

LETTER • OPEN ACCESS

Atmospheric moisture as a proxy for the ISMR variability and associated extreme weather events

To cite this article: P J Nair *et al* 2021 *Environ. Res. Lett.* **16** 014045

View the [article online](#) for updates and enhancements.

You may also like

- [On the decreasing trend of the number of monsoon depressions in the Bay of Bengal](#)
S Vishnu, P A Francis, S S C Shenoi *et al.*
- [Prediction of Indian rainfall during the summer monsoon season on the basis of links with equatorial Pacific and Indian Ocean climate indices](#)
Sajani Surendran, Sulochana Gadgil, P A Francis *et al.*
- [The recycling rate of atmospheric moisture over the past two decades \(1988–2009\)](#)
Liming Li, Xun Jiang, Moustafa T Chahine *et al.*



Breath Biopsy® OMNI®

The most advanced, complete solution for
global breath biomarker analysis

TRANSFORM YOUR
RESEARCH WORKFLOW



Expert Study Design
& Management



Robust Breath
Collection



Reliable Sample
Processing & Analysis



In-depth Data
Analysis



Specialist Data
Interpretation

ENVIRONMENTAL RESEARCH
LETTERS

LETTER

OPEN ACCESS

RECEIVED

27 October 2020

ACCEPTED FOR PUBLICATION

2 December 2020

PUBLISHED

12 January 2021

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.

Atmospheric moisture as a proxy for the ISMR variability
and associated extreme weather eventsP J Nair¹ , H Varikoden², P A Francis³, A Chakraborty¹ and P C Pandey¹¹ CORAL, Indian Institute of Technology Kharagpur, West Bengal, India² Indian Institute of Tropical Meteorology, Ministry of Earth Sciences, Pashan, Pune, India³ Indian National Centre for Ocean Information Services, Ministry of Earth Sciences, Hyderabad, IndiaE-mail: pjnnair@yahoo.in**Keywords:** ISMR, moisture transport, PWC, MLR, FAR, PODSupplementary material for this article is available [online](#)

Abstract

This study explores the potential of atmospheric moisture content, its transport and its divergence over the ocean and land as proxies for the variability of Indian summer monsoon rainfall (ISMR) for the period 1950–2019. The analyses using multiple linear regression reveal that the interannual and intraseasonal variability of ISMR and the mean ISMR is largely controlled by Arabian Sea moisture flux and Ganga river basin moisture content, and these parameters exhibit statistically significant high correlations in most regions. The regression model and the parameters are statistically significant and the model could explain rainfall variability of about 12%–50% in various regions. The model shows a false alarm rate (FAR) of 0.25–0.45 and a probability of detection (POD) of 0.43–0.50 for wet years in West Central, North West and North Central India. The FAR and POD are about 0.06–0.32 and 0.60–0.70, respectively for dry years in those regions. The model reproduces flood and drought years of about 32%–50% and 55%–70% in those regions. Also, the moisture indices could clearly identify the majority of wet and dry years that occurred during the period. The ISMR variability associated with moisture indices is unaffected by El Niño Southern Oscillation. Henceforth, this study demonstrates the significance of atmospheric moisture on regional rainfall distribution and suggests that these parameters can be used in both statistical and dynamical models to better predict monsoon and global precipitation.

1. Introduction

The hydrological cycle is an important natural process in which water reaches the atmosphere from water bodies through evaporation and plants through evapotranspiration, which eventually returns to the ground as precipitation. Generally, the Hadley cell in the austral winter supplies the water vapor required for the boreal summer. Therefore, precipitation in the northern hemisphere during boreal summer is made available by moisture in the austral winter, particularly in tropical regions (Peixóto and Oort 1983). In India, there are two main rainy seasons: the southwest (June through September—JJAS) and northeast (October through December—OND) monsoons. The southwest monsoon is the major rainy season bringing moist air from the oceans to the Indian subcontinent, known as the Indian Summer

Monsoon (ISM) (Pant and Kumar 1997). During this season, moisture transport is regulated by southwest winds followed by the cross-equatorial flow that normally decides the strength of moisture transport and thus the nature of the ISM to some extent (Ramesh Kumar *et al* 1999).

It is well known that the relationship between ISM rainfall (ISMR) and El Niño Southern Oscillation (ENSO) has been strong in previous decades (Annamalai and Liu 2005, Mishra *et al* 2012, Ashok *et al* 2001), but weak in the recent decades (Krishna Kumar *et al* 1999, Ashok *et al* 2001, Pai 2004, Hrudya *et al* 2020, Seetha *et al* 2020); while ISMR is correlated well with the Indian Ocean Dipole events (Wang *et al* 2015, Gadgil and Francis 2016, Yun and Timmerman 2018). Several proxies like Atlantic multidecadal oscillation, Atlantic zonal mode, El Niño modoki and extratropical sea surface temperature

that affect ISMR have also been introduced recently (Goswami *et al* 2006, Zhang *et al* 2006, Kucharski *et al* 2008, Pottapinjara *et al* 2014, Chattopadhyay *et al* 2015, Feifei *et al* 2011, Garfinkel *et al* 2013), but were not very successful in accurate prediction of ISMR. Therefore, it is important to investigate other factors that can be used for predicting ISMR together with the existing climate forcings. Several studies (e.g. Gautam and Pandey 1995, Fasullo and Webster 2002) reported the role of moisture transport in deciding the onset and withdrawal of ISMR and wet or dry monsoon years in India. It also modulates the frequency of monsoon depressions over the Bay of Bengal, which is a major factor in deciding the strength of ISMR (Vishnu *et al* 2016, Vishnu *et al* 2018). A study by Luis and Pandey (2004) noted surface atmospheric moisture convergence as a predictor for ISMR. Similarly, Ramakrishna *et al* (2017) showed the influence of moisture divergence for the low rainfall over India in June 2014. Apart from these, the total moisture content, measured as the precipitable water content (PWC), is also considered as a precursor of onset and withdrawal of Indian monsoon in a study by Puviarasan *et al* (2015). Similarly, the key moisture source regions and their contributions to ISMR have already been identified in previous works (Shukla and Misra 1977, Mei *et al* 2015, Pathak *et al* 2017) and are the western, central and upper Indian Ocean, the Ganga river basin and the Red Sea.

Atmospheric moisture is a key driver to extreme weather events, and the hydrological cycle is essential for life on Earth and, therefore, their changes have to be monitored to predict extreme rainfall in the context of global warming. The abovementioned studies emphasize the importance of atmospheric moisture content, its transport and divergence, and their connections with ISMR. However, the application of moisture-related factors as proxies of rainfall changes in India has not been examined thoroughly. Henceforth, this study presents new indices with respect to the atmospheric moisture parameters of PWC, vertically integrated moisture flux (VIMF) and vertically integrated moisture flux divergence (VIMFD) over various source regions, and analyze their influences on the regional rainfall changes in India using a multiple linear regression (MLR) model. The interannual and intraseasonal variations of ISMR and the changes in mean ISMR associated with these moisture indices are investigated. Then, the potential of these new indices in explaining extreme weather events is also examined. The influence of ENSO on the moisture parameters is analyzed, and their combined effect in improving the model and interpreting the regional rainfall variability is assessed.

2. Data and methods

2.1. Rainfall data

The India Meteorological Department (IMD) gridded daily rainfall measurements made from rain

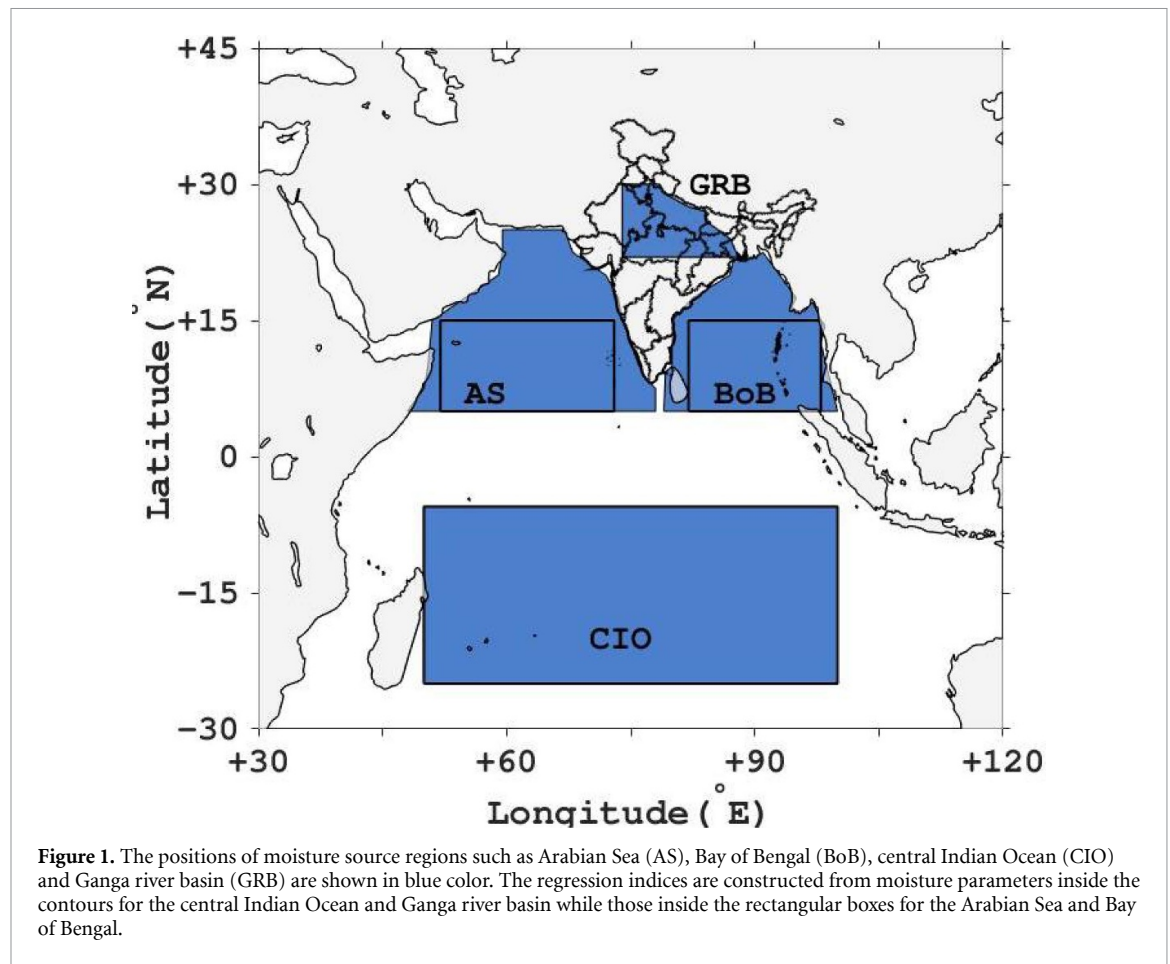
gauges installed at different places in India are used for the period 1950–2019. The data available on a $0.25^\circ \times 0.25^\circ$ latitude \times longitude horizontal resolution are area-averaged over different geographic locations: Peninsular India (PI), West Central India (WCI), North West India (NWI), North Central India (NCI) and North East India (NEI). The exact locations of these regions are given in figure S1. PI includes Andhra Pradesh, Tamil Nadu, Kerala, Karnataka and Goa. WCI includes Maharashtra, Madhya Pradesh and Telangana. NWI includes Gujarat, Rajasthan, Punjab and Haryana. NCI includes Delhi, Uttar Pradesh, Bihar, Jharkhand, Chhattisgarh and Odisha. NEI includes West Bengal, Sikkim, Meghalaya, Tripura, Mizoram, Manipur, Nagaland, Arunachal Pradesh and Assam.

