Weather and Climate Extremes 12 (2016) 43–68

Contents lists available at ScienceDirect

Weather and Climate Extremes

journal homepage: www.elsevier.com/locate/wace

Trends and variability of droughts over the Indian monsoon region

Ganeshchandra Mallya [a], Vimal Mishra [b], Dev Niyogi [c,*], Shivam Tripathi [d], Rao S. Govindaraju [a]

[a] School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA
[b] Department of Civil Engineering, Indian Institute of Technology, Gandhinagar, Gujarat 382424, India
[c] Department of Agronomy and Department of Earth, Atmospheric and Planetary Sciences, Purdue University, West Lafayette, IN 47907, USA
[d] Department of Civil Engineering, Indian Institute of Technology, Kanpur, UP 208016, India

ARTICLE INFO

Article history:
Received 7 December 2015
Accepted 29 January 2016
Available online 15 February 2016

Keywords:
Drought index
Hidden Markov model
Standard precipitation index
Gaussian mixture model
Indian monsoon
Uncertainty analysis
Drought vulnerability

ABSTRACT

Drought characteristics for the Indian monsoon region are analyzed using two different datasets and standard precipitation index (SPI), standardized precipitation-evapotranspiration index (SPEI), Gaussian mixture model-based drought index (GMM-DI), and hidden Markov model-based drought index (HMM-DI) for the period 1901–2004. Drought trends and variability were analyzed for three epochs: 1901–1935, 1936–1971 and 1972–2004. Irrespective of the dataset and methodology used, the results indicate an increasing trend in drought severity and frequency during the recent decades (1972–2004). Droughts are becoming more regional and are showing a general shift to the agriculturally important coastal south-India, central Maharashtra, and Indo-Gangetic plains indicating higher food security and socioeconomic vulnerability in the region.

1. Introduction

Droughts in the monsoon dominated regions have gained greater importance in the recent past, as monsoons not only define the unique features of the climate, but also affect the socio-economic well-being of more than two third of global population (Niranjan Kumar et al., 2013; Rajeevan et al., 2008). Recent changes in Indian monsoon precipitation have received wide attention (Kripalani et al., 2003; Mishra et al., 2012; Rupa Kumar et al., 2006) with some plausible uncertainty on whether trends associated with summer monsoon precipitation are related to global warming or those due to regional changes (Chung and Ramanathan, 2006; Kishtawal et al., 2010; Niyogi et al., 2010). A number of studies (Kumar et al., 1992; Rajeevan et al., 2008; Stephenson, 2001) have indicated that the mean precipitation during the monsoon season may be unaltered over the Indian monsoon region (IMR), however the extreme precipitation events have shown statistically significant increasing trends in last five decades resulting in modification of drought characteristics over IMR (Goswami et al., 2006; Mishra et al., 2012). Trends associated with the Indian summer monsoon rainfall (ISMR) have also shown a great regional variability where some parts of India have seen an increase in precipitation while others show a reduction in precipitation during the monsoon season (Guhathakurta and Rajeevan, 2008; Niyogi et al., 2010; Roxy et al., 2015). Significant interannual, decadal and long term trends have been observed in the monsoon drought time series over IMR influenced by El Nino Southern Oscillation and global warming (Niranjan Kumar et al., 2013).

Recently, contrasting conclusions were drawn about global drought climatology by two synthesis studies (Sheffield et al., 2012; Dai, 2013). While Sheffield et al. (2012) showed that there was little change in drought climatology in recent years, the study by Dai (2013) concluded that droughts were intensifying as a result of a warming climate. Building off these assessments, Trenberth et al. (2014) summarized that the choice of precipitation dataset and other forcing datasets could influence drought analysis in addition to the choice of model parameterizations being used in deriving the drought indices [e.g. potential evapotranspiration calculations while estimating PDSI as reported in Sheffield et al. (2012)]. These studies highlight the need for using multiple drought indices and datasets for drought climatology, and form the basis for reassessing the drought of the Indian Monsoon Region.

Evaluation of trends and variability associated with retrospective drought events provides a basis to understand regional patterns of severity, duration, and areal extent of droughts. It also enables an understanding of the nature of possible future droughts and potential vulnerabilities. Building off the findings of drought assessments over the IMR in recent years and the recommendations cited in Trenberth et al. (2014) the aims of this paper are

* Corresponding author.
E-mail address: climate@purdue.edu (D. Niyogi).

http://dx.doi.org/10.1016/j.wace.2016.01.002

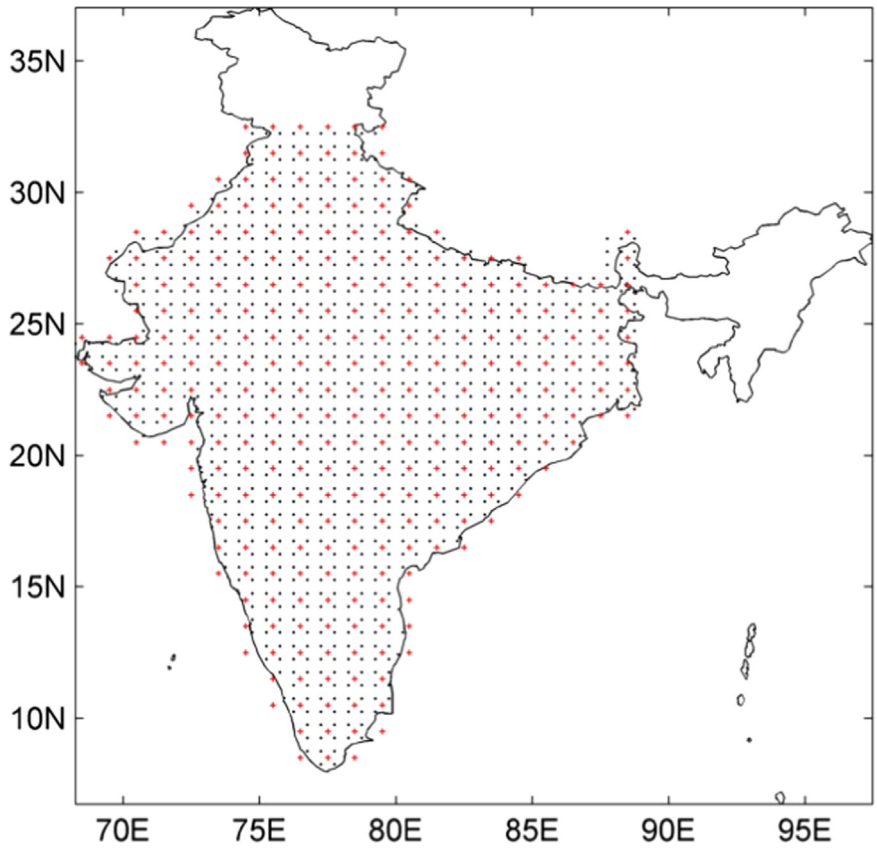


Fig. 1. Study domain showing 1° grid cell locations for India Meteorological Department precipitation dataset as cross-hairs, and 0.5° grid cell locations for University of Delaware precipitation dataset as dots.

(i) to study the retrospective droughts and associated trends over IMR using different precipitation datasets and drought indices, and (ii) to identify regions in IMR that are vulnerable to droughts.

2. Data and methods

We used gridded daily precipitation data from the India Meteorological Department (IMD) (Rajeevan, 2006) available for the period 1901–2004 at 1° spatial resolution (Fig. 1). The daily precipitation data obtained from IMD was then aggregated over monthly time scale. The second dataset used in this study was monthly precipitation data from University of Delaware (UD) available for the period of 1900–2004 (UDel_AirT_Precip data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at ⟨http://www.esrl.noaa.gov/psd/⟩) at 0.5° spatial resolution (Fig. 1). The precipitation data from high-mountainous regions in northern and northeastern parts of the country were not used in the study.

Despite the differences in the spatial resolution, the precipitation datasets show similar patterns in the spatial distribution and variance of precipitation over the study region. Fig. A.1a and b shows the distribution of mean monthly precipitation over the study region, and Fig. A.1c and d compares the standard deviation in monthly mean precipitation between the two datasets. While the overall patterns are similar, the effects of resolution on the magnitudes are evident. For instance, the UD dataset provides more detail in the spatial distribution of precipitation statistics; and a comparison of monthly mean precipitation time series (Fig. A.2) between the two datasets shows that while the overall monthly time series pattern are similar, the precipitation magnitude for IMD grids are lower compared to UD grids during the months June to September, and relatively identical for the remaining months.

Standardized precipitation index (SPI; McKee et al., 1993), standardized precipitation-evapotranspiration index (SPEI; Vicente-Serrano et al., 2010; Niranjan Kumar et al., 2013), Gaussian mixture model-based drought index (GMM-DI; Mallya, 2011), and hidden Markov model-based drought index (HMM-DI; Mallya, 2011; Mallya et al., 2012) were calculated for drought characterization at multiple time scales ending in September (i.e. for 1-month, 4-month, and 12-month moving time-window) and December (i.e. for 7-month moving time-window). The results for 12-month moving time window accounts for precipitation events occurring over both the active monsoon and the non-monsoon months and 7-month time-window ending in December accounts for summer monsoon (JJAS) and winter monsoon (OND) months over the study area and are discussed here in detail. These indices differ in their mathematical formulation and the drought classification technique. While SPI relies on fixed thresholds for drought classification, GMM-DI and HMM-DI employ a probabilistic data-driven approach. SPEI uses temperature (UDel_AirT_Precip, ⟨http://www.esrl.noaa.gov/psd/⟩) for calculating evapotranspiration, thus accounting for any temperature rise in the study area during recent decades. The mathematical formulations of the drought indices are summarized in Appendix A.

The drought index values obtained were analyzed further to extract drought characteristics such as severity, duration, areal extent, and frequency. The drought impact index was then computed for each year, by normalizing the product of mean severity

Fig. 2. Drought characteristics over IMR computed for IMD dataset using (a) SPI, (b) GMM-DI, and (c) HMM-DI for 12-month time window ending in September. In each figure the top-panel shows time-series plot of moderate drought severity averaged over all grids. Middle-panel shows the bar-plot of areal extent of moderate droughts represented as percentage of total area in the IMR. Bottom-panel shows the bar-plot of drought impact index for moderate droughts. Solid line represents the median value and dotted line represents slope during the sub-periods 1902–1935, 1936–1970 and 1971–2004 respectively.

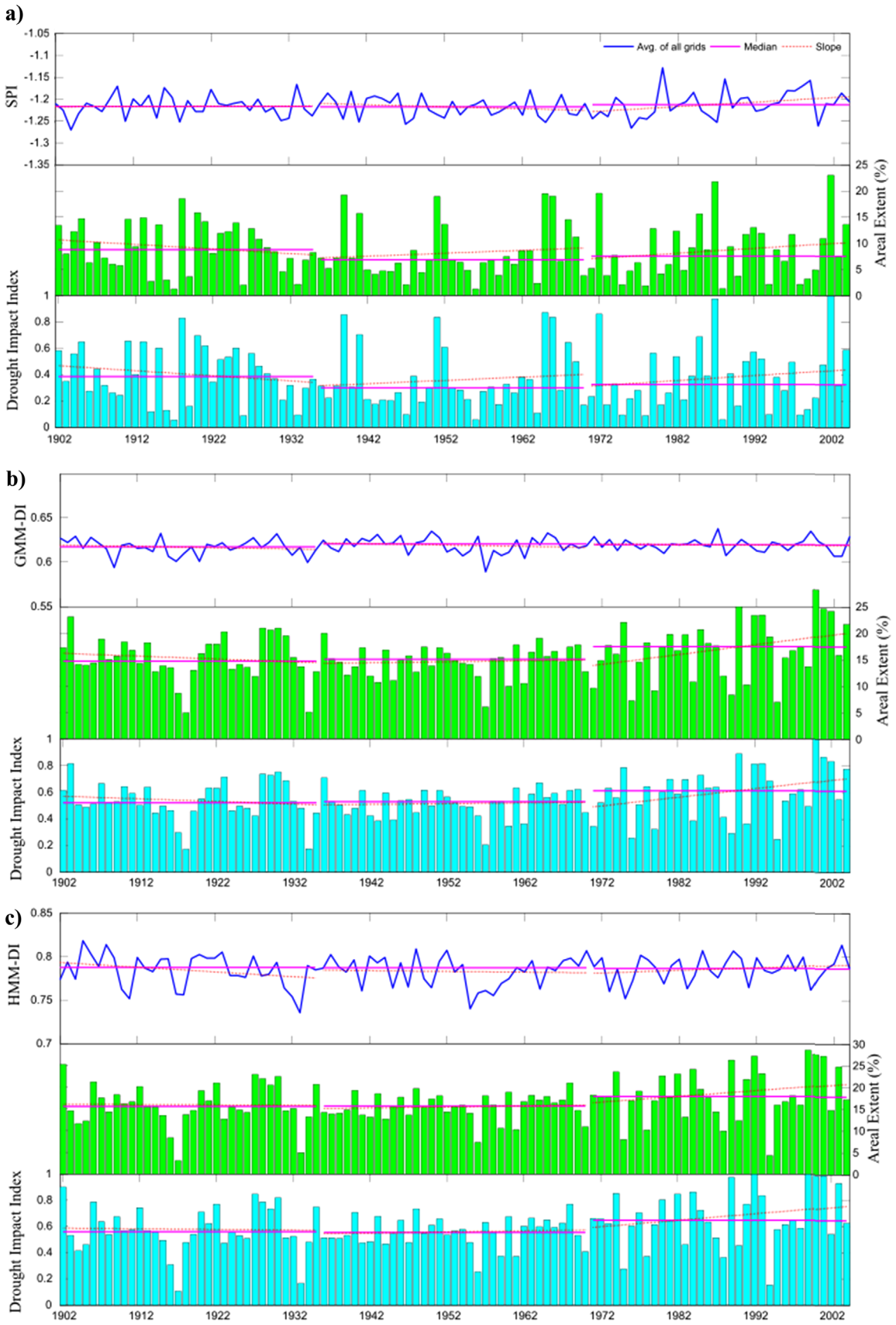


Fig. 3. Same as Fig. 2, but using 0.5° University of Delaware precipitation dataset.

