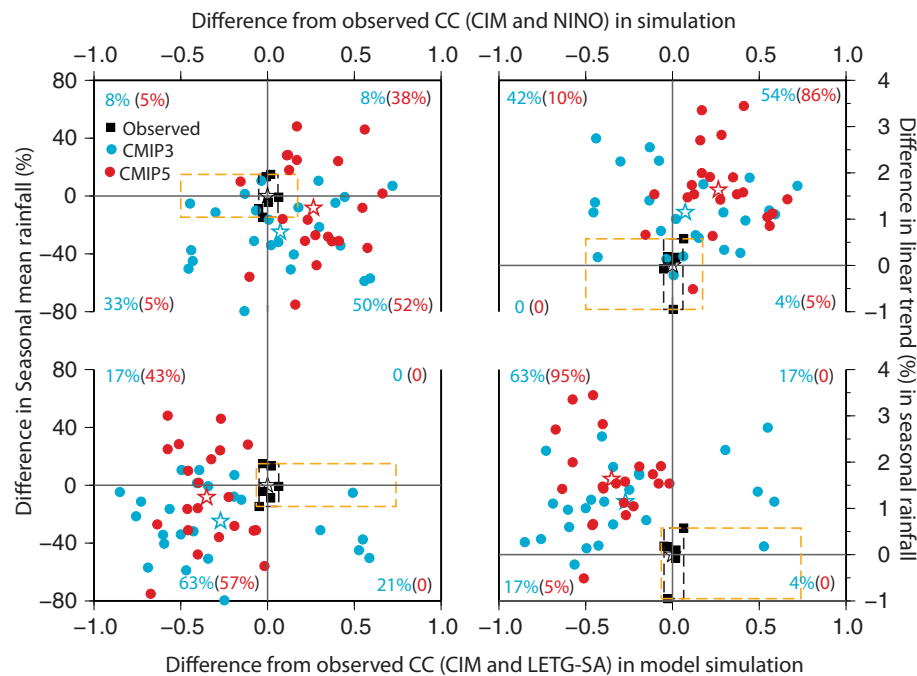
Figure 5 | Distribution of the climate simulations from CMIP3 (blue dots) and CMIP5 (red dots) in the plane spanned by the difference (error) in simulated and observed correlation (x-axis) and the difference in simulated and observed seasonal mean rainfall (left y-axis) and trend (right y axis) during 1951–2005. The values in the left y-axis are shown as percentage of their respective observed composite. The errors in simulated historical trends are shown as % of the observed trend on the right y-axis. The black rectangle defines the spread in observed values along the axis; the orange rectangle defines the acceptable range along each axis by including spread in the observation plus the significance of the correlation.

simulations show distinct negative bias (underestimation) in seasonal mean rainfall (Fig. 5, left panels); while this may be related to the basic design (forcings) of CMIP3 and CMIP5, a clear interpretation is not available. With respect to the linear trend in the observed CIM (Fig. 5, right panels) CMIP3 simulations (blue) are nearly equally distributed about the zero line, while the CMIP5 simulations show clear positive bias. In general, more CMIP3 simulations are found within the acceptability bands, with the CMIP3 ensemble closer to the observed ensemble than the CMIP5 ensemble (Fig. 5). One of the possible sources of error (difference) in simulations is the model's ability to simulate various regional and large-scale climate processes. A large spread exists in both Indian and Australian average monsoon rainfall and their interannual variability in CMIP3 and CMIP5[29]. Most CMIP5 simulations generally produce weaker than observed ENSO-CIM index, and stronger than observed LETG-SA index (Fig. 5). With respect to regional processes, both CMIP5 and CMIP3 simulations show stronger than observed LETG-IND index with comparable IOD index (Fig. S2).

It is important to note that all the simulations that satisfy the condition of significant negative trend in the historical simulation of CIM rainfall also possess LETG-IND index similar to that in the observation (Table 3); no such consistent signal is seen for the other two regional indices DTTD (Table 3) and IOD (Fig. 4b). Indeed all the simulations that satisfy some acceptability criteria also show LETG-IND index close to observation (Table S2), although some models are found with weaker acceptability criteria (2Δ) in terms of annual and seasonal mean rainfall. In terms of large-scale indices, all the seven simulations that satisfy the criteria in terms of seasonal and annual mean as well as significant negative trend in the historical simulation of CIM rainfall show significant (99% in five cases, 95% in two cases) ENSO index as observed, while no consistent signal emerges for the LETG-SA index (Table 3). It appears that successful simultaneous simulation of some indices is necessary for reproducing accurate trend in the historical data and acceptable annual and seasonal mean; overall, the importance of the land-ocean

contrast is consistent with earlier results[36]. However, while these may provide necessary conditions, they are clearly not sufficient. An important feature is that there is no simulation that shows significant negative trend in seasonal rainfall without simulating the LETG-IND index and ENSO index similar to the observed values (Table 3).