The anomaly of monthly accumulated rainfall time series (in mm/month) for the months from June to September (JJAS) is computed by subtracting monthly climatology (1950–2019) from the accumulated data for the corresponding month over the study regions. Similarly, anomaly of JJAS seasonal rainfall, both accumulated in mm/season and averaged in mm/day is calculated by subtracting corresponding seasonal climatology from the respective data over the mentioned locations. The percentage deviation (in %) of JJAS rainfall is evaluated by dividing the accumulated seasonal climatology from the accumulated seasonal rainfall anomaly time series. Thus, four sets of rainfall anomaly time series (monthly accumulated, accumulated JJAS rainfall, mean JJAS rainfall and percentage deviation of JJAS rainfall) are made for MLR analysis with moisture indices.

2.2. Regression indices

The zonal and meridional wind components and specific humidity data at pressure levels of 1000, 925, 850, 700, 600, 500, 400 and 300 hPa on a $2.5^\circ \times 2.5^\circ$ spatial resolution for 6 h intervals taken from the National Center for Environmental Prediction and National Center for Atmospheric Research (NCEP/NCAR) reanalysis (Kalnay *et al* 1996) are used for the moisture flux analyses. The multivariate ENSO index (MEI) data taken from <https://www.esrl.noaa.gov/psd/data/climateindices/list/> for the period 1950–2018 and updated from <https://www.esrl.noaa.gov/psd/enso/mei/data/meiv2.data> for 2019 are used as ENSO indices. The moisture indices are created from the atmospheric moisture parameters of PWC, VIMF and VIMFD over the moisture source regions as mentioned below. These moisture parameters are directly connected to atmospheric dynamics and regional climate. Therefore, the indices made from these factors would also represent changes in climate and atmospheric dynamics.

The moisture source regions are selected from the VIMFD values at 1000–850 and 1000–300 hPa levels as illustrated in figure S2 and are Arabian Sea (5°S – 15°N ; 48° – 78°E), Bay of Bengal



(5°S–15°N; 80°–100°E), Central Indian Ocean (25°–5°S; 50°–100°E) and Ganga River Basin (22°–30°N; 74°–88°E) in which the first three are oceanic regions and the fourth is a terrestrial source region. The locations of these source regions are demarcated in blue in figure 1. The moisture parameters inside the rectangular boxes are considered for creating indices with respect to the Arabian Sea and Bay of Bengal. The regression indices are computed by averaging moisture parameters over the marked regions for each month, and then monthly climatology is subtracted from it and normalized with the standard deviation over 1950–2019. Thus, 12 new moisture indices are made from VIMF, VIMFD and PWC over the Arabian Sea (denoted as VIMF_A, VIMFD_A and PWC_A, respectively), Bay of Bengal (VIMF_B, VIMFD_B and PWC_B), central Indian Ocean (VIMF_C, VIMFD_C and PWC_C) and Ganga river basin (VIMF_G, VIMFD_G and PWC_G). The temporal evolution of the moisture indices at 1000–300 hPa level is provided in figure 2 and that at 1000–850 hPa level is given in figure S3.

2.3. Moisture flux calculation

The PWC (W) between the surface (P_s) and t th level (P_t) atmospheric pressure (P) is calculated using the acceleration due to gravity (g) and atmospheric specific humidity (q) as given in Ullah and Gao (2012) and is:

$$W = \frac{1}{g} \int_{P_t}^{P_s} q dP. \quad (1)$$

The VIMF or the total instantaneous moisture flux transport (Q) is calculated as:

$$\vec{Q} = \frac{1}{g} \int_{P_t}^{P_s} q \vec{v} dP. \quad (2)$$

The vertical integration is performed over the lower troposphere (1000–850 hPa) and vertical column (1000–300 hPa) using the trapezoidal rule. This VIMF is decomposed into divergent and rotational components (Chakraborty *et al* 2006) as:

$$\begin{aligned} \vec{Q} &= \hat{k} \times \nabla \psi + \nabla \chi, \\ \nabla^2 \psi &= \hat{k} \cdot \nabla \times \vec{Q}, \\ \nabla^2 \chi &= \nabla \cdot \vec{Q}. \end{aligned} \quad (3)$$

The divergence of VIMF ($\nabla \cdot \vec{Q}$) is termed as VIMFD.

2.4. Formulation of the multiple linear regression model

The influence of atmospheric moisture parameters and moisture sources on ISMR is diagnosed by applying a statistical technique using the MLR developed by Nair *et al* (2018). As a first step, the correlation between the proxies and accumulated JJAS rainfall is computed at different regions, and

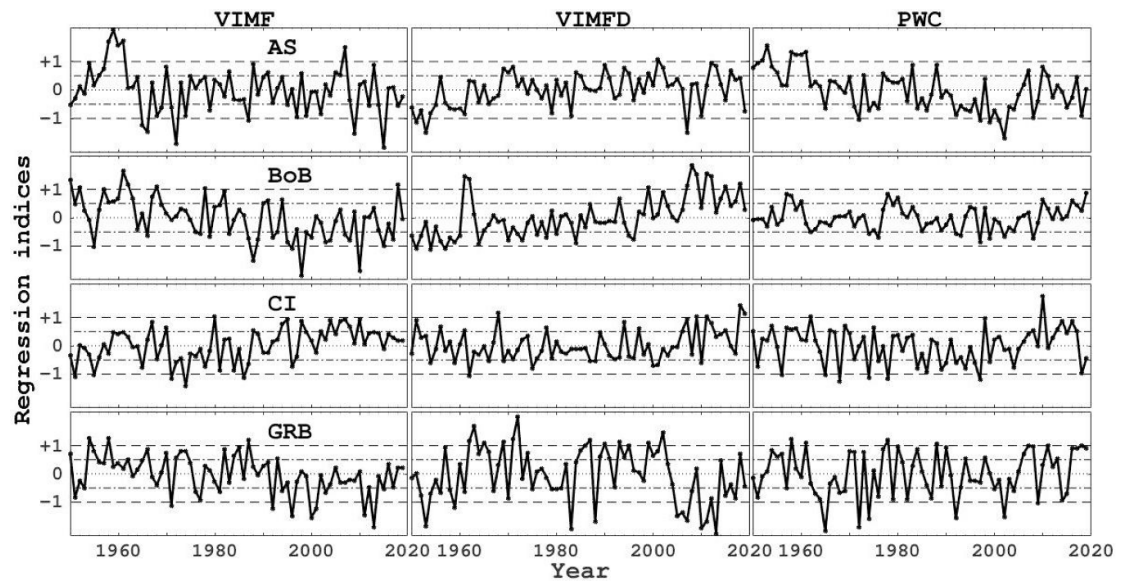


Figure 2. The regression indices created from the moisture parameters of vertically integrated moisture flux (VIMF—left panels), vertically integrated moisture flux divergence (VIMFD—middle panels) and precipitable water content (PWC—right panels) at 1000–300 hPa level over the moisture source regions of Arabian Sea (AS—top panels), Bay of Bengal (BoB—second panels from top), Central Indian Ocean (CI—third panels from top) and Ganga river basin (GRB—bottom panels). The dash-dotted lines indicate ± 5 , the dashed lines represent ± 1 and dotted lines signify zero.

then correlation among the proxies is calculated to check the multicollinearity. The proxies that correlate well with other indices are exempted from the analysis. The stepwise regression procedure is used to choose the best model from significant predictors independent of each other (Draper and Smith 2015). The parameter that is highly correlated with the rainfall and satisfying the limits of statistical significance (as mentioned below) is included first, and then the next most highly correlated variable is added. The statistical significance of the model and parameters is tested along with the improvement of the model. All parameters are thus examined at each stage of adding a new proxy and thereby removing variables that are not significant or do not improve the model performance (<https://statisticsbyjim.com/regression/model-specification-variable-selection/>).