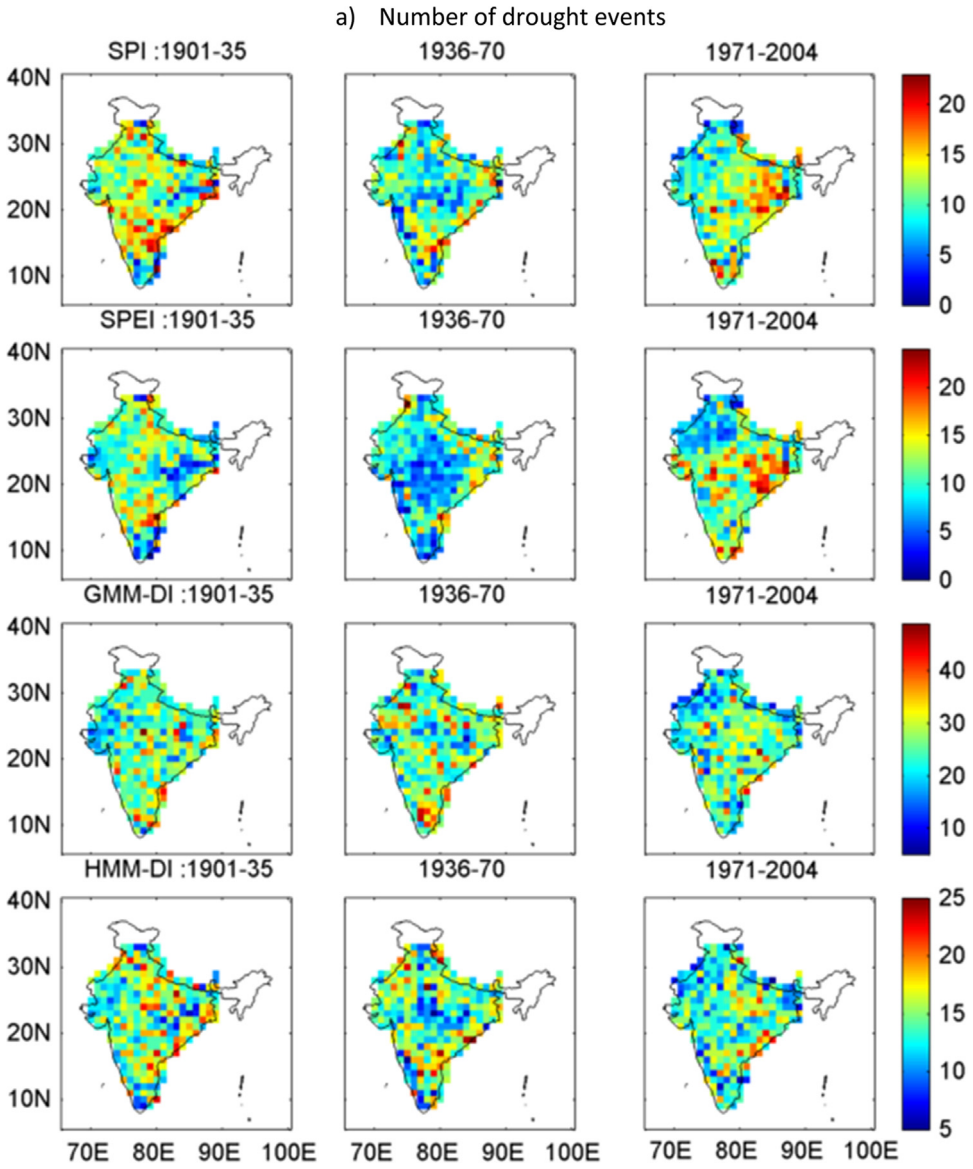


Fig. 4. Epochal variation in drought statistics over IMR using IMD dataset where (a) number of drought events, (b) average intensity of drought, and (c) duration of drought in months. In each sub-plot top panel represents SPI, followed by SPEI, GMM-DI, and HMM-DI.

and the areal extent of drought.

The study period was divided into three segments (1901–1935, 1936–1970, and 1971–2004) to understand the trends and variability associated with retrospective droughts. This was done because droughts have a multiyear influence, and the three periods chosen approximately correspond to periods where IMR experienced significant droughts (e.g. 1918, 1965, 1972, 1987, and 2002). Dividing the entire 104 years (1901–2004) of data into three periods (35, 35, and 34 years) was expected to provide a sufficient length of time series to estimate trends and other statistical values.

A modified Mann-Kendall trend test that accounts for autocorrelation in time-series (Kulkarni and von Storch, 1995; Hamed and Rao, 1998) was used to detect trends in the annual SPI, SPEI, GMM-DI, and HMM-DI values. Trends were estimated on the annual time series for the entire period and for each sub-periods (i.e. 1901–1935, 1936–1970, and 1971–2004) using a 5% significance test. The effect of spatial correlations in the data (Burn and Elnur, 2002; Yue and Wang, 2002) on the trend results were accounted by using false discovery rate (FDR) (Benjamini and Hochberg, 1995; Ventura et al., 2004).

3. Results

3.1. Drought characterization

The drought indices were able to capture (Fig. 2 and Fig. A.3a) the major documented drought events over IMR (De et al., 2005). For the study period, the six most notable moderate-droughts occurred in 1905, 1946, 1965, 1974, 1979, and 1984. In the figure moderate drought refers to SPI values between -1.0 and -1.49 (Charusombat and Niyogi, 2011). Three of the most severe historic droughts occurred during the recent period of 1971–2004. The drought characteristics showed an increasing trend during the same period. Modified Mann-Kendall trend test was performed to test the statistical significance of the trends in these average

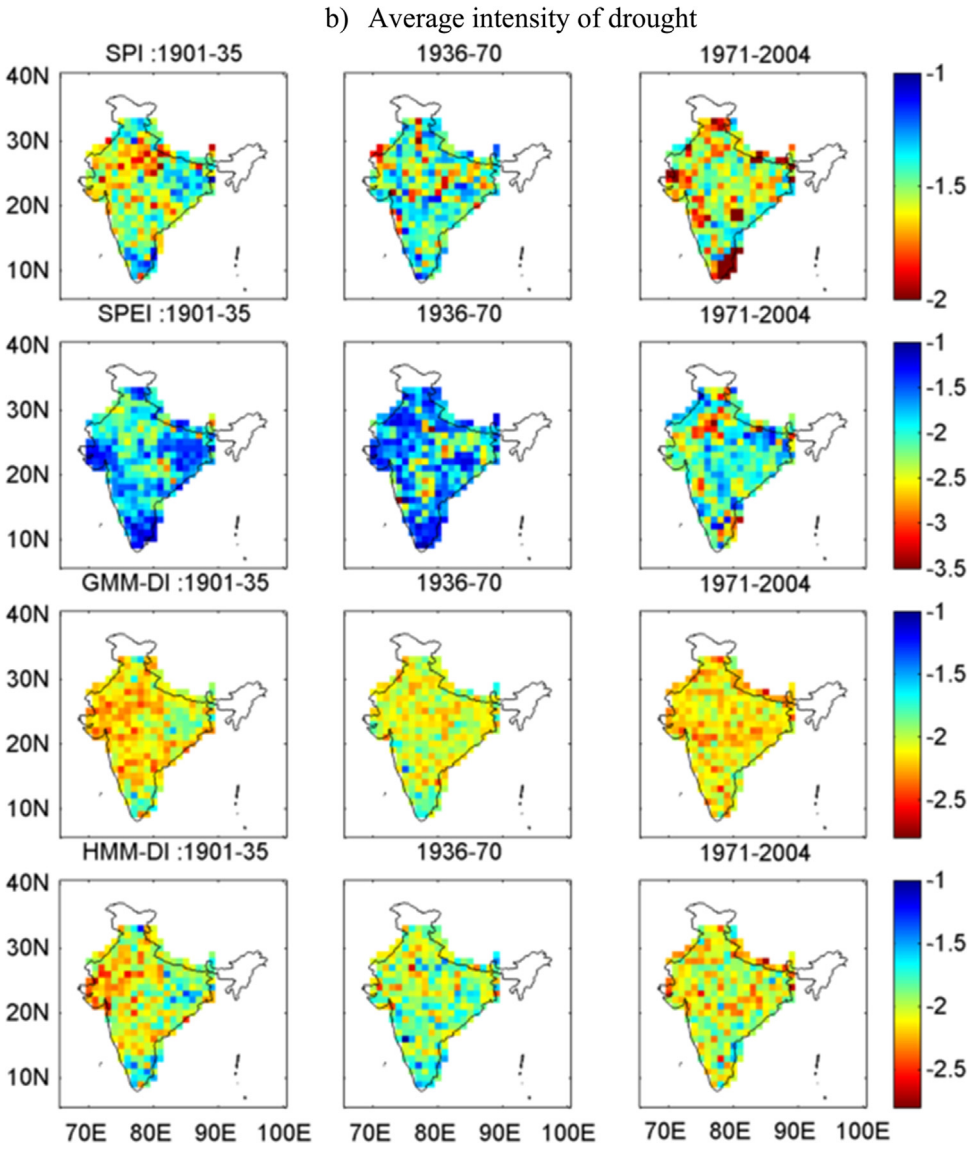


Fig. 4. (*continued*)

drought statistics. For example, SPI analyses (Fig. 2a) showed a drying trend in the mean severity of moderate droughts (-0.04/decade, p-value > 0.05) during the period 1971–2004, indicating increased drying. During the same period the areal extent and drought impact index of moderate droughts also showed increasing trends. Similar trends were observed for SPEI, GMM-DI and HMM-DI analyses (Fig. A.3a and Fig. 2b, c). These trends are consistent with the precipitation trends documented in other studies (Guhathakurta and Rajeevan, 2008; Kripalani et al., 2003; Rupa Kumar et al., 2006).

The trends were reanalyzed in the 0.5° resolution UD precipitation data, thus providing means to compute and validate trends in drought characteristics at a relatively finer spatial resolution over IMR. As in the case of IMD dataset-SPI, SPEI, GMM-DI and HMM-DI were computed. The drought characteristics such as mean severity, areal extent, and drought impact index were computed for each drought index. Again the drought indices were able to capture (Fig. 3 and Fig. A.3b) the major drought events documented over IMR (De et al., 2005) during the period of 1901–2004 and agree well with IMD dataset results (Fig. 2 and Fig. A.3a). There are broad similarities and also specific differences in the characteristics revealed by the choice of index and data. For example, the SPI and SPEI yields a relatively smaller drought impact index as compared to GMM-DI and HMM-DI.

3.2. Spatial and temporal variability in drought characteristics

To study the spatiotemporal variability in droughts; average drought characteristics based on SPI, SPEI, GMM-DI, and HMM-DI values were obtained for each epoch over all grids in IMR by computing the mean number, severity and duration of droughts (e.g. 1901–1935; 1936–1970 and 1971–2004). For the IMD dataset and 12-month time window, during the period 1901–35 there were many widespread droughts (Fig. 4) mainly in the northern, central and the Deccan Plateau regions of IMR. While more number of drought events were observed in the Deccan region (Fig. 4a), the drought duration and intensity were higher in northern and central regions of IMR. During the epoch of 1936–1970 the droughts were more active in the western region and parts of Deccan Plateau of IMR. Compared to 1901–35, droughts

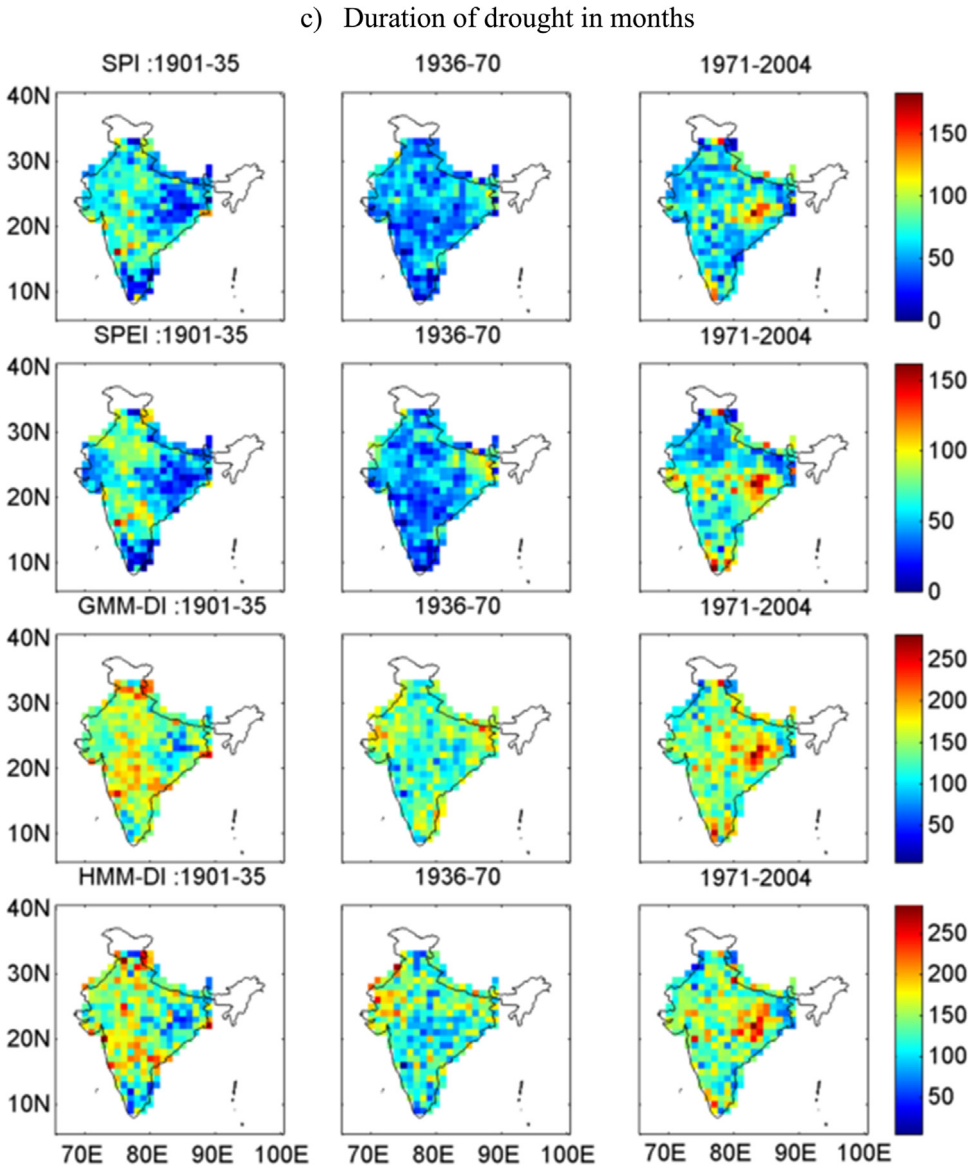


Fig. 4. (*continued*)

were less frequent during this epoch (1936–1970). During 1971–2004, the number of drought events and their duration increased in the central and eastern Indo-Gangetic plain (IGP; 20N-28N), and southern parts of IMR. High drought intensities were recorded in central and eastern IGP, south-India, and parts of western-India (that include states of Maharashtra, Gujarat, and Rajasthan). Drought patterns were mostly similar for all four drought indices in each epoch – while GMM-DI showed more wide spread droughts; SPI, SPEI, and HMM-DI were better able to distinguish the drought hotspots.

Results for the UD dataset were similar to those obtained for IMD dataset. There were many widespread droughts in the western and central parts of IMR during 1901–1935 (Fig. 5). During 1936–1970, except for some parts of western, central and southern India, most of the IMR was the wettest and droughts were infrequent. As in case of IMD dataset (Fig. 4), the number of droughts and duration of droughts increased in the central and eastern IGP (20N–28N), and southern parts of India during 1971–2004. The drought intensities were higher in interior parts of south-India, western parts of India (Maharashtra, Gujarat, and Rajasthan) and IGP. The drought indices – SPI, SPEI, GMM-DI and HMM-DI were able to consistently capture the space and time evolution of drought characteristics over the IMR during the entire study period. A notable west to east migration in the drought severity and extent over the last century is seen.

Similar comparison of epochal drought characteristics over IMR for 7-month time window using IMD (Fig. A.4) and UD (Fig. A.5) datasets showed that during 1901–1935 droughts were more intense and frequent in parts of Deccan Plateau, western and northern parts of India. Droughts were comparatively less frequent during the epoch 1936–1970 according to SPI and SPEI, however GMM-DI and HMM-DI analysis shows that droughts continue to be intense and frequent in western-India and parts of Deccan Plateau. During 1971–2004 central-India, eastern IGP, and parts of south-India emerge as drought hotspots – along with high intensity but short-term droughts in western-India.