It is interesting to note that the models L, U and h that satisfy the criteria in terms of seasonal and annual mean as well as significant negative trend in the historical simulation of CIM rainfall project weak or significant negative trend for the period 2006–2030 (Table 2). In contrast, the three simulations (s) with low annual and seasonal mean (but with significant negative trend in the historical simulation of CIM rainfall) project high and positive trend for 2006–2030 for most scenarios (Table 2).

The monsoon rainfall also exhibits prominent multi-decadal variability[37]; thus epochal trends are modulated by the phase of the multi-decadal variability. The quality of simulations of multi-decadal variability can be thus another important measure of quality of the CMIP simulations. The multi-decadal climate variability in area averaged (Land only: 70-85E; 10-30N) seasonal rainfall for the period 1901–2005, represented as Cramer's t-statistic[38,39] shows (Fig. 6) that both CMIP3 and CMIP5 simulations (selected and ensemble) exhibit multi-decadal variability with characteristics similar to those in composite observation. It also clarifies the hazards of trying to break up the inter decadal behavior with an arbitrary dividing date such as 1975 (Fig. 3b and Fig. 6). The selected simulations show a much higher degree of coherence (similar phase) than the rest of the simulations; in particular, the all-simulation ensemble of CMIP3 does not exhibit the observed characteristics (Fig. 6); interestingly, the all-simulation ensemble of CMIP5 shows a decadal variability but with phases essentially opposite to those of observation or ensemble of selected CMIP5 models (Figure 6). While 76% of CMIP5 simulations simulate positive phase of slow time-scale variability during 1920–1955 and 67% show negative phase for the period 1956–2000, only 62% of simulations show both phases (Fig. 6d). On the other hand, CMIP3 simulates ~50% of the simulations show
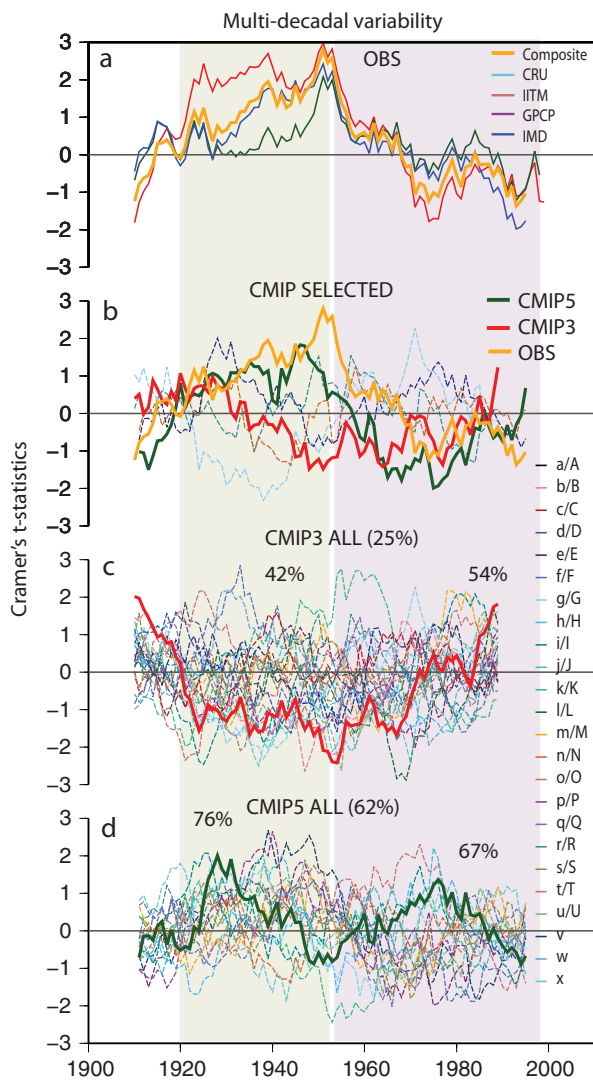
Figure 6 | Multi-decadal variability expressed in terms of Cramer's t-statistics in observation and climate simulations in terms of area averaged (D2, continental India) seasonal rainfall with running average from (a) multiple observations and their composite (b) climate simulations from CMIP5 and CMIP3 based on the climate models ability to simulate the observed seasonal trends and mean (both the seasonal and annual) rainfall (c) CMIP3 (24 models) simulations (d) CMIP5 (21 models) simulations. The thick line represents respective composite. The percentage of climate simulations that simulated the phase of multi-decadal variability is shown in both CMIP3 and CMIP5, while the simulations simulate both phases are shown near to the sub-title.