Several statistical methods are adopted to verify the significance of parameters and model. The statistical significance of parameters is assessed using Student's t -test and the overall significance of the model is tested using the F -statistic. The overall fit of the model is examined using R^2 and adjusted R^2 , and its overall accuracy is analyzed using root-mean-square error (RMSE) and bias. The adjusted R^2 and F -statistic decide the improvement of the model as their values are increased only if the added proxy improves the model. The multicollinearity is analyzed using tolerance and the variance inflation factor (VIF). The tolerance should be greater than 0.1 or 0.2 (Lin 2008) and the VIF should be less than 2.5 (Senaviratna and Cooray 2019) for the model to be free from multicollinearity. The autocorrelation

in the residuals is tested using the Durbin–Watson (DW) statistic (Durbin and Watson 1950, Durbin and Watson 1951) and its satisfied range is between 1.5 and 2.5. Therefore, appropriate regression models are constructed from the significant parameters that explain maximum variance and are given in equations (4)–(8). The indices corresponding to the Bay of Bengal are not used in the MLR model as these correlate with other indices and are statistically insignificant.

$$PI(t) = C_Z Z(t) + C_{VIMF_A} VIMF_A(t) + C_{PWC_G} PWC_G(t) + C_{VIMFD_C} VIMFD_C(t) + \epsilon(t), \quad (4)$$

$$WCI(t) = C_Z Z(t) + C_{VIMF_A} VIMF_A(t) + C_{PWC_G} PWC_G(t) + C_{VIMFD_A} VIMFD_A(t) + \epsilon(t), \quad (5)$$

$$NWI(t) = C_Z Z(t) + C_{VIMF_A} VIMF_A(t) + C_{PWC_G} PWC_G(t) + C_{VIMF_C} VIMF_C(t) + \epsilon(t), \quad (6)$$

$$NCI(t) = C_Z Z(t) + C_{VIMF_A} VIMF_A(t) + C_{PWC_G} PWC_G(t) + \epsilon(t), \quad (7)$$

$$\begin{aligned}
 NEI(t) = & C_Z Z(t) + C_{VIMF_G} VIMF_G(t) \\
 & + C_{PWC_A} PWC_A(t) + C_{PWC_C} PWC_C(t) \\
 & + \epsilon(t),
 \end{aligned}
 \quad (5)$$

where the left-hand side of the equation is rainfall data at various regions and is treated as predictand, t is years, Z is a constant level term (taken as value 1) and C_Z is the intercept that helps to adjust the fit not passing through the origin when a new proxy is added to the equation and is used to adjust the shift of mean predicted value with that of mean observed value (Krzywinski and Altman 2015). C_{VIMF_A} , C_{VIMF_C} and C_{VIMF_G} are regression coefficients of VIMF_A, VIMF_C and VIMF_G, respectively. Similarly, C_{VIMFD_A} and C_{VIMFD_C} are regression coefficients of VIMFD_A and VIMFD_C, respectively. C_{PWC_A} , C_{PWC_C} and C_{PWC_G} are in turn regression coefficients of PWC_A, PWC_C and PWC_G, and ϵ is the residual. The terms on the right-hand side, except the intercept and residual, are considered as predictors. Another point is that the moisture parameters at 1000–300 hPa are used as regression indices in the PI, WCI, NWI and NCI regions whereas those at 1000–850 hPa are used in NEI (details are provided in section 3.1) for MLR analysis.

2.5. Methodology

The rainfall and proxy data are detrended for removing long-term trends from the data. The indices are scaled to unity amplitude before performing regression analysis so that regression coefficients are obtained in the same unit of input data (Gopalapillai 2012). The parameters are solved using the least squares method by minimizing the error of predictors (Press *et al* 1989). The regression analysis is done on four sets of rainfall data and are (1) accumulated JJAS rainfall anomaly time series (2) accumulated monthly rainfall anomaly time series for the months from June to September, (3) mean JJAS rainfall and (4) percentage deviation of JJAS rainfall. Similarly, regression analysis is performed in three ways for the accumulated JJAS rainfall anomaly with (1) the moisture indices, to see the influence of moisture indices on ISMR, (2) moisture indices and ENSO index, to check the combined impact of ENSO and moisture indices, and (3) ENSO index alone, to see the individual impact of ENSO on ISMR.

3. Results and discussion

3.1. Correlation analysis

Table 1 presents correlation coefficients between the accumulated JJAS rainfall and moisture indices at 1000–300 and 1000–850 hPa levels along with the ENSO index at different regions. The correlations, statistically significant at a 95% confidence interval, are given in bold numbers. MEI shows statistically significant correlations of about -0.30 , -0.39 , -0.35 and -0.54 with the rainfall in PI, WCI, NWI and NCI,

respectively and a weak positive correlation of about 0.03 in NEI. The other studies also find the similar correlation of about -0.54 between ENSO Niño 3.4 index and all-India ISMR over 1958–2010 (Sajani *et al* 2015) and about -0.2 to -0.6 between ENSO and ISMR for the period 1950–1999 (Li *et al* 2017).

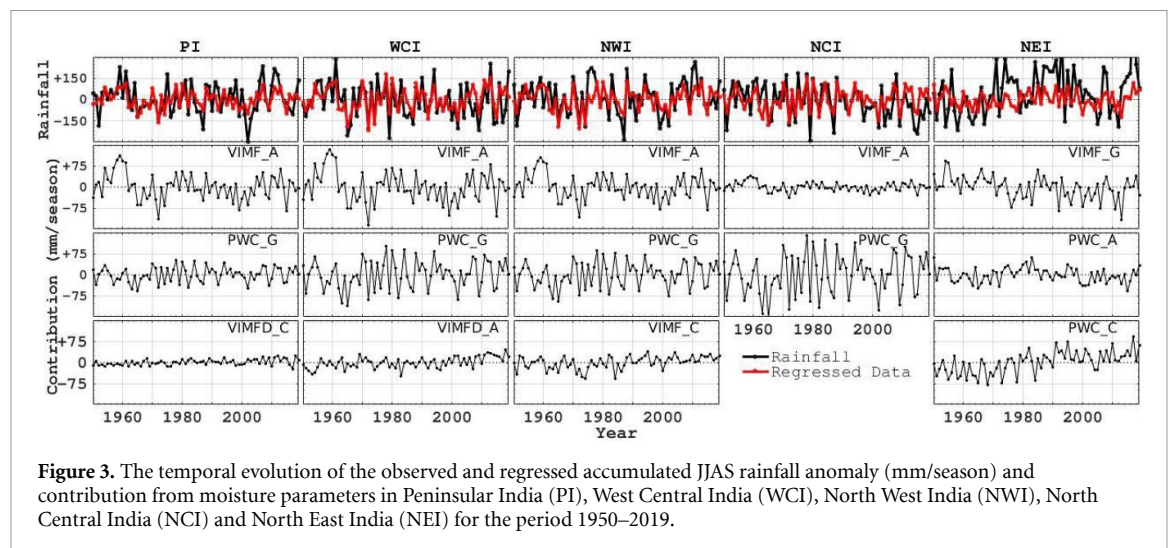
The statistically significant correlation of JJAS rainfall with VIMF and PWC is positive while that with VIMFD is negative in PI, WCI, NWI and NCI (except VIMF_G at 1000–850 hPa level in NWI). The VIMF over the Arabian Sea, VIMFD over Ganga river basin (except over NCI) and PWC over the Arabian Sea and Ganga river basin show statistically significant and strong correlations with rainfall in those regions. In addition, a significant correlation is shown by the Central Indian Ocean VIMF and PWC at 1000–300 hPa in NWI, and the central Indian Ocean VIMFD at 1000–850 hPa in PI. The Bay of Bengal PWC at 1000–300 hPa also shows a significant correlation in PI. The correlation coefficients of the moisture indices at 1000–300 hPa level are greater than those computed at 1000–850 hPa with rainfall and thus are used for MLR analysis in PI, WCI, NWI and NCI. The rainfall shows the highest correlation with the Ganga river basin VIMFD in PI (-0.54) and NWI (-0.64) whereas with Arabian Sea VIMF in WCI (0.65) and with Ganga river basin PWC in NCI (0.56) at 1000–300 hPa. These correlations are higher than that deduced between rainfall and MEI.

In NEI, rainfall shows a statistically significant correlation with the Bay of Bengal VIMF (-0.25) at 1000–300 hPa, the Bay of Bengal VIMF (-0.25), Arabian Sea VIMFD (0.29) and Ganga river basin PWC (0.25) at 1000–850 hPa. The moisture parameters integrated over 1000–850 hPa level are used as proxies as these provide better results in comparison to those at 1000–300 hPa level. Note that NEI is an exception here as the features of monsoon rainfall are opposite to that of other regions in India (Goswami *et al* 2010, Nair *et al* 2018).

Table 2 illustrates the correlation coefficients between moisture indices themselves at 1000–300 hPa level and also with MEI. In general, correlations between the proxies are small and insignificant, although statistically significant and strong correlations exist between VIMFD and PWC over the Arabian Sea (-0.64) and the Bay of Bengal (-0.83). Similarly, the correlation between the Arabian Sea VIMFD and the Bay of Bengal VIMFD is about 0.5 and that between the Ganga river basin VIMFD and the Arabian Sea PWC is about -0.56 . The MEI anticorrelates with VIMF and PWC and correlates with VIMFD. It shows statistically significant correlations with the Arabian Sea VIMF (-0.42), VIMFD ($+0.27$) and PWC (-0.44), Ganga river basin VIMF (-0.25) and PWC (-0.38), and central Indian Ocean VIMFD (0.30). Similarly, correlations between the moisture indices at 1000–850 hPa level along with MEI are given in table S1 (is available online

Table 1. The correlation coefficients between the accumulated JJAS rainfall and non-detrended explanatory parameters such as VIMF, VIMFD and PWC over the Arabian Sea (VIMF_A, VIMFD_A, PWC_A), Bay of Bengal (VIMF_B, VIMFD_B, PWC_B), central Indian Ocean (VIMF_C, VIMFD_C, PWC_C) and Ganga river basin (VIMF_G, VIMFD_G, PWC_G) at 1000–300 hPa and 1000–850 hPa level along with multivariate ENSO Index (MEI) in Peninsular India (PI), West Central India (WCI), North West India (NWI), North Central India (NCI) and North East India (NEI) for the period 1950–2019. The statistically significant correlations are given in bold numbers.