A decadal comparison of 12-month time window drought characteristics over IMR using IMD dataset (Fig. A.6) shows higher level of drought activity in northern-India, western-India and Deccan Plateau during the 1901–10, 1911–20, with more

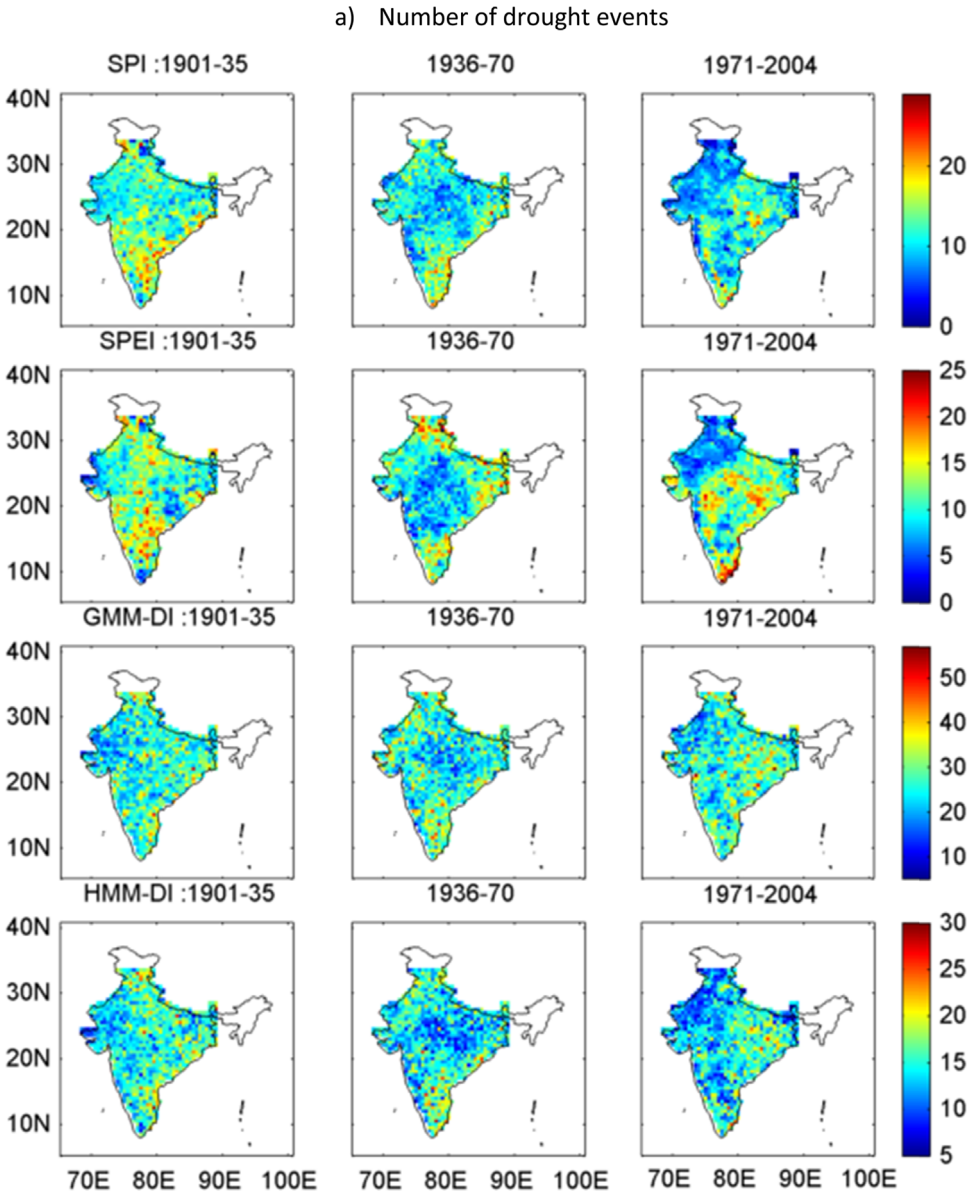


Fig. 5. Same as Fig. 4, but using 0.5° University of Delaware precipitation dataset.

intensification during 1921–30. The subsequent two decades (1931–40 and 1941–50) were amongst the wettest in the past century. Droughts started to emerge in the eastern-IGP during late 1951–60 and intensified in IGP and parts of western-India during 1961–70. During 1971–80 droughts continued to persist over eastern IGP, and in the following decade (1981–90) additional hotspots emerged in south-India and parts of western-India. During 1991–2000 and onwards, eastern-IGP and parts of central-India continue to be the drought hotspot. Similar patterns in drought characteristics were observed in our analysis when using UD dataset, and for different time windows.

3.3. Trends

Figs. 6 and 7 show the trends in drought intensity computed using modified Mann-Kendall trend test, for SPI, SPEI, GMM-DI, and HMM-DI analysis for the IMD and UD datasets respectively, for 12-month time window ending in September. In the IMD dataset, for SPI analysis, during the epoch 1936–1970 (Fig. 6a) drought intensity increased (trend is towards negative SPI values as its magnitude is negative) in the eastern Indo-Gangetic plain and parts of south-India. During the recent epoch 1971–2004, additional grids showed an increase in drought intensity in south-India (parts of coastal Tamilnadu and coastal Karnataka) and western-Rajasthan. These results are consistent with Niyogi et al. (2010) who have shown using empirical orthogonal functions and genetic algorithm-based analyses that anthropogenic land use modifications due to agricultural intensification may have resulted in significant decline in precipitation in north/northwest India and increasing patterns over east central India. Similar conclusions could be drawn from SPEI analysis (Fig. 6b), GMM-DI analysis (Fig. 6c) and HMM-DI analysis (Fig. 6d). Thus parts of eastern Indo-Gangetic plain, western-Rajasthan, and parts of coastal south-India emerge as the current hotspots for droughts.

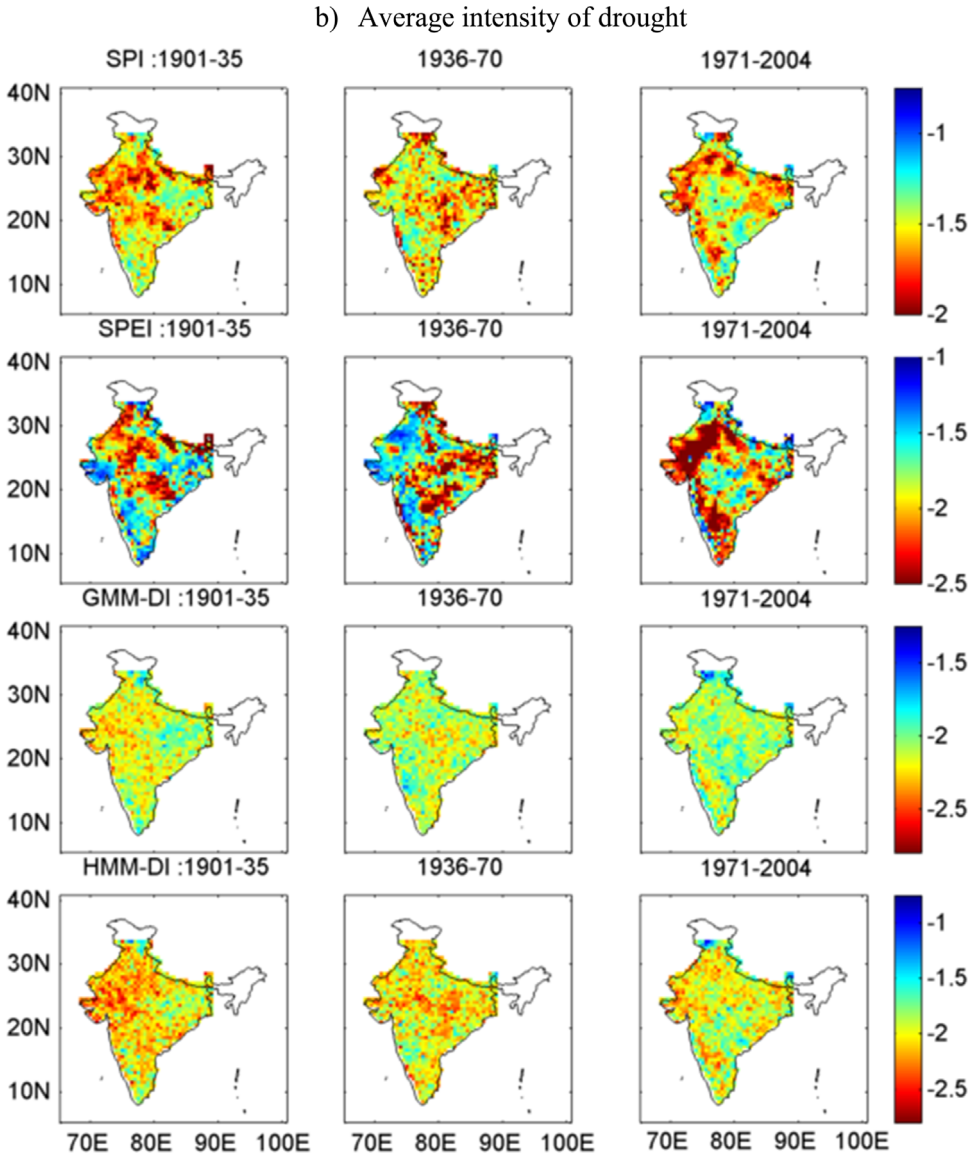


Fig. 5. (*continued*)

For the finer-resolution UD dataset (Fig. 7), using a 12-month time window ending in September, each of the four drought indices show an increasing trend in drought intensity during the period 1936–70 over the eastern Indo-Gangetic plain. However, during 1971–2004 trends in drought intensity also show an increase in south-India (parts of coastal Tamilnadu, coastal Karnataka, and central Maharashtra) and western-Rajasthan, in addition to central and eastern IGP. Thus as in case of IMD dataset, we can conclude that parts of eastern Indo-Gangetic plain, and parts of coastal south-India are emergent vulnerable regions to droughts.

At shorter time scales (e.g. 7-months ending in December) it was found that in addition to central- and eastern-IGP and costal south-India, interior parts of Maharashtra and central India were emerging as vulnerable regions to droughts for both IMD (Fig. A.7) and UD datasets (Fig. A.8).

3.4. Drought frequency

Hypothesis tests were carried out to investigate whether the number of droughts had significantly increased during the recent epoch 1971–2004, when considering 12-month droughts ending in September. A right tailed t-test with significance level (α) of 5% was used. Fig. 8a and b shows the results of the hypothesis test at each grid of the IMD and UD datasets for SPI, SPEI, GMM-DI, and HMM-DI, respectively. The results indicate that the hypothesis test was significant, or in other words the number of droughts had shown a statistically significant increase at several grids in the study region. To account for the bias induced in the hypothesis test due to spatial correlation in the gridded meteorological data, a FDR test (Ventura et al., 2004; Wilks, 2006) was performed. The FDR test further confirmed that the number of droughts showed a statistically significant increase in the Indo-Gangetic plains, coastal south-India, and central Maharashtra during the recent period 1971–2004.

Similarly, for 7-month time window ending in December, it was found that the number of droughts showed a statistically significant increase in the central and eastern IGP and interior parts of Maharashtra during the recent epoch (1971–2004) for both IMD (Fig. A.9a) and UD datasets (Fig. A.9b).

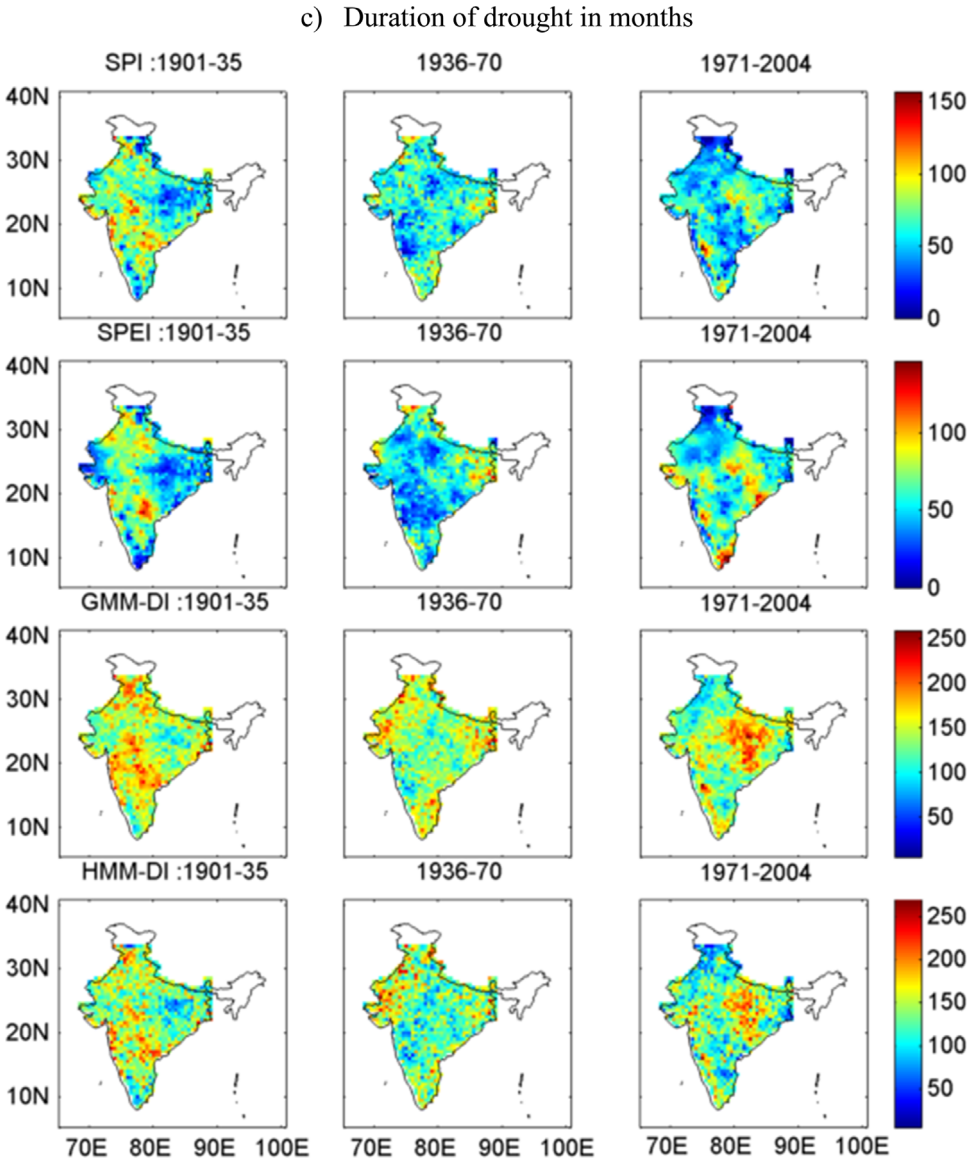


Fig. 5. (*continued*)

3.5. Drought vulnerability

Fig. 9a and b shows the regions over IMR that were vulnerable to droughts (defined as SPI < -1.0) using IMD and UD precipitation datasets for the three study periods, considering a 12-month time window. Using the gridded population estimates available (CIESIN, 2005), an estimate of the population affected by droughts for the three periods was made. According to SPI, during the recent period of 1971–2004 approximately 405 million people were in the drought affected region. This is equivalent to a GDP of USD 208 billion. The population and GDP estimates are calculated after defining a threshold for drought intensity below which a drought is considered to have negative impact on the economy and society. The values in the bar plot (see inset in Fig. 9a and b) correspond to an intensity threshold of -1.0 for SPI. Similar computations using SPEI, GMM-DI and HMM-DI resulted in consistently higher estimates compared to SPI for each of the three periods. This may be due to the choice of threshold and the differences in the methodology used in their computation.

4. Conclusions

Recent studies have highlighted that IMR has a steady increase in the drought patterns. Motivated by the cautionary conclusions of Trenberth et al. (2014), a reassessment of the drought patterns using multiple data sources and methods was desired. Accordingly, we examined the long-term retrospective drought variability over the Indian Monsoon Region (IMR) using two gridded precipitation datasets that differ in their primary data source and spatial resolution. Moreover, we compared several drought characteristics (severity, duration, areal extent, and frequency) using SPI, SPEI, GMM-DI, and HMM-DI to assess the variability in the results.