the phases individually; it poorly simulates (25%) both the phases of slow variability (Fig. 6c). While the epochal trends may have various uncertainties such as due to low-frequency (decadal) variability, and cannot be used readily for model evaluation, it is interesting, and encouraging, that simulations that satisfy the other acceptability criteria also perform well in terms of simulation of low-frequency variability. In particular, the models selected based on hierarchical assessment, one each from CMIP5 (L) and CMIP3 (s), simulate both the phases and the amplitudes reasonably well, while others do not.

Discussion

An accurate simulation of the observed trends is a primary indicator of reliability of projections for future climate. Our results show that no significant progress has been achieved in our ability to simulate basic quantities like observed seasonal mean and trend, and hence to

project the regional climate system, namely CIM, with reasonable certainty. It is normally expected that an ensemble of simulations, statistically, would provide a more reliable result than any of the individual simulations; given that the (equal-weight) ensemble from either CMIP3 or CMIP5 does not provide a more accurate simulation, there is an urgent need for redesigning multi-model ensemble. Given that CMIP5 have poorer quality than the CMIP3 in simulating the observed features of CIM, a critical look at our strategy for model improvement is also required. A silver lining is the consistency across the large number of simulations in terms of association between simulations and climate indices. Such consistency provides an objective way of assessing reliability of climate projections based on physical and mechanistic understanding; such approaches are necessary to offset any effect of non-stationarity of observed trends in assessing projections. However, while the quality of the simulations in terms of various metrics might not have improved, the CMIP5 simulations can be argued to be based on a more comprehensive and refined knowledge base[17], implying large applicability.

A major aspect of hierarchical evaluation is based on validation of simulated trends against observed trends in CIM rainfall for assessing confidence in future-climate simulations. We should emphasize that such a validation of simulated trends certainly provides a necessary requirement; but not necessarily a sufficient condition. Analysis of climate indices shows that in both CMIP5 and CMIP3 certain common processes at large and regional scales show consistency with skill in simulation (Table S2); in particular, these climate simulations also simulate the observed relationship between CIM and ENSO/LETG-IND. Similarly, other parameters in addition to phases of multi-decadal oscillation and multi-epoch trend can be included for enhanced confidence.

We have refrained from assigning any definite cause(s) for the decrease in rainfall during 1951–1975 as there are many interacting complex processes as well as a spectrum of natural variability at different time periods. It could be argued, based on the similarity of trends during 1951–1975 and those during 1975–2005, that aerosol or anthropogenic sources may not have significant contribution to the declining trends in CIM in the current epoch; however, there is some evidence[40], that aerosols may have significant effects in the recent epoch, 1951–1996. However, definitive attribution is difficult as the precise amplitudes and the phases of the low-frequency variability are not known. The enhanced scope in the CMIP5 models may enable addressing a wider spectrum of applications[5,16]. Although we have considered CIM for a case study, arguments presented earlier indicate that assessments are needed for other regional systems.

The current analysis is based on simulations with a single initial condition from each model. However, use of multiple simulations with different initial conditions, may not change the basic conclusions, especially since the current simulations also use different initial conditions; however, this issue needs to be explored. As against the natural system, the correlations, trends etc in the simulations are a result of the model dynamics that is not necessarily identical across the models. Thus, each model will need independent analysis and improvement. Finally, our results suggest a metric (application) based evaluation of climate models rather than looking for a best model that performs well for all metrics. Quite clearly, a metric for CIM rainfall may not be adequate for comparing spatial patterns that are more useful for application; however, more elaborate comparison will require climate simulations of adequate skill. While these issues are definitely challenging, they need to be properly addressed given the importance of climate projections.