Parameters	PI	WCI	NWI	NCI	NEI	PI	WCI	NWI	NCI	NEI
MEI	−0.30	−0.39	−0.35	−0.54	0.03					
	1000–300 hPa					1000–850 hPa				
VIMF_A	0.53	0.65	0.54	0.50	−0.14	0.47	0.63	0.49	0.50	0.07
VIMF_B	0.01	0.21	−0.12	0.18	−0.25	−0.00	0.23	−0.11	0.17	−0.25
VIMF_C	0.20	0.19	0.32	0.06	−0.07	0.08	0.19	0.20	0.12	−0.06
VIMF_G	−0.15	−0.03	−0.19	0.01	−0.00	−0.16	−0.09	−0.27	0.11	0.06
VIMFD_A	−0.34	0.07	−0.17	0.01	0.21	−0.40	−0.10	−0.18	−0.16	0.29
VIMFD_B	0.04	−0.11	−0.00	−0.15	0.15	0.16	0.13	0.04	0.18	−0.11
VIMFD_C	0.18	0.05	0.00	−0.09	−0.04	0.28	0.07	0.15	−0.05	0.03
VIMFD_G	−0.54	−0.37	−0.64	−0.15	0.07	−0.53	−0.37	−0.62	−0.12	0.07
PWC_A	0.44	0.35	0.39	0.29	−0.21	0.43	0.24	0.41	−0.04	0.05
PWC_B	0.33	0.21	0.24	0.09	−0.16	0.15	0.07	0.13	−0.20	0.00
PWC_C	0.17	0.11	0.24	0.03	−0.17	0.10	−0.05	0.18	−0.21	0.09
PWC_G	0.44	0.63	0.57	0.56	0.18	0.34	0.52	0.50	0.39	0.25



at stacks.iop.org/ERL/16/014045/mmedia). The combination of highly correlated proxies (greater than 0.4) is avoided for MLR analysis. However, the influence of MEI on ISMR has been tested keeping the Arabian Sea VIMF in the model, even though these exhibit correlation of -0.42 , as Arabian Sea VIMF contributes significantly to the ISMR and therefore cannot be excluded (discussed in more detail below).

3.2. Contribution of moisture parameters

Figure 3 exhibits the temporal evolution of accumulated JJAS rainfall anomaly and regressed data (top panel) along with contributions from moisture parameters (second to bottom panels) at different regions for the period 1950–2019. The regression model reproduces the features of the rainfall anomaly quite well. The Arabian Sea VIMF and Ganga river basin PWC are common in PI, WCI, NWI and NCI, and these factors contribute mainly to the rainfall. The Arabian Sea VIMF follows a similar pattern with rainfall in the PI, WCI and NWI regions whereas

the Ganga river basin PWC follows the rainfall in NCI. The Ganga river basin PWC also shows a similar pattern with rainfall except in the 1960s in WCI and NWI. The contributions of central Indian Ocean VIMFD, Arabian Sea VIMFD and central Indian Ocean VIMF are comparatively small in the respective regions of PI, WCI and NWI. In NEI, Ganga river basin VIMF, Arabian Sea PWC and central Indian Ocean PWC contribute significantly to the rainfall. These three regression indices show large interannual variations compared to the parameters in other regions.

3.3. Variability of moisture content/transport/divergence associated with ISMR

Figure 4 (a) depicts the variability of atmospheric moisture parameters regressed with the accumulated JJAS rainfall anomaly in mm/season ($\sigma[\text{regression coefficient} \times \text{time series of proxy}]$) for the period 1950–2019. It explains the influence of parameters

Table 2. The correlation coefficients of the proxies such as multivariate ENSO Index (MEI), VIME, VIMFD and PWC over Arabian Sea (VIMF_A, VIMFD_A, PWC_A), Bay of Bengal (VIMF_B, VIMFD_B, PWC_B), central Indian Ocean (VIMF_C, VIMFD_C, PWC_C) and Ganga river basin (VIMF_G, VIMFD_G, PWC_G) at 1000–300 hPa level among themselves for the period 1950–2019. The correlations that are statistically significant at a 95% confidence interval are given in bold numbers.

Parameters	VIMF_A	VIMF_B	VIMF_C	VIMF_G	VIMFD_A	VIMFD_B	VIMFD_C	VIMFD_G	PWC_A	PWC_B	PWC_C	PWC_G	MEI
VIMF_A	1.00												
VIMF_B	0.12	1.00											
VIMF_C	0.14	−0.26	1.00										
VIMF_G	−0.12	−0.15	0.01	1.00									
VIMFD_A	−0.17	0.16	0.17	0.02	1.00								
VIMFD_B	−0.15	−0.28	0.18	0.22	0.50	1.00							
VIMFD_C	−0.08	0.17	0.13	0.03	0.30	0.26	1.00						
VIMFD_G	−0.42	0.15	−0.27	0.14	0.24	−0.03	−0.12	1.00					
PWC_A	0.47	−0.03	−0.08	0.04	−0.64	−0.32	−0.12	−0.56	1.00				
PWC_B	0.14	0.43	−0.04	−0.27	−0.40	−0.83	−0.05	−0.02	0.27	1.00			
PWC_C	−0.10	−0.41	0.47	0.01	0.11	0.35	−0.33	−0.20	−0.00	−0.29	1.00		
PWC_G	0.39	−0.11	0.01	0.09	0.00	0.13	0.09	−0.47	0.48	−0.14	0.04	1.00	
MEI	−0.42	0.22	−0.14	−0.25	0.27	0.17	0.30	0.21	−0.44	−0.11	−0.02	−0.38	1.00

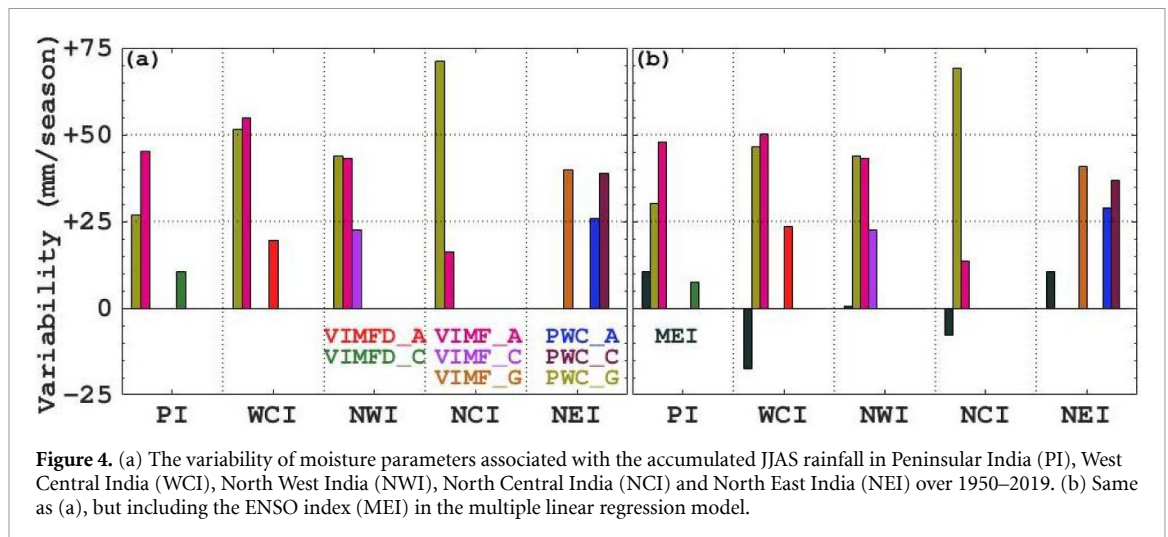


Figure 4. (a) The variability of moisture parameters associated with the accumulated JJAS rainfall in Peninsular India (PI), West Central India (WCI), North West India (NWI), North Central India (NCI) and North East India (NEI) over 1950–2019. (b) Same as (a), but including the ENSO index (MEI) in the multiple linear regression model.

on the interannual variability of rainfall. Among the moisture source regions, the Arabian Sea and Ganga river basin provide significant variability to rainfall in PI, WCI, NWI and NCI. In NEI, the Ganga river basin and the central Indian Ocean are predominant, and then the Arabian Sea. Similarly, PWC and VIMF explain the highest rainfall variability compared to VIMFD. The Arabian Sea VIMF and Ganga river basin PWC show the highest influence on rainfall variability in PI (45 mm/season) and NCI (70 mm/season), respectively. On the other hand, both parameters affect the rainfall equally in WCI (52–55 mm/season) and NWI (45 mm/season). The other parameters that influence rainfall variability are central Indian Ocean VIMFD, Arabian Sea VIMFD and central Indian Ocean VIMF in PI, WCI and NWI, respectively. NEI is entirely different from other regions where Ganga river basin VIMF and central Indian Ocean PWC contribute significantly, about 40 mm/season and then Arabian Sea PWC, about 25 mm/season. These analyses suggest the control of moisture parameters as well as moisture source regions to the JJAS rainfall variability.

3.4. Performance of the regression model

The performance of the regression model and the parameters are verified using several statistical tests mentioned earlier. The *t*-test and probability evaluated for the coefficients of moisture parameters regressed with the accumulated JJAS rainfall are given in table 3(a). The probability is less than 0.01 here, suggesting that parameters used in the model are all significant at a 99% confidence interval. Additionally, regression coefficients and their uncertainty ($2 \times$ standard deviation) are shown in table S2. The regression coefficients that are statistically significant at a 95% confidence interval are given in bold numbers. The Arabian Sea VIMF and Ganga river basin PWC are significant at a 95% confidence interval in PI, WCI and NWI. In NCI, Ganga river basin PWC is significant at 95% confidence interval whereas Arabian sea VIMF is significant at a 85% confidence interval.

The Arabian Sea VIMFD and Central Indian Ocean VIMF are significant at a 90% confidence interval in WCI and NWI, respectively. In NEI, Ganga river basin VIMF and central Indian Ocean PWC are significant at a 90% confidence interval.

Table 4(a) provides the performance of regression models as evaluated from various statistical methods such as R^2 , adjusted R^2 , RMSE, bias, *F*-statistic, probability, tolerance, VIF and DW-statistic. The R^2 values suggest that the regression model could explain rainfall variability of about 32%, 50%, 41%, 48% and 12% in PI, WCI, NWI, NCI and NEI, respectively. The adjusted R^2 values are comparable to the R^2 values; implying that the employed proxies enhance the performance of the model. The RMSE and bias of the model show small values of about 82–154 and $-6e-15$ to $+0.7e-15$ mm/season, respectively, indicating the good performance of the model. The high *F* values, greater than its critical value at a level of 0.05 are computed from the *F*-table, and the probabilities are less than 0.01. This points out that the model results are highly significant (at 99% confidence interval) and explain noticeable variance in the PI, WCI, NWI and NCI regions. The model is significant at a 95% confidence interval (probability ~ 0.03) in NEI. Note that the significance of model results increases as *F* value increases. The tolerance is greater than 0.1 and VIF is less than 2.5 in all regions; indicating that the model is free from multicollinearity problems. The DW statistic is also between 1.5 and 2.5 everywhere. These statistical tests corroborate that the model is robust for evaluating regional variability of monsoon rainfall and henceforth could be a helpful tool for forecasting ISMR. A study by Pandey *et al* (2020) also stated the importance of global warming mode along with ENSO in improving the skill of ISMR prediction models.