The 104 year (1901–2004) SPI, SPEI, GMM-DI, and HMM-DI were analyzed for three periods 1901–1935, 1936–1970, and 1971–

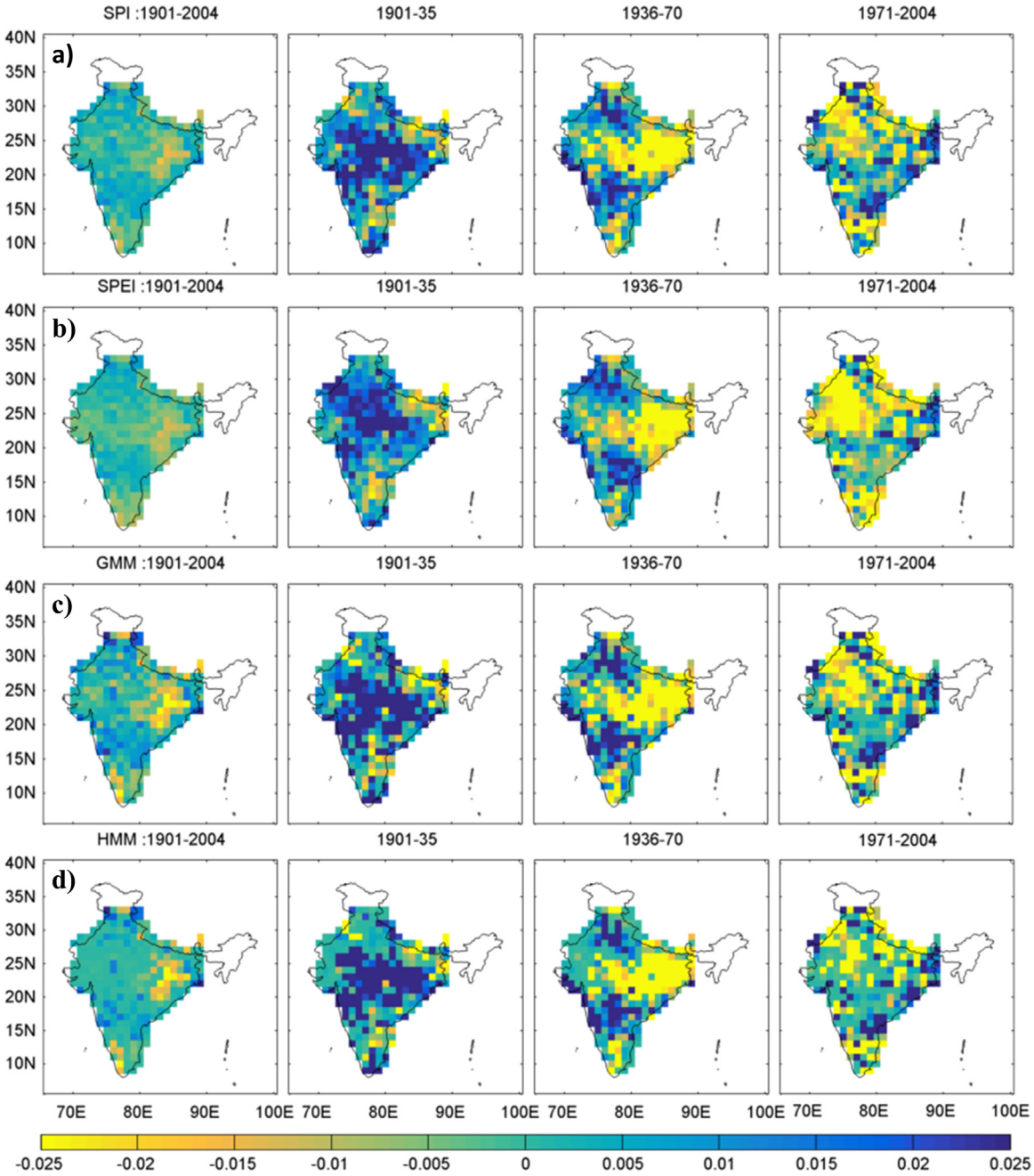


Fig. 6. Mann-Kendall trend slope for 12-month droughts ending in September over IMR during the periods 1901–2004, 1902–1935, 1936–1970, and 1971–2004. Results correspond to the IMD dataset using (a) SPI, (b) SPEI, (c) GMM-DI, and (d) HMM-DI.

2004. Epochal and decadal variation in drought characteristics over IMR were analyzed. Consistent with the findings from recent studies that indicate that the monsoon precipitation is becoming extreme and regionally varied, we found that there is a significant change in the drought climatology over the IMR. Results indicated that the droughts are becoming much more regional in recent decades and showing a general migration from west to east and the Indo-Gangetic plain. We found an increased duration, severity, and spatial extent in the recent decades and identify the Indo-Gangetic plain, parts of coastal south-India and central Maharashtra as vulnerable regions for recent droughts. Despite some differences in results for the choice of drought indices, the time window chosen for analysis, and/or the precipitation dataset (resolution) used, overall the results and conclusions are consistent.

It is beyond the scope of present study to assess the causal mechanism of droughts, and to find if the observed trends are related to other phenomena such as changes observed in the monsoon break (active-dry spell) periods (Singh et al., 2014). There are a number of possible mechanisms – aerosols, landuse

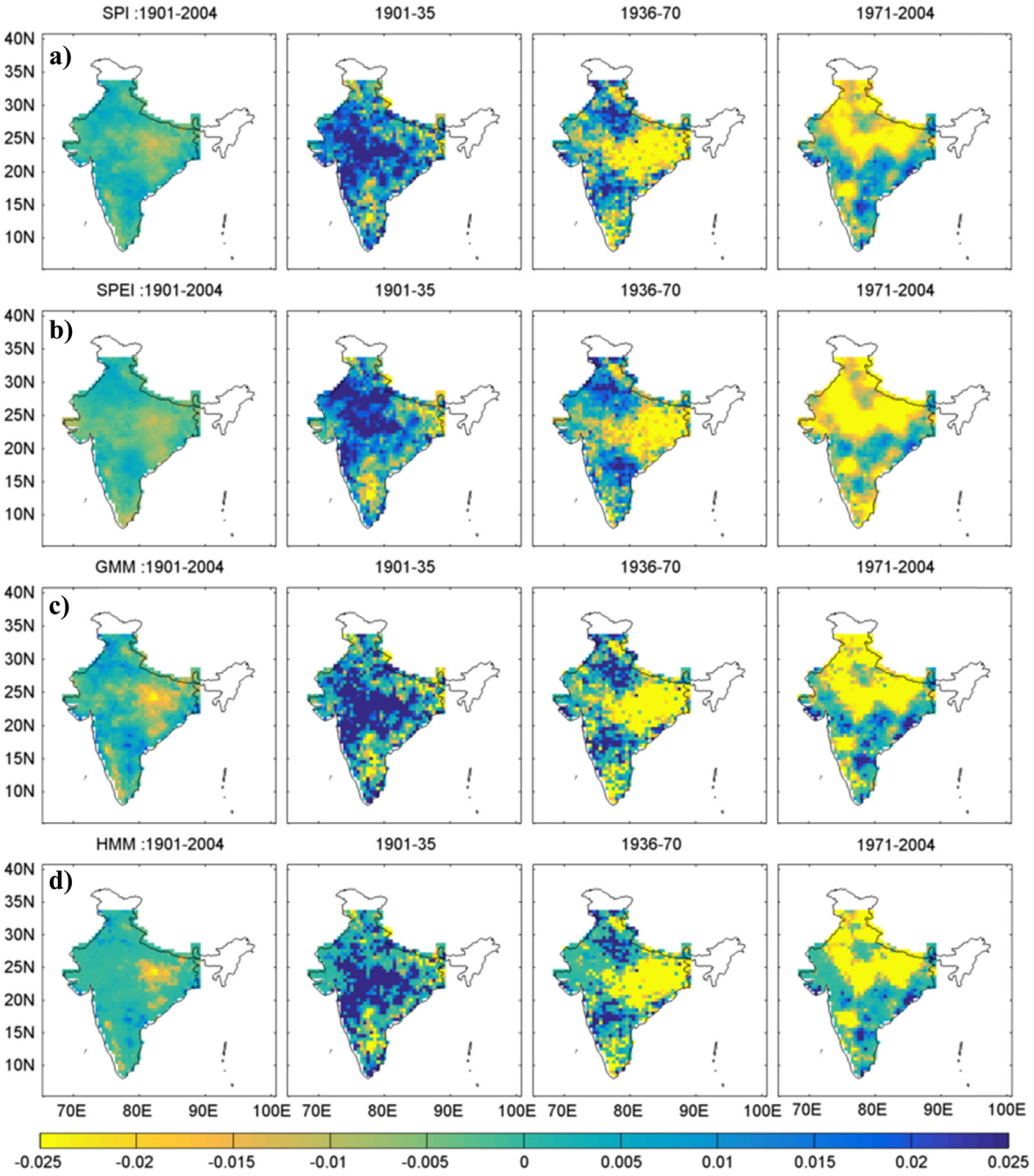


Fig. 7. Same as Fig. 6, but using 0.5° University of Delaware precipitation dataset.

change, SST changes, global changes, thermodynamic feedback due to heating rates (Roxy et al., 2015), as a result, diagnosis and discussion of potential mechanisms will have to be a part of follow-up study using numerical models. The results from this study provide the baseline for future climate change studies, and also provide robust conclusion that irrespective of the datasets and methodology used, the IMR has high potential of droughts and that the droughts appear to be migrating to the agriculturally important regions including Indo-Gangetic plains.

Acknowledgment

The authors acknowledge NSF CAREER (AGS 0847472, Dr. Anjuli Bamzai), NSF INTEROP DriNET (0753116), NIFA USDA Drought Trigger Projects at Purdue through Texas A&M (2011-67019-20042), Purdue Climate Change Research Center, Indo-US Science and Technology Forum (IUSSTF) ⟨http://www.iusstf.org/⟩, and the Information Technology Research Academy-Water (ITRA-W).

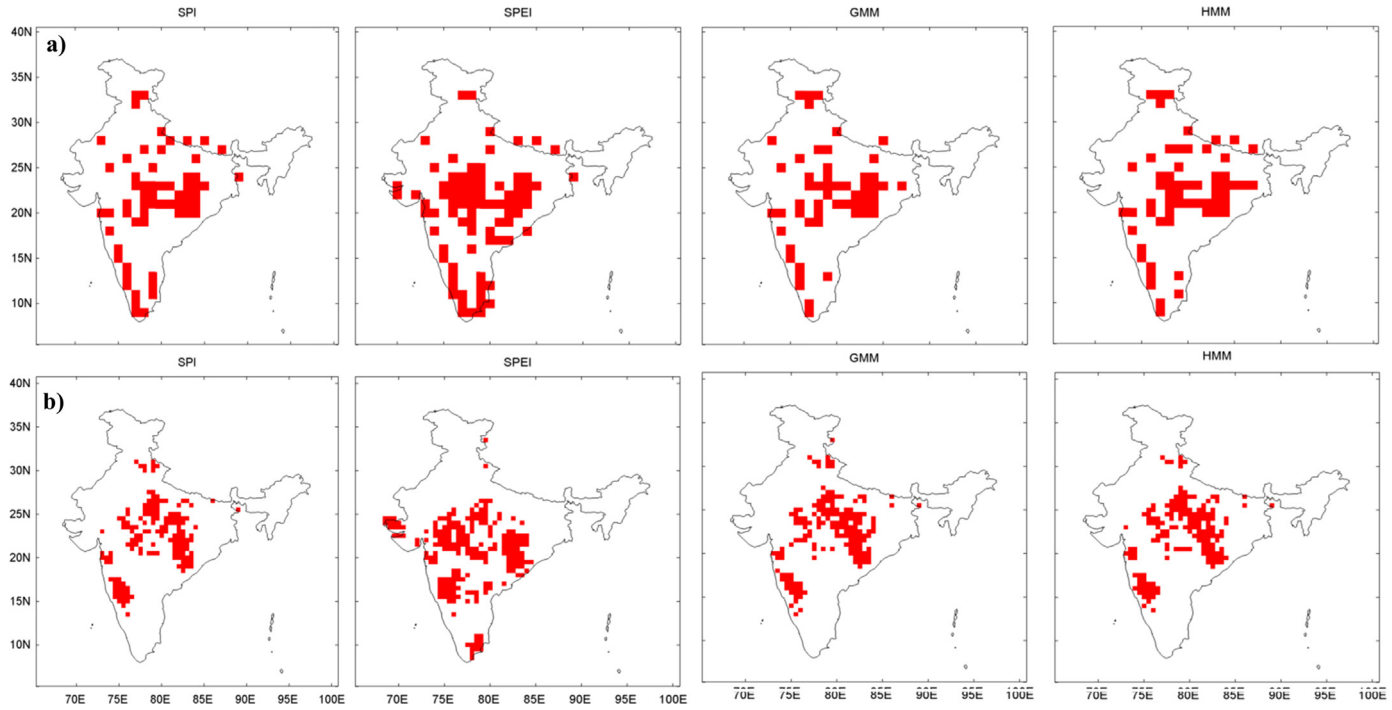


Fig. 8. Hypothesis test to see if the number of droughts (moderate, severe and extreme) of 12-month time window ending in September have increased during the period 1971–2004 in comparison to 1936–1970 for (a) IMD and, (b) UD precipitation datasets according to SPI, SPEI, GMM-DI, and HMM-DI. Grids where the number of droughts show a statistically significant increase at $\alpha = 0.05$ are displayed.

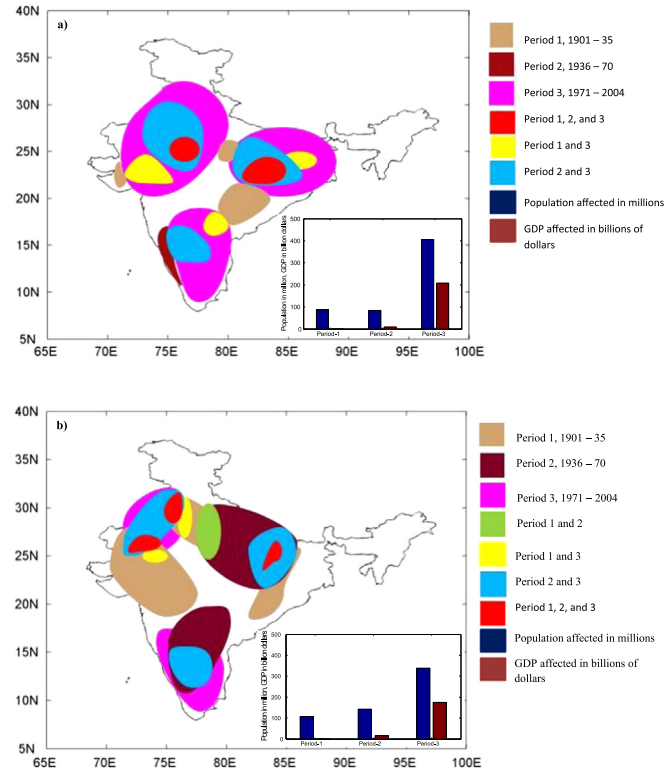


Fig. 9. The estimate of population and GDP affected, and the drought hotspots during the sub-periods 1901–1935, 1936–1970, and 1971–2004 according to SPI < -1.0 for (a) IMD precipitation dataset and (b) UD precipitation dataset.

Appendix A

See Appendix Figs. A.1–A.9.