Methods

gridded (1° × 1°) rainfall dataset is adopted from the India Meteorological Department (IMDG) for the period 1951–2004 over India[22]. Monthly rainfall on a 2.5-degree global grid from 1979 to the present is adopted from Global Precipitation Climatology Project (GPCP) based on over 6,000 rain gauge stations, and satellite geostationary and low-orbit infrared, passive microwave, and sounding observations[24]; monthly rainfall is also adopted from CRU-Climate Research Unit[23]; NCEP reanalysis rainfall[26]; APHRO rainfall[24], All India averaged monthly rainfall from IITM[20] and IMD[21].

In this study we use Cramer's test is to examine the stability of a long-term record in terms of comparison between the overall mean of an entire record and the mean of the certain part of the record (WMO 1966). The statistical significance of the moving mean as well as decadal averages was examined using the Cramer's test statistics as follows:

$$t_k = \sqrt{\left[\frac{n(N-2)}{N-n(1+r_k^2)}\right]}\, r_k \text{ where, } r_k = \left(\overline{R}_k - \overline{R}\right)/\sigma$$

$\overline{R}$ is the mean for the total number of years and σ is the standard deviation of the series for the total number of years (N) under investigation; $\overline{R_k}$ is the mean for each successive n-year. The statistic t_k is distributed as 'Student's t' with N-2 degrees of freedom. This test may be repeated for any desired number and choice of sub periods in the whole record. The time plot of the t-value gives the pictorial representation of variability.