3.5. Intraseasonal variability of ISMR

The intraseasonal variability of ISMR has a major role in the interannual variability of ISMR and on

Table 3. The parameter estimates such as *t*-statistics and probability of the coefficients of regression indices used in the multiple linear regression model. The regression is performed for (a) the accumulated and (b) mean JJAS rainfall data, with moisture parameters, (c) accumulated JJAS rainfall with moisture parameters and ENSO and (d) accumulated JJAS rainfall with ENSO in Peninsular India (PI), West Central India (WCI), North West India (NWI), North Central India (NCI) and North East India (NEI) for the period 1950–2019.

(a) Accumulated JJAS rainfall—moisture parameters										
Predictands	PI		WCI		NWI		NCI		NEI	
Predictors	<i>t</i> -Value	<i>P</i> -value	<i>t</i> -Value	<i>P</i> -value	<i>t</i> -Value	<i>P</i> -value	<i>t</i> -Value	<i>P</i> -value	<i>t</i> -Value	<i>P</i> -value
Constant	−1.37	1.82	0.18	0.86	−0.55	1.42	0.05	0.96	6.94	0.00
PWC_G	51.08	0.00	97.85	0.00	83.45	0.00	135.99	0.00	—	—
VIMF_A	85.37	0.00	102.81	0.00	81.45	0.00	30.88	0.00	—	—
VIMF_C	—	—	—	—	46.21	0.00	—	—	—	—
VIMF_G	—	—	—	—	—	—	—	—	80.07	0.00
VIMFD_A	—	—	40.01	0.00	—	—	—	—	—	—
VIMFD_C	21.70	0.00	—	—	—	—	—	—	—	—
PWC_A	—	—	—	—	—	—	—	—	53.50	0.00
PWC_C	—	—	—	—	—	—	—	—	77.63	0.00
(b) Mean JJAS rainfall—moisture parameters										
Constant	−0.04	1.04	0.01	1.00	−0.02	1.01	0.00	1.00	0.23	0.82
PWC_G	1.68	0.10	3.22	0.00	2.73	0.01	4.46	0.00	—	—
VIMF_A	2.79	0.01	3.36	0.00	2.67	0.01	1.01	0.32	—	—
VIMF_C	—	—	—	—	1.50	0.14	—	—	—	—
VIMF_G	—	—	—	—	—	—	—	—	2.62	0.01
VIMFD_A	—	—	1.30	0.20	—	—	—	—	—	—
VIMFD_C	0.71	0.48	—	—	—	—	—	—	—	—
PWC_A	—	—	—	—	—	—	—	—	1.74	0.09
PWC_C	—	—	—	—	—	—	—	—	2.54	0.01
(c) Accumulated JJAS rainfall—moisture parameters and ENSO										
Constant	−4.74	2.00	6.47	0.00	−0.72	1.53	2.91	0.00	3.50	0.00
MEI	17.82	0.00	−30.71	2.00	0.92	0.36	−14.54	2.00	20.06	0.00
PWC_G	54.09	0.00	85.35	0.00	80.78	0.00	127.77	0.00	—	—
VIMF_A	86.85	0.00	90.14	0.00	77.96	0.00	24.68	0.00	—	—
VIMF_C	—	—	—	—	46.03	0.00	—	—	—	—
VIMF_G	—	—	—	—	—	—	—	—	81.69	0.00
VIMFD_A	—	—	46.39	0.00	—	—	—	—	—	—
VIMFD_C	14.47	0.00	—	—	—	—	—	—	—	—
PWC_A	—	—	—	—	—	—	—	—	57.07	0.00
PWC_C	—	—	—	—	—	—	—	—	72.95	0.00
(d) Accumulated JJAS rainfall—ENSO										
Constant	6.72	0.00	17.74	0.00	13.37	0.00	14.22	0.00	−0.71	1.52
MEI	−38.98	2.00	−102.92	2.00	−77.56	2.00	−82.52	2.00	4.09	0.00

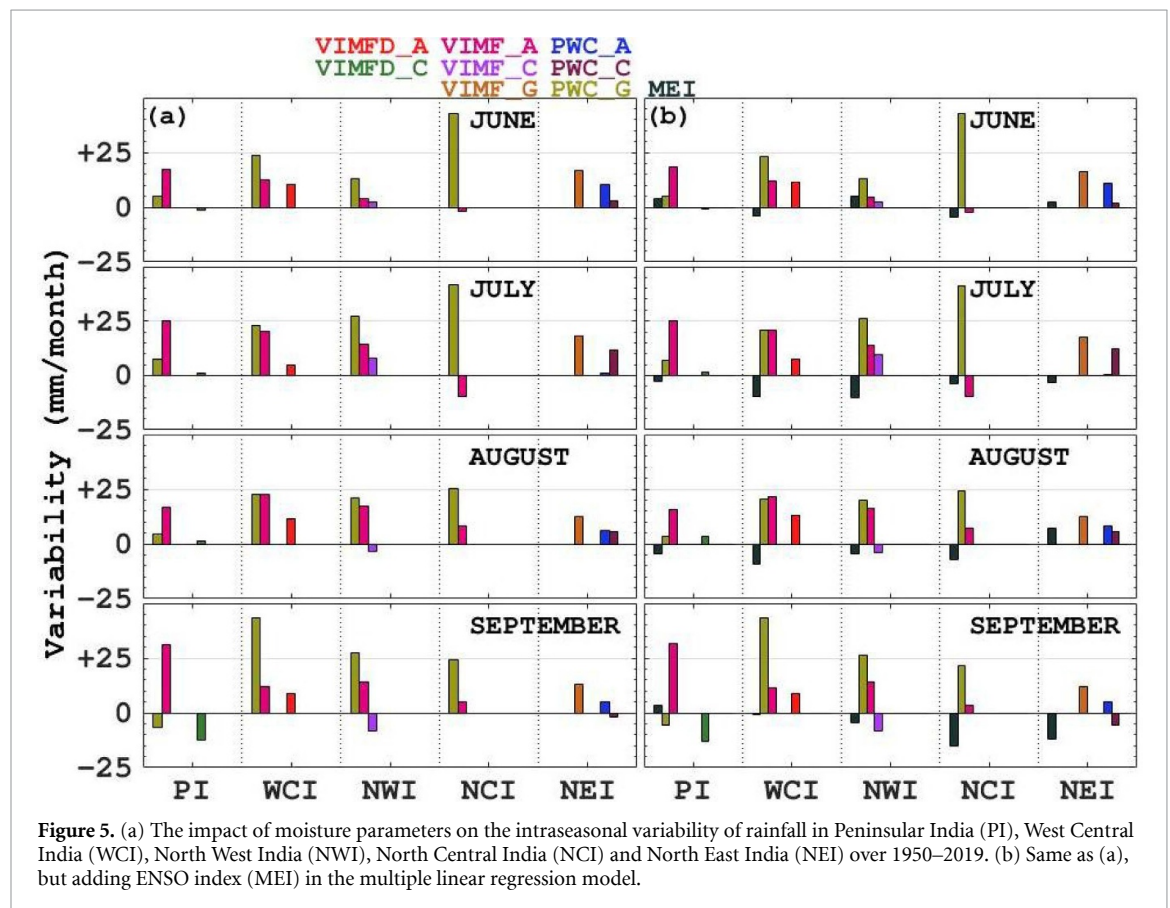
agriculture (Goswami *et al* 2006, Maharana and Dimri 2016). Therefore, the accumulated monthly rainfall anomaly time series is regressed using moisture parameters to explore their influence on the intraseasonal variability of rainfall. Figure 5 shows the impact of moisture parameters on the intraseasonal variability of monsoon rainfall in different regions.

In PI, monthly rainfall variability associated with the Arabian Sea VIMF increases from June to July, decreases in August and peaks in September. The impact of Ganga river basin PWC on intraseasonal rainfall variability remains the same throughout the season with an anticorrelation in September. The central Indian Ocean VIMFD affects rainfall variability only in September when it is anticorrelated with it. In WCI, the Ganga river basin PWC related rainfall variability is high and it peaks in September. The rainfall variability associated with the Arabian Sea VIMF increases from June to August and decreases in September. Similarly, the Arabian Sea

VIMFD shows a positive correlation throughout the season and its influence is highest in June and August. In NWI, all parameters show a small influence on rainfall variability in June and peak variability is exhibited by the Ganga river basin PWC in July and September. The Influence of Arabian Sea VIMF is highest in August. The central Indian Ocean VIMF shows a positive correlation in June and July, and a negative correlation in August and September. In NCI, the Ganga river basin PWC shows large variability which decreases from June to September. The influence of Arabian Sea VIMF is small with an anticorrelation in June and July, but a positive correlation in August and September. In NEI, the influence of the Ganga river basin VIMF is large and similar in June and July, about 16–18 mm/month and also in August and September, about 12–13 mm/month. The Arabian Sea and central Indian Ocean PWC show intraseasonal variability with their maxima in June and July, respectively in NEI. In general, large

Table 4. The summary of the regression model fit such as the coefficient of multiple determination (R^2 in %), adjusted R^2 ($A-R^2$) in %, root-mean-square error (RMSE), bias (Bias), F -statistic (F -stat), probability (Prob), tolerance (TOL), variance inflation factor (VIF) and Durbin–Watson statistic (DW-stat) estimated by regressing the (a) accumulated JJAS rainfall data with moisture parameters, (b) mean JJAS rainfall data with moisture parameters, (c) accumulated JJAS rainfall with moisture parameters and ENSO and (d) accumulated JJAS rainfall with ENSO for the period 1950–2019. The unit of RMSE and bias for the accumulated JJAS rainfall is mm/season and that for the mean JJAS rainfall is mm/day. The F -statistic, probability, tolerance, VIF and DW statistic are unitless quantities.