Time window accounting

Droughts of different time-windows (1-month, 4-months, 7-months, 12-months, and so on) were analyzed as part of this study. In this paper we discuss results for 7-month time window ending in December which accounts for precipitation over the summer monsoon season (June to September or JJAS) and winter monsoon season (October to December or OND). Thus 7-month precipitation ending in December-2000 would account for the total precipitation from June-2000 to December-2000.

We also present results for 12-month time window ending in September as it accounts for the total precipitation during monsoon and non-monsoon months. For 12-month precipitation ending in September 2000, would account for cumulative rainfall from October-1999 to September-2000.

SPI methodology

The SPI, measures the deficit in observed precipitation (McKee et al., 1993) and has been used widely to identify meteorological, agricultural, and hydrological droughts (Mishra and Singh, 2010; Mo, 2008). Precipitation time-series for each grid cell over IMR at any desired time-scale was first used to fit a probability distribution function, and then normalized using a standard inverse Gaussian function to obtain SPI values. Drought severity was identified using the SPI ranges as described by Charusombat and Niyogi (2011). A drought event was classified as *moderate* if SPI

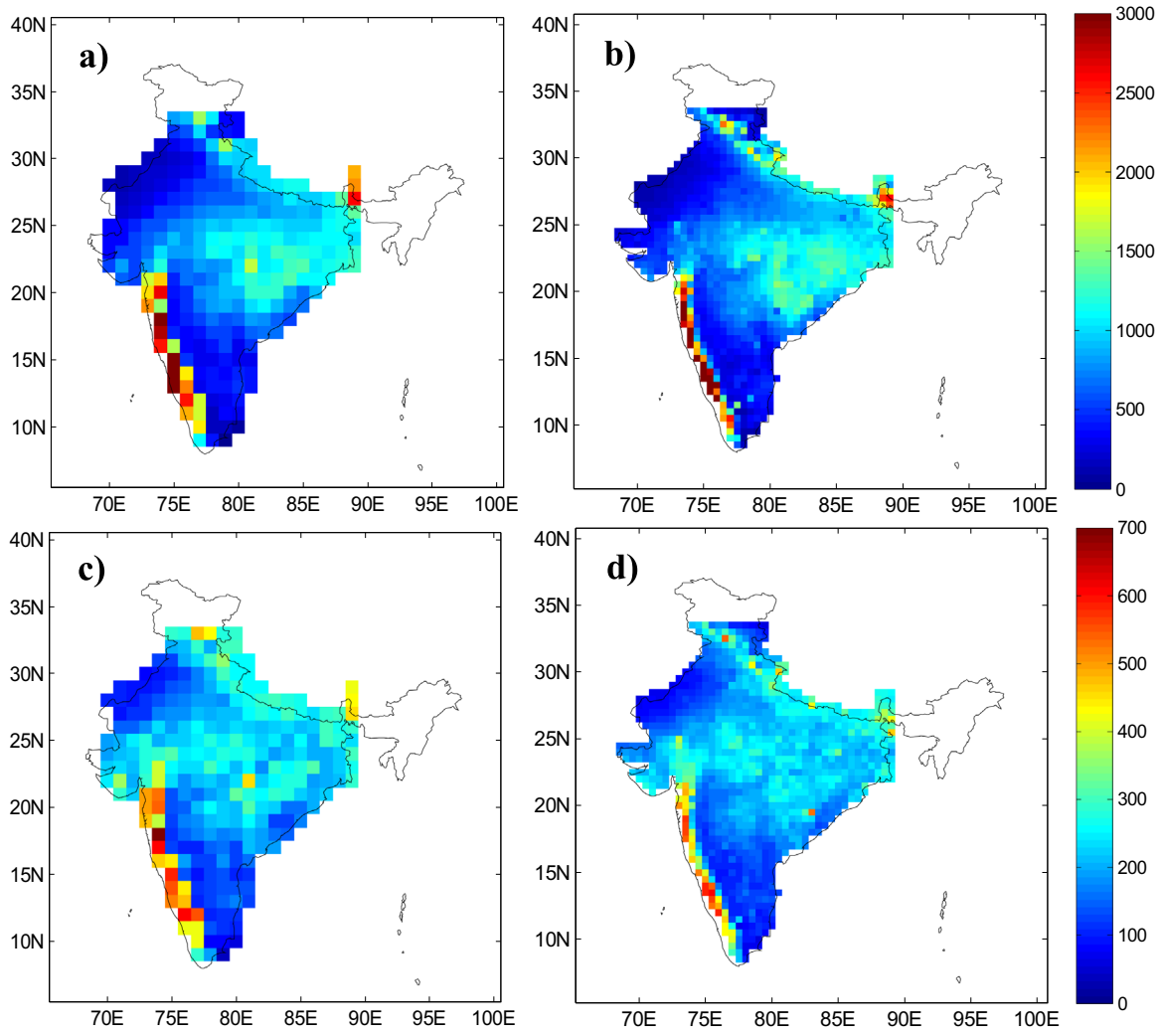


Fig. A.1. Comparison of IMD and UD monthly precipitation statistics over each grid in the study region. Mean of monthly precipitation (in mm) over (a) IMD grids, and (b) UD grids. Standard deviation of monthly precipitation (in mm) over (c) IMD grids, and (d) UD grids.

was between −1.0 to −1.49, *severe* if SPI was between −1.5 to −2.0, and *extreme* if SPI was less than −2.0.

SPEI methodology

SPEI (Vicente-Serrano et al., 2010) first requires the computation of potential evapotranspiration (PET). We have used Thornthwaite's equation (Thornthwaite, 1948) for computing PET, but other popular approaches may also be used (Penman, 1948; Priestley and Taylor, 1972; Allen et al., 1998). After subtracting the PET from precipitation, SPEI may be computed using similar approach as SPI (McKee et al., 1993).

GMM-DI methodology

A GMM is a probabilistic model where the parametric density function is represented as a weighted sum of the Gaussian component densities (Reynolds and Rose, 1995). GMMs have been successfully used in recent studies involving estimation of weather parameters using remotely sensed radar data (Li and Zhang, 2011), and hydrologic forecasting studies (Liang et al., 2011). In this study, each individual Gaussian component is assumed to represent the underlying distribution of the hidden drought (or wet) classes with a mean μ, and a covariance matrix Σ. Such a GMM model closely represents a hidden Markov model (HMM) with equal transition probabilities amongst all hidden states. The mathematical formulation of the GMM used in this study is described below.

Let the precipitation at time t be denoted by x_t, $t = 1, \ldots, N$ $\{x_t \in R$ and $X=[x_t,\ldots,x_N]^T = x_{1:N}\}$. If the total number of components of the GMM, M, are known *a priori*, then the weighted sum of Mcomponent GMM is given by the equation,

$$p(x \mid \lambda) = \sum_{i=1}^{M} w_i g(x \mid \mu_i, \Sigma_i), \tag{A.1}$$

where w_i are the mixture weights, and $g(x|\mu_i, \Sigma_i)$ are the component Gaussian densities of the form,

$$g(x \mid \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu_i)'\Sigma_i^{-1}(x-\mu_i)\right\}, \tag{A.2}$$

with mean μ_i and covariance matrix Σ_i. In this study, since only precipitation data are used, the number of dimensions, $D = 1$. Further, the mixture weights satisfy the constraint $\sum_{i=1}^{M} w_i = 1$. The parameter set can be represented by the notation shown below.

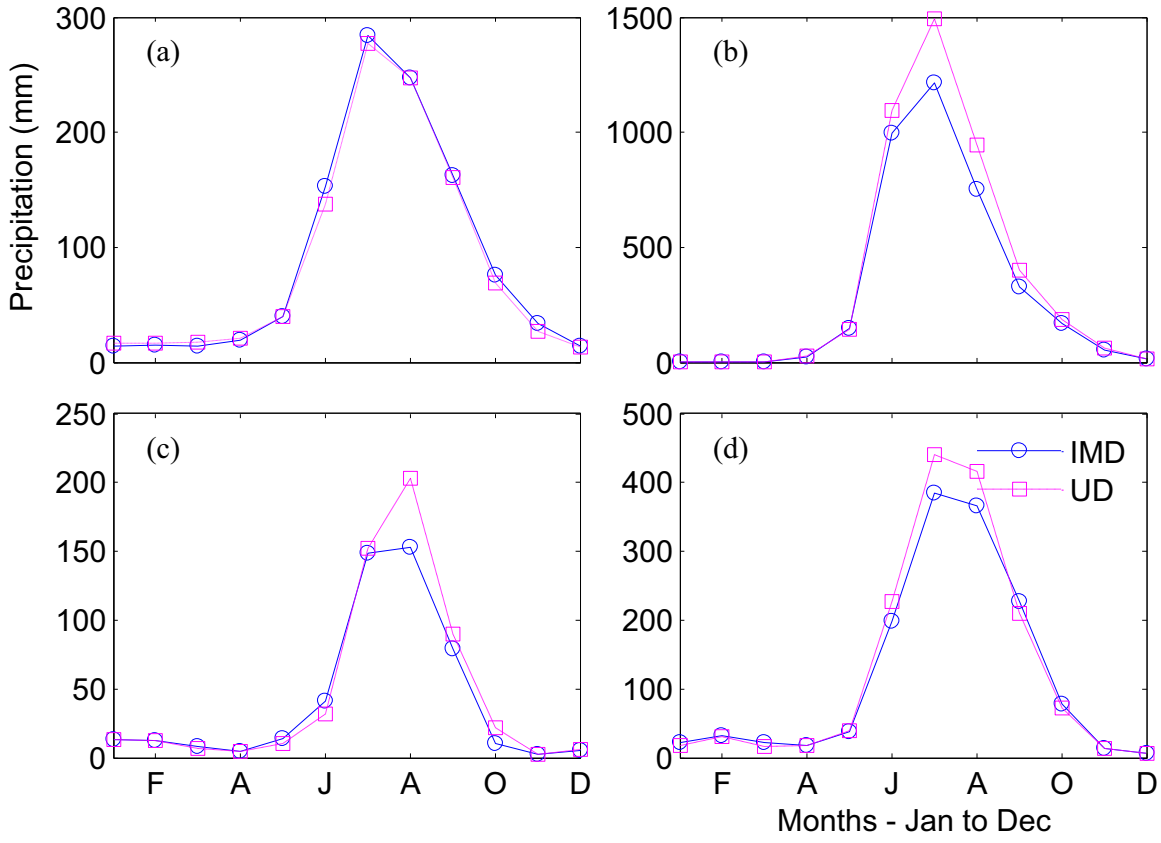


Fig. A.2. Comparison of monthly mean precipitation between IMD and UD datasets for all months of a year (January to December), (a) averaged over all grids in the study region, (b) IMD grid 18 and UD grid 208 over Western Ghats, (c) IMD grid 275 and UD grid 430 over Punjab, and (d) IMD grid 140 and UD grid 1005 over West Bengal.

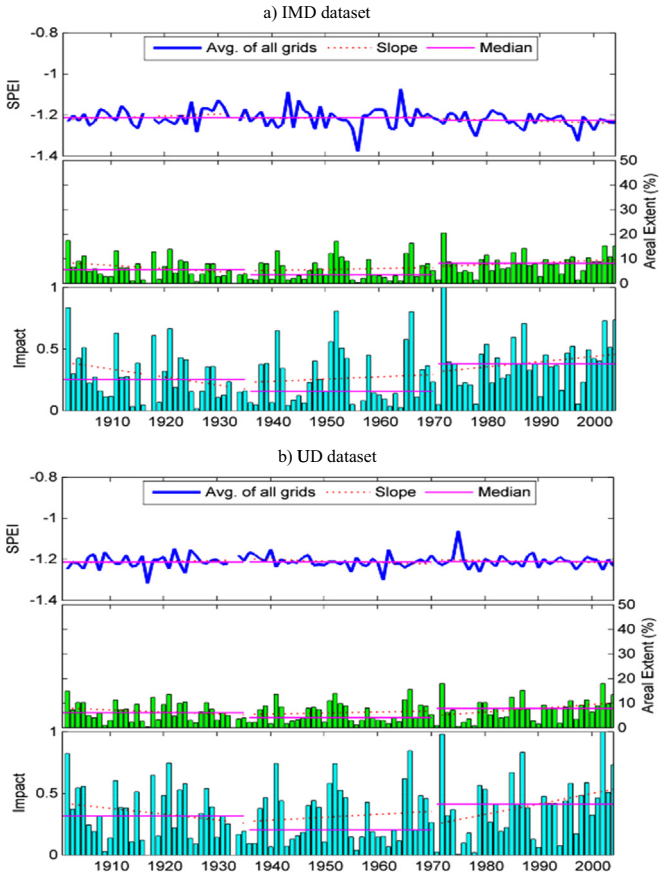


Fig. A.3. Drought characteristics over IMR computed for (a) IMD dataset and (b) UD dataset using SPEI values for 12-month time window ending in September. In each figure the top-panel shows time-series plot of moderate drought severity averaged over all grids. Middle-panel shows the bar-plot of areal extent of moderate droughts represented as percentage of total area in the IMR. Bottom-panel shows the bar-plot of drought impact index for moderate droughts. Solid line represents the median value and dotted line represents slope during the sub-periods 1902–1935, 1936–1970 and 1971–2004 respectively.

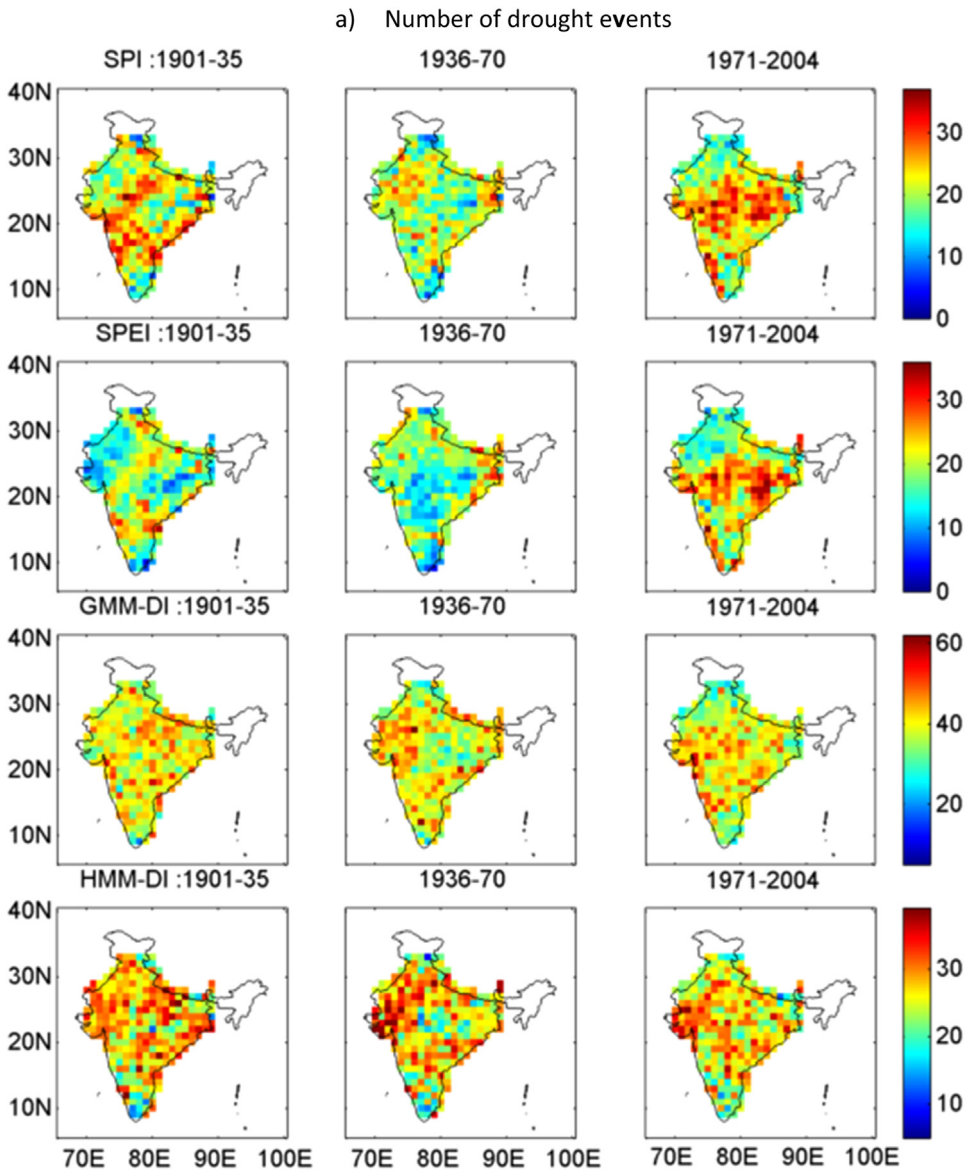


Fig. A.4. Epochal variation in 7-month drought statistics over IMR using IMD dataset where (a) number of drought events, (b) average intensity of drought, and (c) duration of drought in months. In each sub-plot top panel represents SPI, followed by SPEI, GMM-DI, and HMM-DI.

$$\lambda = \{w_i, \mu_i, \Sigma_i\},\ i=1, \ldots, M \tag{A.3}$$

Expectation-maximization (EM) algorithm (Dempster et al., 1977; McLachlan and Krishnan, 1997) was used to estimate the parameters of the GMM using a maximum likelihood approach. The *a posteriori* probability for component i was given by

$$P(i \mid x_t, \lambda) = \frac{w_i p(x_t \mid \mu_i, \Sigma_i)}{\sum_{k=1}^{M} w_k p(x_t \mid \mu_k, \Sigma_k)} \tag{A.4}$$

To compare the results of GMM-DI with SPI, the number of mixture components, M, was set to 7 (3 drought states+1 normal state+3 wet states).