1. Webster, P. J. et al. Monsoons: processes, predictability, and the prospects for prediction. J. Geophys. Res. 103, 14451–14510 (1998).
2. Meehl, G. A. et al. Global Climate Projections. In Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (eds Solomon, S. et al.). Cambridge University Press, Cambridge, UK, 747–845 (2007).
3. Turner, A. G. & Annamalai, A. Climate change and the South Asian summer monsoon. Nat. Clim. Change. 2, 587–595 (2012).
4. DelSole, T. & Shukla, J. Climate models produce skillful predictions of Indian summer monsoon rainfall. Geophys. Res. Lett. 39, L09703 (2012).
5. Deser, C., Knutti, R., Solomon, S. & Phillips, A. S. Communication of the role of natural variability in future North American climate. Nat. Clim. Change 2, 775–779 (2012).
6. Goswami, B. N., Madhusoodanan, M. S., Neema, C. P. & Sengupta, D. A physical mechanism for North Atlantic SST influence on the Indian summer monsoon. Geophys. Res. Lett. 33, L02706 (2006).
7. Kitoh. et al. Monsoons in a changing world: A regional perspective in a global context. J. Geophys. Res. Atmos. 118, 3053–3065 (2013).
8. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. Bull. Amer. Meteor. Soc. 93, 485–498 (2012).
9. Solomon, S. et al. IPCC Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Inter- governmental Panel on Climate Change. Cambridge University Press 1–996 (2007).
10. Sperber, K. R. et al. The Asian summer monsoon: An intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century. Clim. Dyn. 1–34 (Published online 2012).
11. Kripalani, R. H., Oh, J. H., Kulkarni, A., Sabade, S. S. & Chaudhari, H. S. South Asian summer monsoon precipitation variability: Coupled climate model simulations and projections under IPCC AR4. Theor. App. Clim. 90, 133–159 (2007).
12. Lee, J. & Wang, B. Future change of global monsoon in the CMIP5. Clim. Dyn. 1–19 (2012).
13. Hsu, P.-C., Li, T., Luo, J.-J., Murakami, H., Kitoh, A. & Zhao, M. Increase of global monsoon area and precipitation under global warming: A robust signal? Geophys. Res. Lett. 39, L06701 (2012).
14. Menon, A., Levermann, A., Schewe, J., Lehmann, J. & Frieler, K. Consistent increase in Indian monsoon rainfall and its variability across CMIP-5 models. Earth Syst. Dynam. Discuss. 4, 1–24 (2013).
15. Li, J.-L. F. et al. An observationally-based evaluation of cloud ice water in CMIP3 and CMIP5 GCMs and contemporary reanalyses using contemporary satellite data. J. Geophys. Res. 117, D16105 (2012).
16. Krishna Kumar. et al. The Once and Future Pulse of Indian Monsoonal Climate. Clim. Dyn. 36, 2159–2170 (2011).
17. Knutti, R. & Sedlacek, J. Robustness and uncertainties in the new CMIP5 climate model projections. Nature Climate Change 3, 369–373 (2013).
18. Syroka, J. & Toumi, R. Recent lengthening of the south Asian summer monsoon season. Geophys. Res. Lett. 29, L1458 (2012).
19. Ramesh, K. V. & Goswami, P. Reduction in temporal and spatial extent of the Indian summer monsoon. Geophys. Res. Lett. 34, L23704 (2007).
20. Parthasarathy, B., Munot, A. A. & Kothawale, D. R. All-India monthly and seasonal rainfall series: 1871–1993. Theo. and Appl. Clim. 49, 217–224 (1994).
21. All India Monthly and Seasonal (June-September) Rainfall Series, Rainfall Data for SW Monsoon (1901–2010). Indian Meteorological Department (IMD) http://www.imd.gov.in/section/nhac/dynamic/Monsoon_frame.htm (Oct 18 2013).
22. Rajeevan. et al. High resolution daily gridded rainfall data for Indian region: Analysis of break and active monsoon spells. curr. sci. 91, 296–306 (2006).
23. Mitchell, T. D. & Jones, P. D. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int. J. Climatology 25, 693–712 (2005).
24. Huffman, G. J., Adler, R. F., Bolvin, D. T. & Gu, G. Improving the global precipitation record: GPCP Version 2.1. Geophys. Res. Lett. 36, L17808 (2009).
25. Yatagai, A. et al. APHRODITE: Constructing a Long-term Daily Gridded Precipitation Dataset for Asia based on a Dense Network of Rain Gauges. Bull. Amer. Meteor. Soc. 93, 1401–1415 (2012).
26. Kistler, R. et al. The NCEP-NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation. Bull. Amer. Meteor. Soc. 82, 247–268 (2001).
27. Wang, B. & Ding, Q. Changes in global monsoon precipitation over the past 56 years. Geophys. Res. Lett. 33, L06711 (2006).
28. Wang, B., Liu, Y., Kim, H.-J., Webster, P. J. & Yim, S.-Y. Recent change of the global monsoon precipitation (1979–2008). Clim.Dynam. 39, 1123–1135 (2012).
29. Jourdain. et al. The Indo-Australian monsoon and its relationship to ENSO and IOD in reanalysis data and the CMIP3/CMIP5 simulations. Clim. Dynam. 1–30 (Published online 2013).
30. Kodra, E., Ghosh, S. & Ganguly, A. R. Evaluation of global climate models for Indian monsoon climatology. Env. Res. Lett. 7, 014012 (2012).
31. Knutti, R. & Sedlacek, J. Robustness and uncertainties in the new CMIP5 climate model projections. Nature Climate Change 3, 369–373 (2012).
32. Li, C. & Yanai, M. The onset and interannual variability of the Asian summer monsoon in relation to land-sea thermal contrast. J. Clim. 9, 358–375 (1996).
33. Webster, P. J., Moore, A., Loschnigg, J. & Leban, M. Coupled ocean-atmosphere dynamics in the Indian Ocean during 1997–98. Nature 401, 356–360 (1999).
34. Ihara, C. et al. Indian summer monsoon rainfall and its link with ENSO and Indian Ocean climate indices. Int. J. Clim. 27, 179–187 (2007).
35. Bollasina, M. A., Ming, Y. & Ramaswamy, V. Anthropogenic aerosols and the weakening of the South Asian Monsoon. Science 334, 502–505 (2011).
36. Fasullo, J. A mechanism for land-ocean contrasts in global monsoon trends in a warming climate. Clim. Dyn. 39, 1137–1147 (2012).
37. Kripalani, R. H. & Kulkarni, A. Climatic impact of El Nino/La Nina on the Indian monsoon: A new perspective. Weather 52, 39–46 (1997).
38. World Meteorological Organization climate change. WMO Tech. Note No. 79, WMO No. 195-TP-100, 1–53 (1966).
39. Trenberth, K. E. Decadal variability in ENSO an extratropical teleconnections: Evidence of global warming? In: Proceedings of International Scientific Conference on TOGA Programme, 2–7 Apnl 1995, Melbourne, Ausnalia WCRP 91, WMO/TD No. 717, 57–61 (1995).
40. Rotstayn, L. D. et al. Have Australian Rainfall and Cloudiness Increased Due to the Remote Effects of Asian Anthropogenic Aerosols? J. Geo Phys Res. 112, D09202 (2013).

Acknowledgments

Author contributions

P.G. and K.V.R. wrote the main manuscript text and K.V.R. & P.G. prepared the figure. All authors reviewed the manuscript.

Additional information