(a) Accumulated JJAS rainfall-moisture parameters						(b) Mean JJAS rainfall—moisture parameters				
Statistics	PI	WCI	NWI	NCI	NEI	PI	WCI	NWI	NCI	NEI
R^2	32.11	49.92	41.13	48.39	12.44	32.04	49.90	41.26	48.40	12.35
$A-R^2$	29.03	47.65	38.46	46.85	8.46	28.95	47.62	38.59	46.86	8.36
RMSE	91.57	90.80	95.07	82.70	154.36	0.75	0.74	0.78	0.68	1.26
Bias	$-1.7e-15$	$-2.2e-15$	$6.5e-16$	$-1.2e-15$	$-6.0e-15$	$2.7e-17$	$-1.7e-17$	$-3.4e-18$	$-1.2e-17$	$-6.7e-18$
F -stat	10.41	21.93	15.37	31.41	3.13	10.37	21.91	15.45	31.42	3.10
Prob	<0.001	<0.001	<0.001	<0.001	0.03	<0.001	<0.001	<0.001	<0.001	0.03
TOL	0.68	0.50	0.59	0.52	0.88	0.68	0.50	0.59	0.52	0.88
VIF	1.47	2.00	1.70	1.94	1.14	1.47	2.00	1.70	1.94	1.14
DW-stat	1.96	2.20	1.83	2.23	1.85	1.97	2.20	1.83	2.22	1.86
(c) Accumulated JJAS rainfall-moisture parameters & ENSO						(d) Accumulated JJAS rainfall—ENSO				
R^2	32.74	51.31	41.13	48.77	12.80	2.98	15.60	9.50	12.27	0.01
$A-R^2$	28.60	48.32	37.51	46.44	7.43	1.56	14.36	8.17	10.98	-1.46
RMSE	91.84	90.22	95.80	83.02	155.22	107.84	116.14	116.13	107.03	162.50
Bias	$-2.2e-15$	$-2.2e-15$	$4.4e-16$	$-2.4e-15$	$-1.7e-15$	$1.3e-15$	$3.8e-15$	$-1.3e-15$	$-5.2e-16$	$-2.9e-15$
F -stat	7.91	17.13	11.35	20.94	2.39	2.09	12.57	7.14	9.51	0.01
Prob	<0.001	<0.001	<0.001	<0.001	0.06	0.15	<0.001	0.01	<0.001	0.92
TOL	0.67	0.49	0.59	0.51	0.87	0.97	0.84	0.91	0.88	1.00
VIF	1.49	2.05	1.70	1.95	1.15	1.03	1.18	1.10	1.14	1.00
DW-stat	2.02	2.23	1.84	2.23	1.83	1.59	2.27	1.68	2.40	1.70



intraseasonal variability is shown by PWC compared to VIMF and VIMFD, and the central Indian Ocean among the source regions. Also, parameters tend to show anticorrelation mainly in September.

3.6. MLR analysis on mean ISMR

The mean ISMR is very important in forecasting rainfall and has a strong impact on agriculture. It is also demonstrated that the interannual variability is

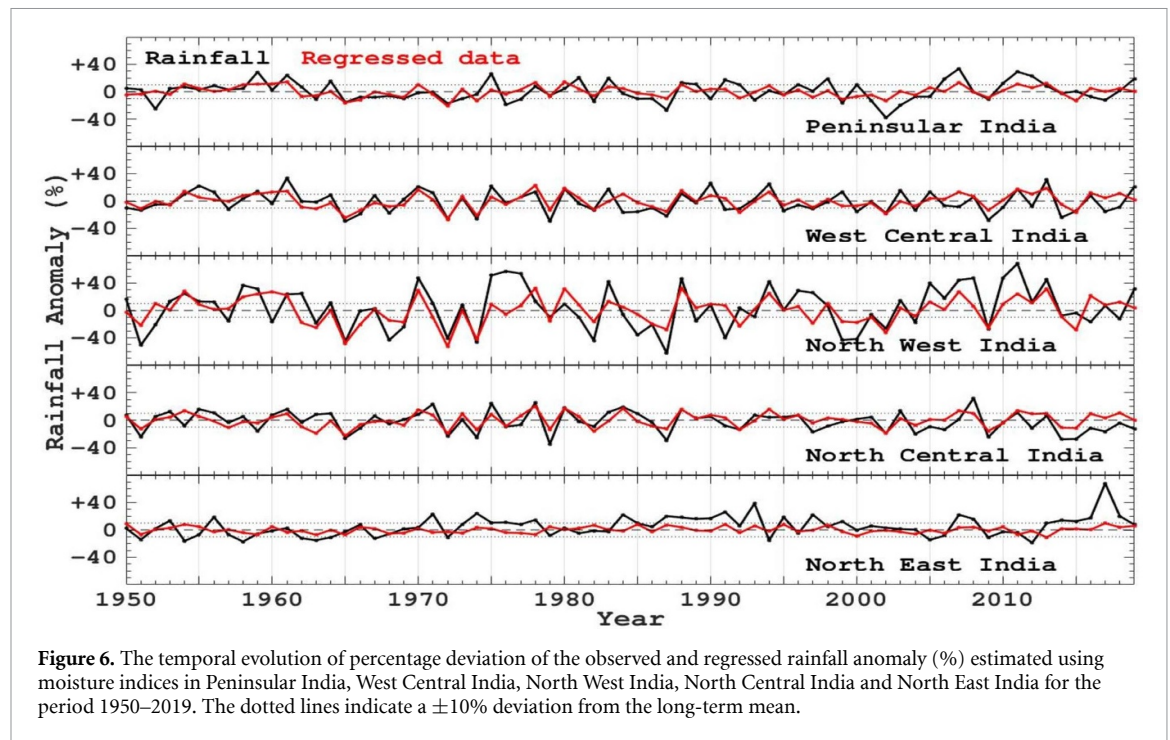


Figure 6. The temporal evolution of percentage deviation of the observed and regressed rainfall anomaly (%) estimated using moisture indices in Peninsular India, West Central India, North West India, North Central India and North East India for the period 1950–2019. The dotted lines indicate a $\pm 10\%$ deviation from the long-term mean.

decisive to predict mean ISMR (Goswami *et al* 2006). Here, mean JJAS rainfall is regressed with moisture indices at different regions for the period 1950–2019. The variability of moisture indices associated with the mean JJAS rainfall is shown in figure S4. The temporal evolution of the mean JJAS rainfall and regressed rainfall anomaly is shown in figure S5. The pattern of the observed and regressed anomaly of mean JJAS rainfall is similar to that of the accumulated JJAS rainfall as shown in figure 3. The statistical significance of the parameters and MLR model for the mean JJAS rainfall is given in tables 3(b) and 4(b), respectively. The difference is that only the dominant parameters like Arabian Sea VIMF and Ganga river basin PWC are significant at 95% confidence interval in PI, WCI and NWI. In NCI, only Ganga river basin PWC is significant, while in NEI, Ganga river basin VIMF and central Indian Ocean PWC are significant. Despite this, the model is significant at a 99% confidence interval in PI, WCI, NWI and NCI as the probability is less than 0.001 and it is significant at a 95% confidence interval in NEI. The computed RMSE and bias (in mm/day) are also very small. Note that as long as the pattern of input and proxy data is not changing, the model output remains the same. This strongly attests that the model is highly stable and very good in explaining regional variability of rainfall.

3.7. Percentage deviation of ISMR: detection of wet and dry years

Global warming enhances the moisture holding capacity of the atmosphere which in turn increases the frequency of extreme weather events (Mukherjee *et al* 2018). Therefore, the percentage departure of

JJAS rainfall anomaly is regressed with the moisture indices to test the applicability of the model in reproducing extreme events. The temporal evolution of regressed data along with observed rainfall anomaly for the period 1950–2019 is illustrated in figure 6. The dotted lines represent $\pm 10\%$ deviation from the long-term mean, a condition for determining wet and dry years. If a year meets the condition of anomaly greater (less) than 10 (-10)%, that year is considered as a wet (dry) year (Kumar *et al* 2013).

The analysis on regional average unveils a number of extreme events with total wet years ($>10\%$ anomaly) of about 19, 20, 31, 14 and 24, and dry years ($<-10\%$ anomaly) of about 20, 23, 24, 20 and 12 in PI, WCI, NWI, NCI and NEI, respectively. Out of these, the regression model reproduces about 6, 10, 15, 7 and 0 wet years and 11, 16, 15, 12 and 3 dry years in the respective regions. The model could explain about 32%, 50%, 48%, 50% and 4% of wet and 55%, 70%, 63%, 60% and 25% of dry years in PI, WCI, NWI, NCI and NEI, respectively. The performance of the model in reproducing extreme rainfall events is calculated using a number of hits (H), false alarms (F) and misses (M), false alarm rate (FAR) and the probability of detection (POD). If both measurements and model satisfy the condition of the extreme rainfall event, it is treated as a hit. If the model does not capture the observed extreme event, it is considered as a miss. The false alarm is that the model shows an extreme event that is not observed (Ashrit *et al* 2015, Sofiati and Nurlatifah 2019, McBride and Ebert 2000). The FAR is the ratio of a number of false alarms to the number of forecasts ($H+F$) and POD is the ratio of a number of hits to the number of observed events ($H+M$). The FAR should be low

Table 5. (a) The number of hits (H), false alarms (F), misses (M), false alarm rate (FAR) and the probability of detection (POD) of wet/dry years deduced from the accumulated JJAS rainfall anomaly regressed using moisture parameters in Peninsular India (PI), West Central India (WCI), North West India (NWI), North Central India (NCI) and North East India (NEI) for the period 1950–2019. (b) Same as (a), but regression is performed with moisture and ENSO indices.

(a) Moisture parameters											(b) Moisture parameters and ENSO									
Region	Wet					Dry					Wet					Dry				
	H	F	M	FAR	POD	H	F	M	FAR	POD	H	F	M	FAR	POD	H	F	M	FAR	POD
PI	5	7	14	0.58	0.26	11	3	9	0.21	0.55	5	6	14	0.55	0.26	10	4	10	0.29	0.50
WCI	10	6	10	0.38	0.50	16	1	7	0.06	0.70	10	4	10	0.29	0.50	16	1	7	0.06	0.70
NWI	15	5	16	0.25	0.48	15	7	9	0.32	0.62	15	4	16	0.21	0.48	15	7	9	0.32	0.62
NCI	6	5	8	0.45	0.43	12	3	8	0.20	0.60	6	8	8	0.57	0.43	12	4	8	0.25	0.60
NEI	0	0	24	NaN	0.00	3	1	9	0.25	0.25	0	0	24	NaN	0.00	3	1	9	0.25	0.25

(~ 0), whereas POD should be high (~ 1) for a model to be good (McCann 1983).