HMM-DI methodology

The mathematical formulation of the hidden Markov model-based drought index (HMM-DI) is described in detail in Mallya (2011). The precipitation at time t be denoted by x_t, $t = 1, \ldots, N$ $\{x_t \in R$ and $X=[x_t,\ldots,x_N]^T = x_{1:N}\}$. In a HMM, the precipitation x_t is assumed to depend only on the state variable z_t, $\{Z=[z_1,\ldots,z_N]^T\}$ that denotes a drought or wet state, is hidden (not observed), and follows the first order Markov property. The state variable z_t is a K-dimensional binary random variable. If the number of states, K, are known a priori, the standard HMM can be parameterized using the following three distributions: (i) The conditional distribution of precipitation given the drought state, $p(x_t \mid z_t)$, referred to as the emission distribution. (ii) The conditional distribution of the present drought state given the previous state i.e. $p(z_t \mid z_{t-1})$. Because z_t is a K dimensional binary variable, the conditional distribution is given by a $K \times K$ transition matrix A whose element $A_{jk}=p(z_{tk}=1 \mid z_{t-1,j}=1)$. (iii) The marginal distribution of the drought state at the first time step, $p(z_1)$, is given by a K dimensional vector

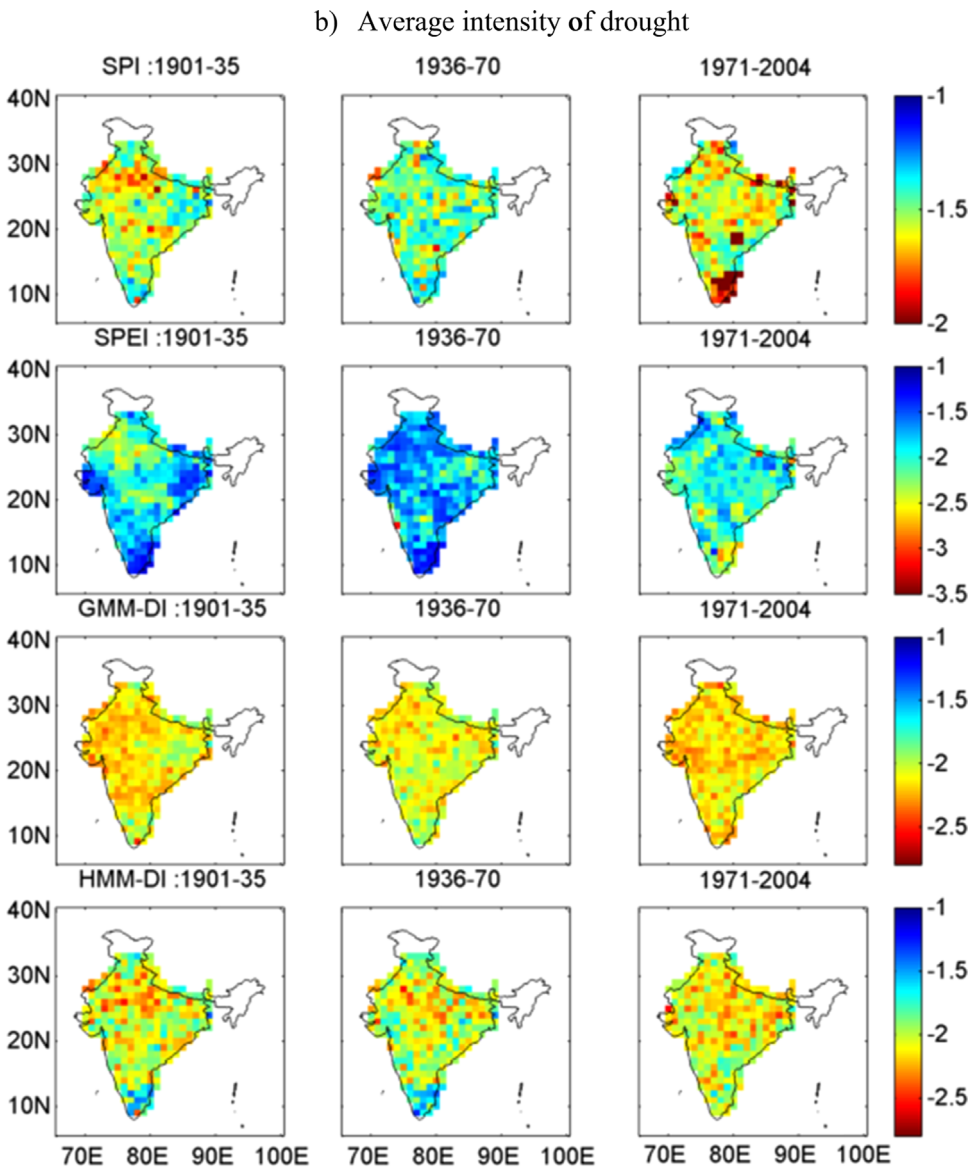


Fig. A.4. (*continued*)

Π whose element $\pi_k = p(z_{1k}=1)$.

The precipitation data at the desired time-scale was transformed to represent percentage deviation from their long term mean. The HMM model was then applied to this transformed data. The probability density function of the emission distribution was selected to be a Gaussian distribution of the form

$$p(x_t \mid z_t) = \prod_{k=1}^{K} \mathrm{N}(x_t \mid \mu_k, \sigma_k^2)^{z_{tk}} \tag{A.5}$$

where μ_k and σ_k^2 are the mean and the variance of a Gaussian distribution, respectively. The μ_k's and σ_k's were considered to be free parameters and were estimated along with other parameters of HMM. Since the results of HMM-DI are to be compared with SPI and GMM-DI, the number of states (components in the Gaussian mixture) K was set to 7 (3 drought states + 1 normal state + 3 wet states).

The underlying emission distributions are not known before hand and were assumed to be Gaussian. This was done for mathematical convenience, and also because many processes combine to create droughts, and one may expect that their combined influence expressed through deviations from the mean could be Gaussian. Additionally, if there is no temporal dependence, the HMM automatically collapses to a Gaussian mixture for which the theories are well developed.

Both GMM-DI and HMM-DI provides probabilities of belonging to each drought class. To get drought intensity values we have multiplied an intensity factors (e.g. assumed intensity factors for Extreme drought = −3.0, Severe drought = −2.0, Moderate drought = −1.0 and so on) with corresponding probability measures.

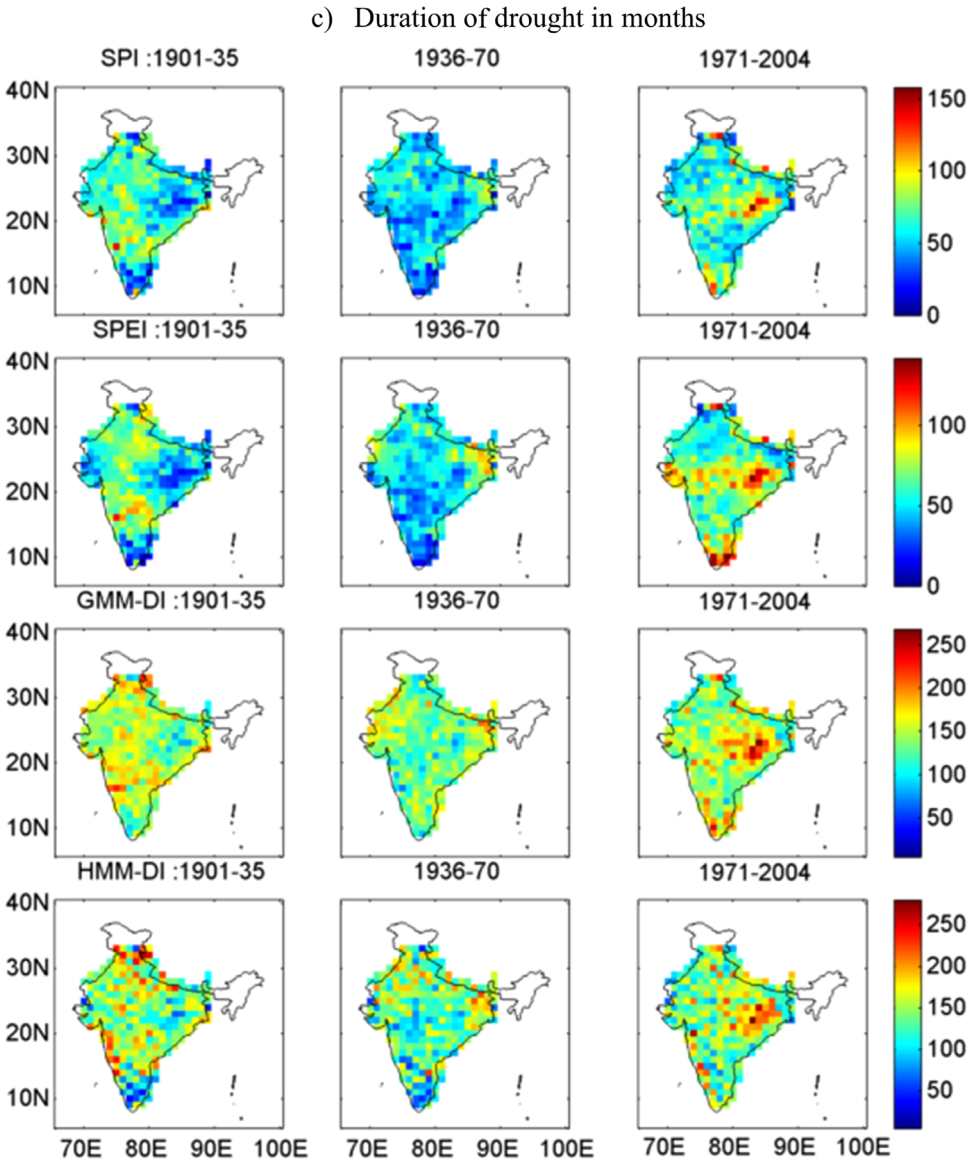


Fig. A.4. (*continued*)

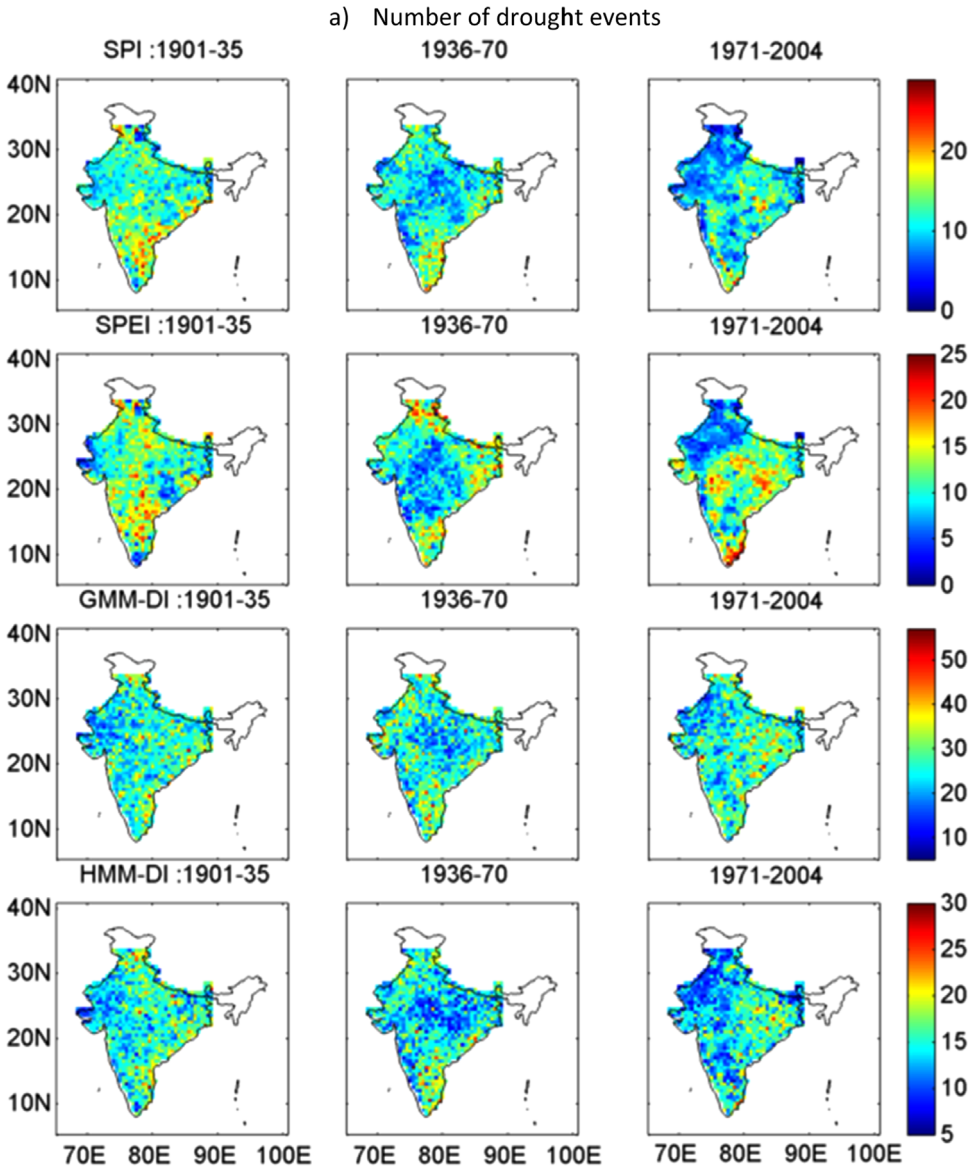


Fig. A.5. Epochal variation in 7-month drought statistics over IMR using UD dataset where (a) number of drought events, (b) average intensity of drought, and (c) duration of drought in months. In each sub-plot top panel represents SPI, followed by SPEI, GMM-DI, and HMM-DI.

b) Average intensity of drought

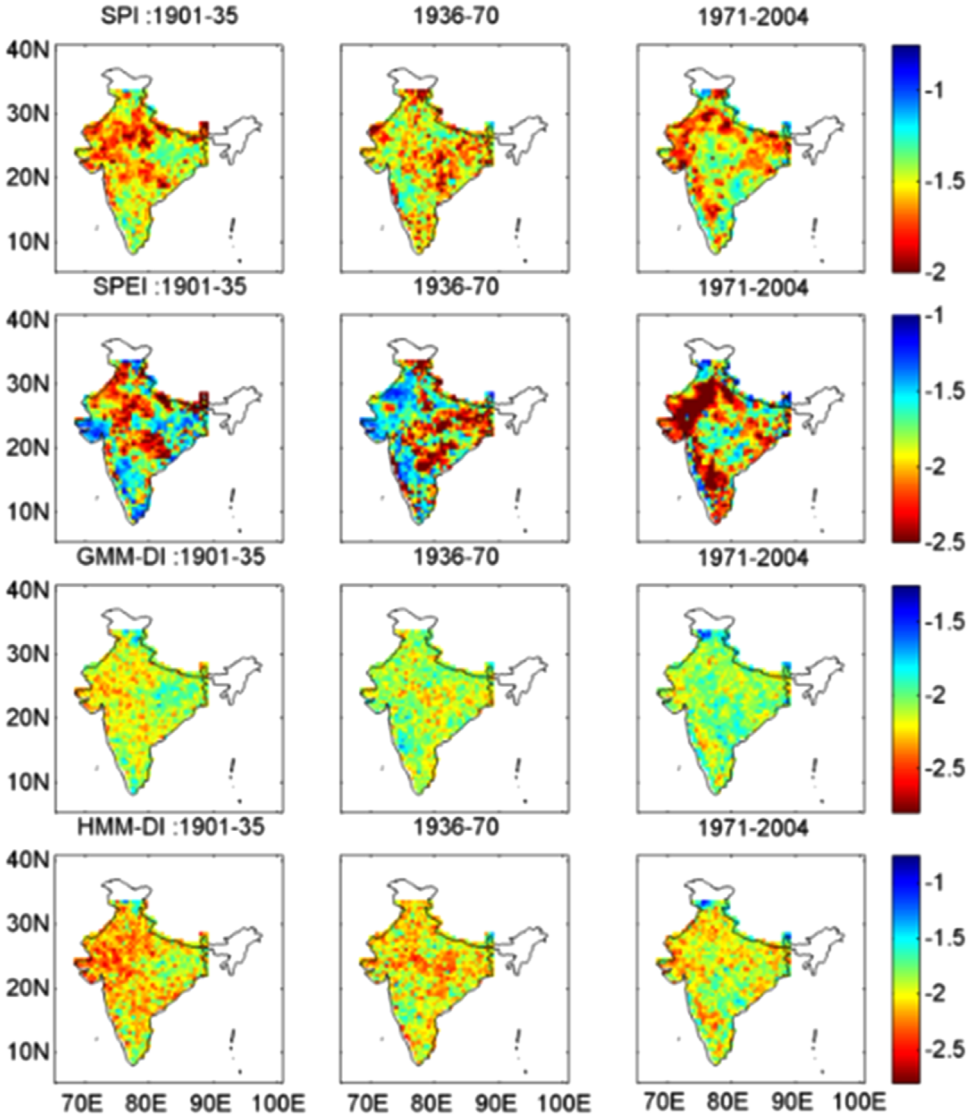


Fig. A.5. (*continued*)

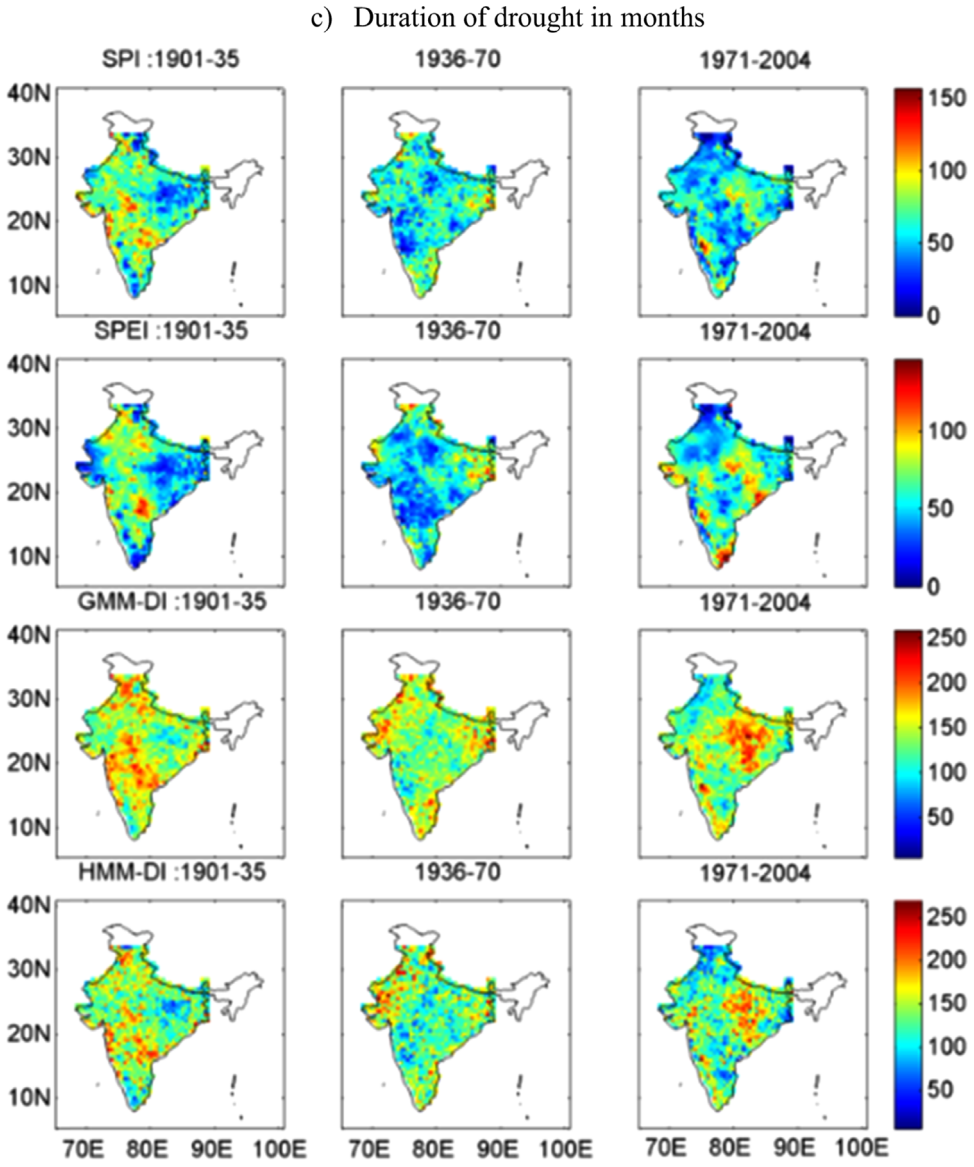


Fig. A.5. (*continued*)

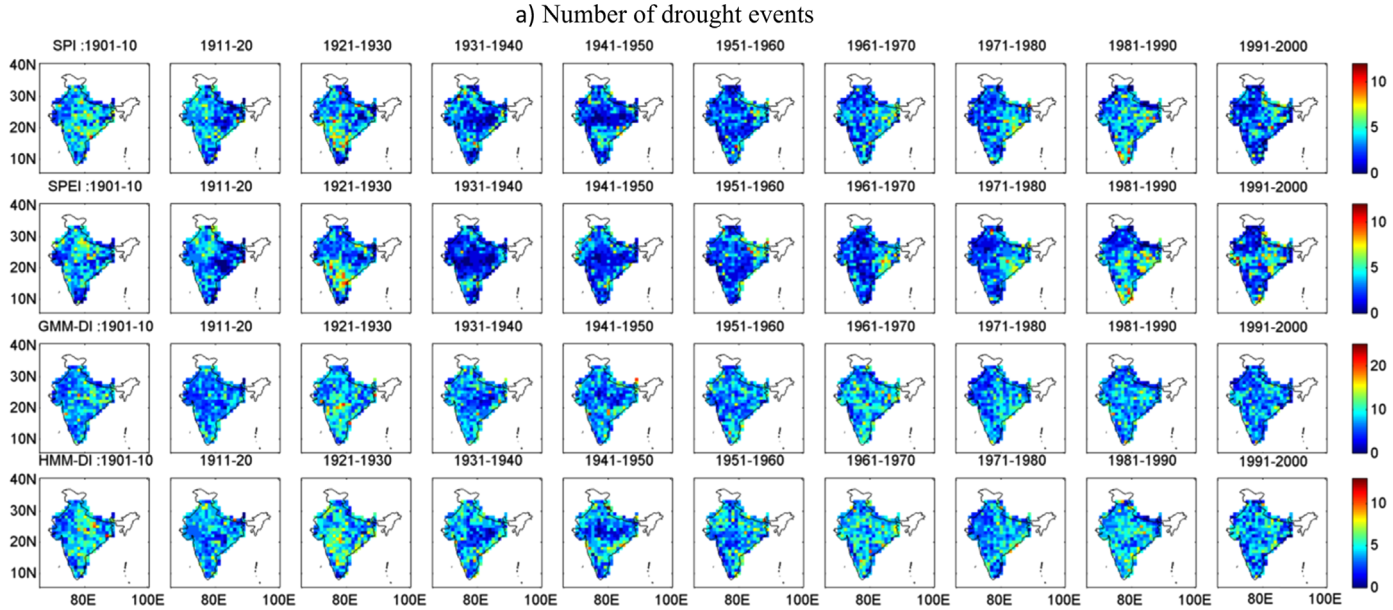
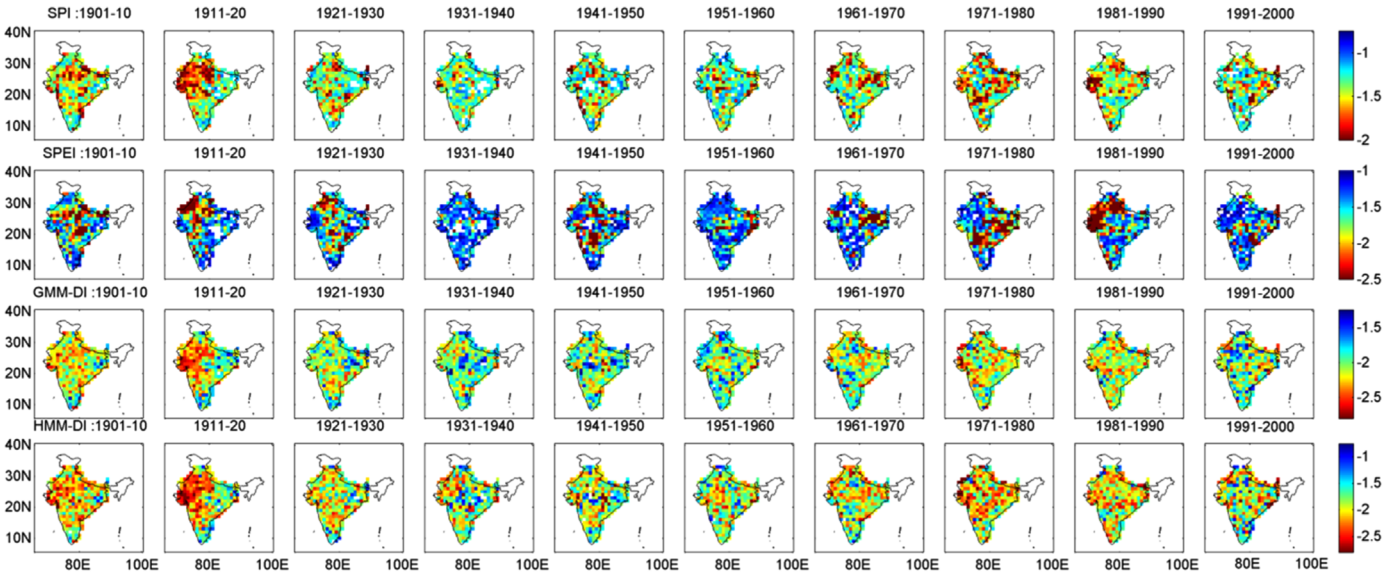


Fig. A.6. Decadal variation in 12-month drought statistics over IMR using IMD dataset where (a) number of drought events, (b) average intensity of drought, and (c) duration of drought in months. In each sub-plot, top panel represents SPI, followed by SPEI, GMM-DI, and HMM-DI.

b) Average intensity of drought



c) Duration of drought in months

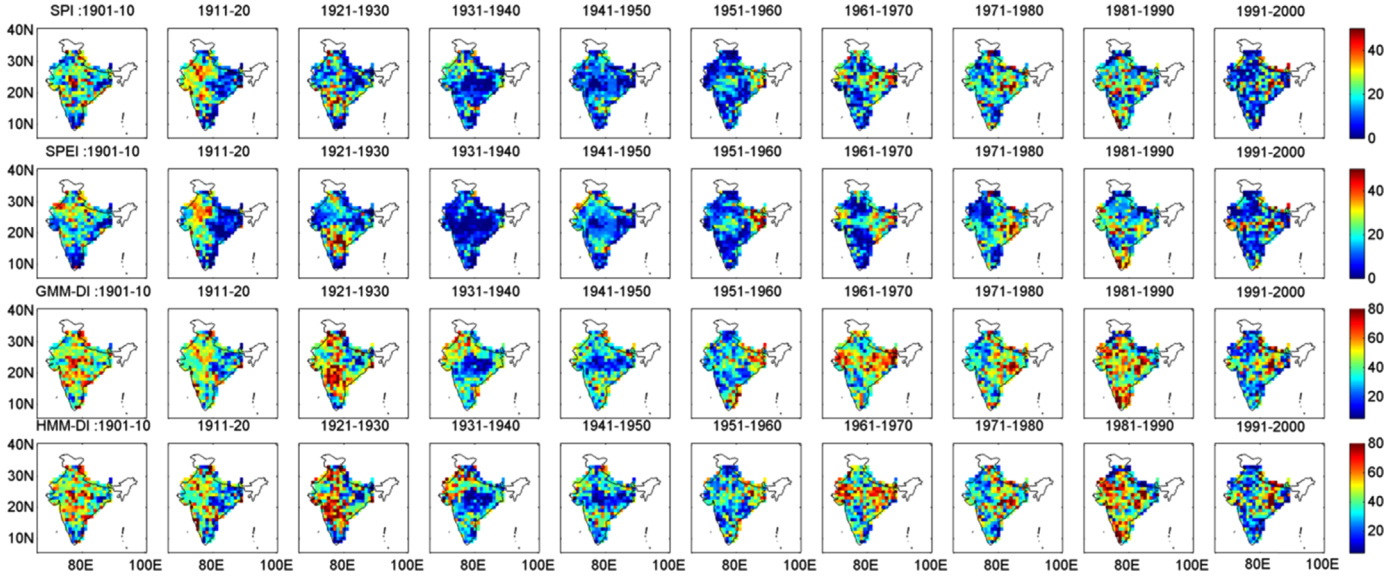


Fig. A.6. (*continued*)