Table 5(a) shows H, F, M, FAR and POD computed for the wet and dry years. The estimated FAR is about 0.25–0.45 and POD is about 0.43–0.50 for wet years in WCI, NWI and NCI. In PI, the number of hits is lower than that of misses and so FAR is a little bit high (0.58) and POD is low (0.26) for wet years. In general the number of hits is greater than the number of misses for dry years in PI, WCI, NWI and NCI, indicating that the model could reproduce a good number of dry years there. The FAR ranges from 0.06 to 0.32 and POD varies from 0.55 to 0.70 in these regions. On the other hand, the model could not explain well wet and dry years in NEI. Therefore, the developed model is good in explaining extreme events in West Central, North West and North Central India. The Peninsular and North East India demand more proxies to better interpret the rainfall variability. A study by Rajeevan *et al* (2007) showed zero false alarm and 77%–100% of POD for the models computed for the period 1981–2004. Similarly, Sharma *et al* (2017) presented FAR of about 0.5–0.8 and POD of 0.3 for the period 2007–2015. Another study by Pandey *et al* (2015) presented FAR of about 0.14–0.63 and POD of about 0.33–0.60 over 1982–2013.

The importance is that the model could clearly reproduce severe droughts of 2002, 2004, 2014 and 2015 everywhere except 2002 in PI, 2014 in WCI and 2004 in NCI. Therefore, an attempt is made to find out a condition for determining wet and dry years from moisture indices used in the model. It is found that the indices stay within ± 0.5 in normal years, but exceed this value during extreme events, particularly the indices of VIMF over the Arabian Sea and PWC over the Ganga river basin. A condition is henceforth drawn for finding extreme events from the moisture parameters such that if the index based on Ganga river basin PWC or Arabian Sea VIMF is greater (less) than $+0.5$ (-0.5); then, it will be a wet (dry) year. These indices well satisfy this condition for wet/dry years and even exceed $+1$ (-1) during extreme flood (drought) years; demonstrating the effectiveness and relevance of new indices. It attests to the applicability and potential of our statistical model based on

moisture-related parameters. The observed wet and dry years that are reproduced by the model and corresponding moisture indices are shown in table S3. A study by Wang *et al* (2015) reported that the IMD could not forecast extreme events in 1994, 2002, 2004 and 2009 using statistical models for 1989–2012 and confirms that this failure is due to the lack of proxies regarding global warming in the model.

3.8. Influence of ENSO on ISMR

ENSO is considered as a major driver for explaining the interannual variability of ISMR even though their relationship was weakening after the 1980s. Therefore, the influence of ENSO on moisture parameters and thus on the ISMR is analyzed regressing the accumulated JJAS rainfall anomaly time series with MEI and moisture indices. The statistical significance of parameters estimated using *t*-test is shown in table 3(c) and statistical results of the model are provided in table 4(c) for various regions. The statistical significance of moisture parameters remains intact with the addition of MEI. MEI is significant only in PI and NEI whereas the model is significant in all regions. Ideally in the climate data analysis, although a parameter is not statistically significant, it can be used if it improves the model (<https://statisticsbyjim.com/regression/model-specification-variable-selection/>). The multicollinearity statistics (tolerance and VIF) are also inside the favorable limit. The MEI improved the model with an increase of R^2 and adjusted R^2 by $\sim 2\%$ in WCI, whereas slightly lower adjusted R^2 and *F*-statistic values are estimated in other regions. The rainfall variability associated with MEI and moisture indices is exhibited in Figure 4(b). As provided in the statistical analysis results, MEI does not affect the rainfall variability imposed by the moisture indices although slight changes are observed. Note that the nature of the ENSO index is reversed; it correlates with rainfall in PI and NEI, and is least affected in NWI. The individual influence of ENSO on accumulated JJAS rainfall is also analyzed. The statistical significance of MEI and the corresponding model is provided in tables 3(d) and 4(d), respectively. MEI is significant only in NEI while the model is significant in WCI,

NWI and NCI. The R^2 is very low of about 2.98%, 15.6%, 9.5%, 12.3% and 0.01% in PI, WCI, NWI, NCI and NEI, respectively. It categorically points out that the ISMR variability cannot be explained only with ENSO.

Furthermore, the combined impact of ENSO and moisture parameters on the intraseasonal variability of ISMR is analyzed regressing the accumulated monthly rainfall anomaly from June to September and is demonstrated in figure 5(b). Here, the behavior of moisture indices remains intact, albeit with small changes in absolute values. In general, MEI shows anticorrelation with rainfall throughout the season; however, a positive correlation is found mainly in PI (June and September), NWI (June) and NEI (June and August), suggesting large regional variability. The influence of MEI is similar throughout the season in PI. It is high in July and August in WCI. Similarly, the influence of MEI peaks in July and then decreases in NWI while it increases from July to September in NCI. In NEI, the influence of MEI is high in September. Although the interannual variability is not influenced by ENSO, the intraseasonal variability of rainfall shows a profound influence in NWI. In short, ENSO influences rainfall in July and August in Peninsular, western central and northwestern parts of India while in September in the central and eastern parts of North India.

The wet and dry years are also examined using a model developed from moisture parameters and ENSO, and the resulting FAR, POD, number of hits, false alarms and misses are shown in table 5(b). In the case of wet years, the number of hits, misses and hence the POD is similar to that evaluated from the model based on moisture parameters. However, the number of false alarms and FAR is lower in PI, WCI and NWI while the number of false alarms and FAR is a bit higher in NCI compared to the model formulated on moisture parameters. In the case of dry years, F, H, M, FAR and POD are similar to that found in a model derived from moisture parameters in WCI, NWI, NCI and NEI. In PI, the number of hits is lower by one while the number of false alarms and misses is higher by one and therefore FAR is slightly higher and POD is lower for the model including ENSO.

4. Conclusion

The MLR analysis uncovers the application of atmospheric moisture parameters such as moisture content, moisture transport and its divergence over the moisture source regions of the Arabian Sea, central Indian Ocean and Ganga river basin as proxies of ISMR. The regression model is made from statistically significant moisture parameters that improve model performance. Therefore, indices related to the Bay of Bengal are not used as these are statistically insignificant. The regression is carried out for the seasonal,

monthly, mean and percentage deviation of southwest monsoon rainfall.

The moisture indices provide statistically significant and strong correlations (greater than 0.4) with ISMR. The regression model could explain regional rainfall variability of about 12%–50% for the period 1950–2019. The contributions of moisture indices to JJAS rainfall vary in various regions. Among the moisture source regions, the Arabian Sea and Ganga river basin are the largest contributors to the regional distribution of rainfall. Similarly, moisture content and its transport mainly decide the amount of rainfall in all regions. The moisture transport is prevailing in PI and moisture content is dominant in NCI while both affect equally in West Central, North West and North East India. The robustness of the developed model is checked using a number of statistical tests and the model fulfills the conditions for being a good model. The F -values are highly significant (at 99% confidence interval) at four regions implying that the model explains regional rainfall variability very well. The model could explain the intraseasonal rainfall variability and the variability of mean JJAS rainfall too.

The regression model could reproduce about 6 (11), 10 (16), 15 (15), 7 (12) and 0 (3) wet (dry) years out of the 19 (20), 20 (23), 31 (24), 14 (20) and 24 (12) observed wet (dry) years in PI, WCI, NWI and NCI, respectively. The model explains about 32%–50% of drought years, especially 2002, 2004, 2014 and 2015, and about 55%–70% of flood years occurred in those regions. In addition, the atmospheric moisture indices based on Arabian Sea VIMF and Ganga river basin PWC can be used for detecting wet and dry years such that their index is greater than 0.5 for a wet year and less than -0.5 for a dry year and exceed $+1$ (-1) during extreme flood (drought) years. The model shows a FAR of 0.25–0.58 (0.06–0.32) and POD of 0.26–0.50 (0.55–0.70) for wet (dry) years in all regions except in North East India and that demands careful evaluation considering other relevant proxies in the model.

ENSO shows a good correlation with ISMR, though it is smaller than the correlation between ISMR and moisture indices. The regression analysis reveals that the ISMR variability associated with the moisture indices is unchanged in the presence of the ENSO index while ENSO shows a positive correlation in Peninsular and North East India and a negative correlation in other regions. The ENSO index improved the model by 2% in WCI. ENSO shows a strong influence on the intraseasonal variations though it is not affecting the seasonal variability of rainfall in North West India. It mainly affects the July and August rainfall in Peninsular, West Central and North West India while September rainfall in North Central and North East India. The extreme events evaluated from the model based upon moisture parameters and ENSO are almost similar to that deduced from the model created on moisture parameters.

Henceforth, this study demonstrates the significance of moisture content, its transport and divergence on regional rainfall distribution. This study recommends that these parameters can be used in both statistical and dynamical models to better predict ISMR. It also attests to the importance of local factors in explaining ISMR, as the local factors are affected by the change in global factors. The new indices made from the atmospheric moisture parameters can be employed as proxies for climate change predictions and can be used together with the commonly used parameter of ENSO for improving the prediction of ISMR.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

The first author acknowledges Department of Science & Technology, Government of India for financial support vide reference number SR/WOS-A/EA-11/2018(G) under Women Scientist Scheme to carry out this work and the Indian Institute of Technology Kharagpur for providing infrastructure facilities. The specific humidity, zonal and meridional wind data are taken from the NCEP Reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at <https://www.esrl.noaa.gov/psd/>. The authors thank IMD for rainfall data. MEI is taken from <https://psl.noaa.gov/data/climateindices/> list.

ORCID iD

P J Nair  <https://orcid.org/0000-0002-7768-5154>

References

- Annamalai H and Liu P 2005 Response of the Asian summer monsoon to changes in El Niño properties *Q. J. R. Meteorol. Soc.* **131**, 805–31
- Ashok K, Guan Z and Yamagata T 2001 Impact of the Indian Ocean dipole on the relationship between the Indian monsoon rainfall and ENSO *Geophys. Res. Lett.* **26** 4499–502
- Ashrit R, Sharma K, Dube A, Iyengar G, Mitra A and Rajagopal E N 2015 Verification of short range forecasts of extreme rainfall during monsoon *Mausam* **66** 375–86
- Chakraborty A, Behera S K, Mujumdar M, Ohba R and Yamagata T 2006 Diagnosis of tropospheric moisture over Saudi Arabia and influences of IOD and ENSO *Mon. Weather Rev.* **134** 598–617
- Chattopadhyay R, Phani R, Sabeerali C T, Dhakate A R, Salunke K D, Mahapatra S, Rao A S and Goswami B N 2015 Influence of extratropical sea-surface temperature on the Indian summer monsoon: an unexplored source of seasonal predictability *Q. J. R. Meteorol. Soc.* **141** 2760–75
- Draper N R and Smith H 2015 *Applied Regression Analysis* 3rd edn (Hoboken, NJ: Wiley)
- Durbin J and Watson W S 1950 Testing for serial correlation in least squares regression, I *Biometrika* **37** 409–28
- Durbin J and Watson W S 1951 Testing for serial correlation in least squares regression, II *Biometrika* **38** 159–77
- Fasullo J and Webster P J 2002 A hydrological definition of Indian monsoon onset and withdrawal *J. Clim.* **16** 3200–11
- Feifei L, Li S and Furevik T 2011 The connection between the Atlantic multidecadal oscillation and the Indian summer monsoon in Bergen climate model version 2.0 *J. Geophys. Res.* **116**, D19
- Gadgil S and Francis P A 2016 El Niño and the Indian rainfall in June *Curr. Sci.* **110** 1010–22
- Garfinkel C I, Hurrell M M, Waugh D W and Butler A H 2013 Are the teleconnections of Central Pacific and Eastern Pacific El Niño distinct in Boreal wintertime? *Clim. Dyn.* **41** 1835–52
- Gautam N and Pandey P C 1995 Fluctuations in satellite derived moisture fields over the Indian ocean and their possibilities *Indian J. Radio Space Phys.* **24** 107–12
- Gopalapillai P 2012 Evolution of stratospheric ozone in the mid-latitudes in connection with the abundances of halogen compounds *Atmospheric and Oceanic Physics [physics.ao-ph]* Université Pierre et Marie Curie—Paris VI, English
- Goswami B B, Mukhopadhyay P, Mahanta R and Goswami B N 2010 Multiscale interaction with topography and extreme rainfall events in the northeast Indian region *J. Geophys. Res.* **115** D12114
- Goswami B N, Madhusoodanan M S, Neema C P and Sengupta D 2006 A physical mechanism for North Atlantic SST influence on the Indian summer monsoon *Geophys. Res. Lett.* **33** L02706
- Hrudya P H, Varikoden H, Vishnu R and Kuttippurath J 2020 Changes in ENSO-monsoon relations from early to recent decades during onset, peak and withdrawal phases of Indian summer monsoon *Clim. Dyn.* **55** 1457–71
- Kalnay E et al 1996 The NCEP/NCAR 40-year reanalysis project *Bull. Am. Meteor. Soc.* **77** 437–70
- Krishna Kumar K, Rajagopalan B and Cane M A 1999 On the weakening relationship between the Indian monsoon and ENSO *Science* **284** 5423, 2156–9
- Krzywinski M and Altman N 2015 Multiple linear regression *Nat. Methods* **12** 1103–4
- Kucharski F, Bracco A, Yoo J H and Molteni F 2008 Atlantic forced component of the Indian monsoon interannual variability *Geophys. Res. Lett.* **35** L04706
- Kumar K N, Rajeevan M, Pai D S, Srivastava A K and Preethi B 2013 On the observed variability of monsoon droughts over India *Weather Clim. Extremes* **1** 42–50
- Li Z, Lin X and Cai W 2017 Realism of modelled Indian summer monsoon correlation with the tropical Indo-Pacific affects projected monsoon changes *Sci. Rep.* **7** 4929
- Lin F 2008 Solving multicollinearity in the process of fitting regression model using the nested estimate *Procedure Qual Quant.* **42** 417–26
- Luis A J and Pandey P C 2004 Relationship between surface atmospheric convergence over Indian Ocean and Indian rainfall *Geophys. Res. Lett.* **31** L06208
- Maharana P and Dimri A P 2016 Study of intraseasonal variability of Indian summer monsoon using a regional climate model *Clim. Dyn.* **46** 1043–64
- McBride J L and Ebert E E 2000 Verification of quantitative precipitation forecasts from operational numerical weather prediction models over Australia *Am. Met. Soc.* **15** 103–21
- McCann D W 1983 The enhanced-V: a satellite observable severe storm signature *Mon. Weather Rev.* **111** 887–94
- Mei R, Ashfaq M and Rastogi D 2015 Dominating controls for wetter south Asian summer monsoon in the twenty-first century *J. Clim.* **28** 3400–19
- Mishra V, Smoliak B V, Lettenmaier D P and Wallace J M 2012 A prominent pattern of year-to-year variability in Indian summer monsoon rainfall *PNAS* **109** 7213–17

- Mukherjee S, Aadhar S, Stone D and Mishra V 2018 Increase in extreme precipitation events under anthropogenic warming in India *Weather Clim. Extremes* **20** 45–53
- Nair P J, Chakraborty A, Varikoden H, Francis P A and Kuttippurath J 2018 The local and global climate forcings induced inhomogeneity of Indian rainfall *Sci. Rep.* **8** 6026
- Pai D S 2004 A possible mechanism for the weakening of El Niño-monsoon relationship during the recent decade *Meteorol. Atmos. Phys.* **86** 143–57
- Pandey D K, Rai S, Shahi N K and Mishra N 2015 Seasonal prediction of ISMR and relationship with EL-NINO and IOD in ECMWF system 4 coupled model *Clim. Change* **1** 447–55
- Pandey P, Dwivedi S, Goswami B N and Kucharski F 2020 A new perspective on ENSO-Indian summer monsoon rainfall relationship in a warming environment *Clim. Dyn.* **55** 3307–26
- Pant G B and Kumar K R 1997 *Climates of South Asia* (Chichester, UK: John Wiley and Sons) p 320
- Pathak A, Ghosh S, Martinez J A, Dominguez F and Kumar P 2017 Role of oceanic and land moisture sources and transport in the seasonal and interannual variability of summer monsoon in India *J. Clim.* **30** 1839–59
- Peixoto J and Oort A H 1983 *The Atmospheric Branch of the Hydrological Cycle and Climate: Variations in the Global Water Budget* (Springer Verlag GmbH: Springer Netherlands)
- Pottapinjala V, Girishkumar M S, Ravichandran M and Murtugudde R 2014 Influence of the Atlantic zonal mode on monsoon depressions in the Bay of Bengal during boreal summer *J. Geophys. Res. Atmos.* **119** 6456–69
- Press W H, Flannery B P, Teukolsky S A and Vetterling W T 1989 *Numerical Recipes* (Cambridge, UK: Cambridge University Press) pp 504–8
- Puviarasan N, Sharma A K, Ranalkar M and Giri R K 2015 Onset, advance and withdrawal of southwest monsoon over Indian subcontinent: a study from precipitable water measurement using ground based GPS receivers *J. Atmos. Sol.-Terr. Phys.* **122** 45–57
- Rajeevan M, Pai D S, Anil Kumar R and Lal B 2007 New statistical models for long-range forecasting of southwest monsoon rainfall over India *Clim. Dyn.* **28** 813–28
- Ramakrishna S S V S, Rao B V, Srinivasa Rao B R, Hari Prasad D, Nanaji Rao N and Panda R 2017 A study of 2014 record drought in India with CFSv2 model: role of water vapor transport *Clim. Dyn.* **49** 297–312
- Ramesh Kumar M R, Shenoi S S C and Schluessel P 1999 On the role of the cross equatorial flow on summer monsoon rainfall over India using NCEP/NCAR reanalysis data *Meteorol. Atmos. Phys.* **70** 201–13
- Seetha C J, Varikoden H and Babu C A 2020 Significant changes in the ENSO-monsoon relationship and associated circulation features on multidecadal timescale *Clim. Dyn.* **54** 1491–506
- Senaviratna N A M R and Cooray T M J A 2019 Diagnosing multicollinearity of logistic regression model *Asian J. Probab. Stat.* **5** 1–9 AJPAS.51693
- Sharma K, Ashrit R, Bhatla R, Mitra A K, Iyengar G R and Rajagopal E N 2017 Skill of predicting heavy rainfall over India: improvement in recent years using UKMO global model *Pure Appl. Geophys.* **174** 4241–50
- Shukla J and Misra B M 1977 Relationships between sea surface temperature and wind speed over the central Arabian Sea and monsoon rainfall over India *Mon. Weather Rev.* **105** 998–1002
- Sofiati I and Nurlatifah A 2019 The prediction of rainfall events using WRF (weather research and forecasting) model with ensemble technique *Conf. Series: Earth and Environmental Science* vol 374 p 012036
- Surendran S, Gadgil S, Francis P A and Rajeevan M 2015 Prediction of Indian rainfall during the summer monsoon season on the basis of links with equatorial Pacific and Indian Ocean climate indices *Environ. Res. Lett.* **10** 9
- Ullah K and Gao S 2012 Moisture transport over the Arabian Sea associated with summer rainfall over Pakistan in 1994 and 2001 *Adv. Atmos. Sci.* **29** 501
- Vishnu S, Francis P A, Shenoi S S C and Ramakrishna S S V S 2016 On the decreasing trend of the number of monsoon depressions in the Bay of Bengal *Environ. Res. Lett.* **11** 014011
- Vishnu S, Francis P A, Shenoi S S C and Ramakrishna S S V S 2018 On the relationship between the pacific decadal oscillation and monsoon depressions over the Bay of Bengal *Atmos. Sci. Lett.* **19** e825
- Wang B, Xiang B, Li J, Webster P J, Rajeevan M N, Liu J and Ha K-J 2015 Rethinking Indian monsoon rainfall prediction in the context of recent global warming *Nat. Commun.* **6** 7154
- Yun K-S and Timmermann A 2018 Decadal monsoon-ENSO relationships reexamined *Geophys. Res. Lett.* **45** 2014–21
- Zhang R and Delworth T L 2006 Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes *Geophys. Res. Lett.* **33** L17712