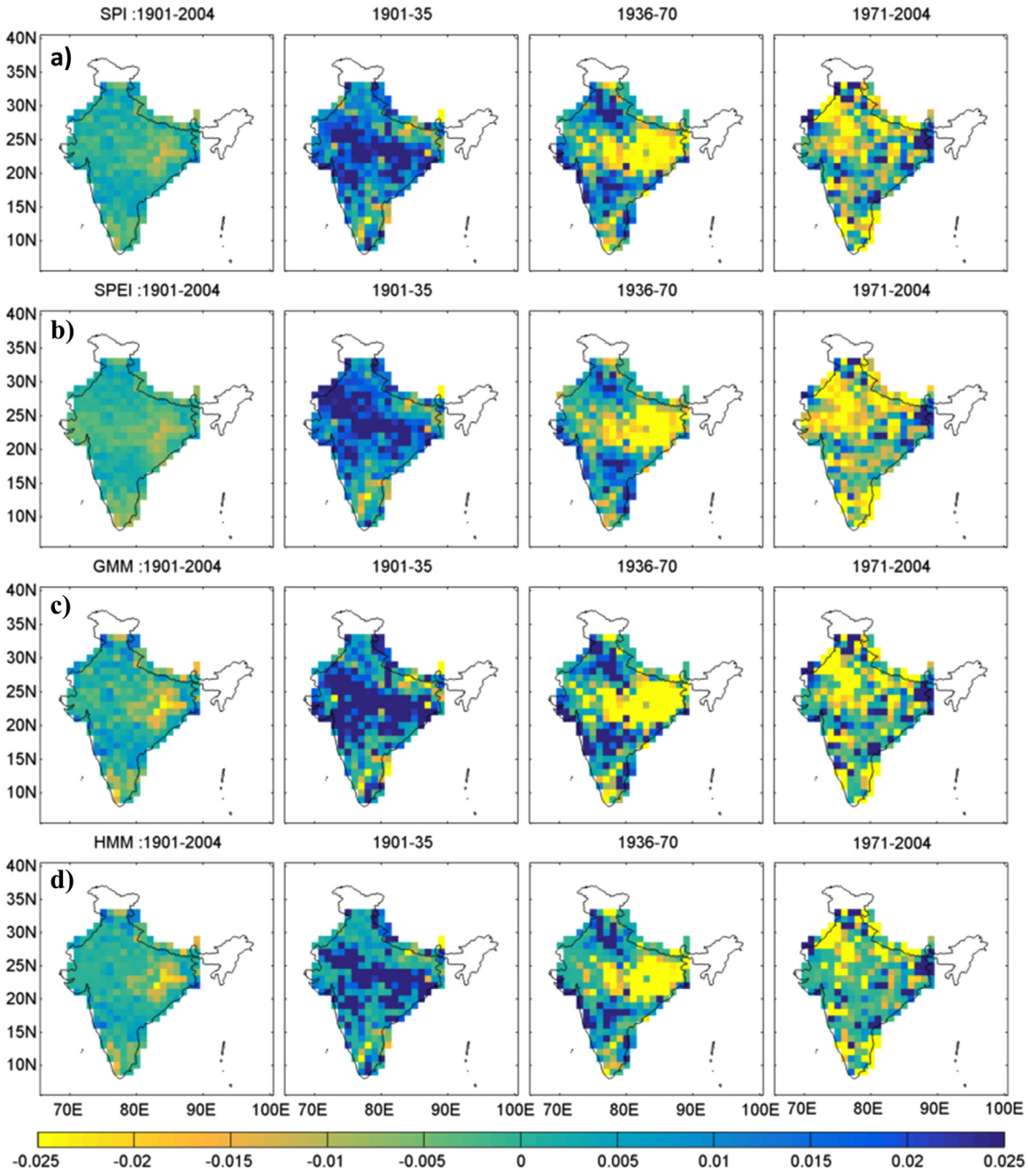


Fig. A.7. Mann-Kendall trend slope for 7-month droughts ending in December over IMR during the periods 1901–2004, 1902–1935, 1936–1970, and 1971–2004. Results correspond to the IMD dataset using (a) SPI, (b) SPEI, (c) GMM-DI, and (d) HMM-DI.

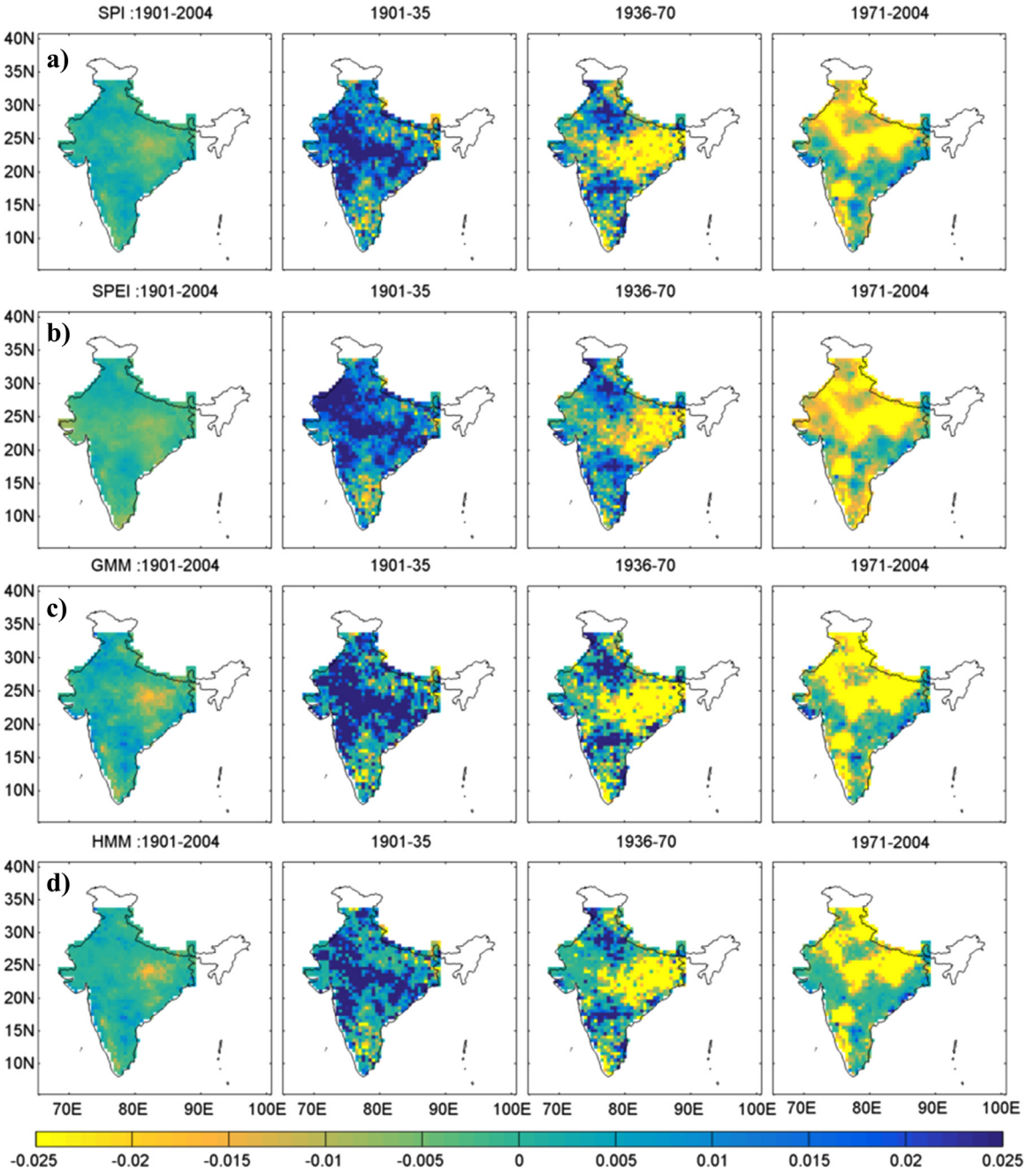


Fig. A.8. Same as Fig. A.7, but using 0.5° University of Delaware precipitation dataset.

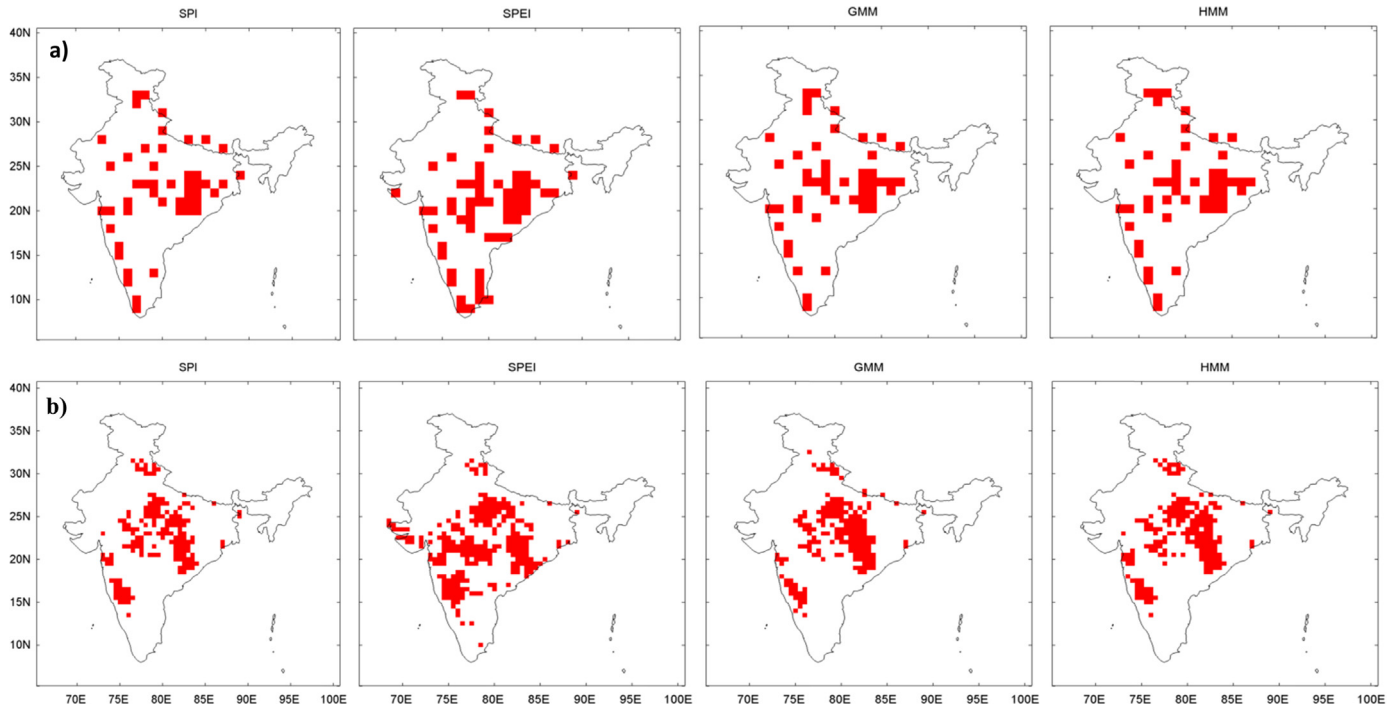


Fig. A.9. Hypothesis test to see if the number of droughts (moderate, severe and extreme) of 7-month time window ending in December have increased during the period 1971–2004 in comparison to 1936–1970 for (a) IMD and, (b) UD precipitation datasets according to SPI, SPEI, GMM-DI, and HMM-DI. Grids where the number of droughts show a statistically significant increase at $\alpha = 0.05$ are displayed.

References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., et al., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. FAO, Rome 300, D05109.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B (Methodological) 57, 289–300.

Burn, D.H., Elnur, M.A.H., 2002. Detection of hydrologic trends and variability. J. Hydrol. 255, 107–122. http://dx.doi.org/10.1016/S0022-1694(01)00514-5.

Center for International Earth Science Information Network-CIESIN-Columbia University, United Nations Food and Agriculture Programme-FAO, Centro Internacional de Agricultura Tropical-CIAT, 2005. Gridded Population of the World, Version 3 (GPWv3): Centroids.

Charusombat, U., Niyogi, D., 2011. A Hydroclimatological Assessment of Regional Drought Vulnerability: A Case Study of Indiana Droughts. Earth Interact. 15, 1–65. http://dx.doi.org/10.1175/2011EI343.1.

Chung, C.E., Ramanathan, V., 2006. Weakening of North Indian SST Gradients and the Monsoon Rainfall in India and the Sahel. J. Climate 19, 2036–2045. http://dx.doi.org/10.1175/JCLI3820.1.

Dai, A., 2013. Increasing drought under global warming in observations and models. Nat. Clim. Change 3, 52–58. http://dx.doi.org/10.1038/nclimate1633.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B (Methodological) 39, 1–38.

De, U.S., Dube, R.K., Rao, P.G.S., 2005. Extreme weather events over India in the last 100 years. J. Indian Geophys. Union 9, 173–187.

Goswami, B.N., Venugopal, V., Sengupta, D., Madhusoodanan, M.S., Xavier, P.K., 2006. Increasing trend of extreme rain events over India in a warming environment. Science 314, 1442–1445. http://dx.doi.org/10.1126/science.1132027.

Guhathakurta, P., Rajeevan, M., 2008. Trends in the rainfall pattern over India. Int. J. Climatol. 28, 1453–1469. http://dx.doi.org/10.1002/joc.1640.

Hamed, K.H., Rao, A.R., 1998. A modified Mann-Kendall trend test for autocorrelated data. J. Hydrol. 204, 182–196. http://dx.doi.org/10.1016/S0022-1694(97)00125-X.

Kishtawal, C.M., Niyogi, D., Tewari, M., Pielke Sr, R.A., Shepherd, J.M., 2010. Urbanization signature in the observed heavy rainfall climatology over India. Int. J. Climatol. 30, 1908–1916. http://dx.doi.org/10.1002/joc.2044.

Kripalani, R.H., Kulkarni, A., Sabade, S.S., Khandekar, M.L., 2003. Indian monsoon variability in a global warming scenario. Nat. Hazard. 29, 189–206.

Kulkarni, A., von Storch, H., 1995. Monte Carlo experiments on the effect of serial correlation on the Mann-Kendall test of trend. Meteorol. Zeitschrift 4, 82–85.

Kumar, K.R., Pant, G.B., Parthasarathy, B., Sontakke, N.A., 1992. Spatial and subseasonal patterns of the long-term trends of Indian summer monsoon rainfall. Int. J. Climatol. 12, 257–268. http://dx.doi.org/10.1002/joc.3370120303.

Liang, Z., Wang, D., Guo, Y., Zhang, Y., Dai, R., 2011. Application of Bayesian Model Averaging Approach to Multi-Model Ensemble Hydrologic Forecasting. J. Hydrol. Eng. 1. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000493.

Li, Z., Zhang, Y., 2011. Application of Gaussian mixture model and estimator to radar-based weather parameter estimations. IEEE Geosci. Remote Sens. Lett. 8, 1041–1045. http://dx.doi.org/10.1109/LGRS.2011.2151250.

Mallya, G., 2011. Hidden Markov Model Based Probabilistic Assessment of Droughts. Purdue University, West Lafayette, IN, p. 47907.

Mallya, G., Tripathi, S., Kirshner, S., Govindaraju, R., 2012. Probabilistic assessment of drought characteristics using a hidden Markov model. J. Hydrol. Eng. . http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000699

McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales. In: Proceedings of Conference of Applied Climatology, American Meteorological Society, Anaheim, CA.

McLachlan, G., Krishnan, T., 1997. The EM Algorithm and Extensions. John Wiley & Sons, Inc., New York.

Mishra, A.K., Singh, V.P., 2010. A review of drought concepts. J. Hydrol. 391, 202–216. http://dx.doi.org/10.1016/j.jhydrol.2010.07.012.

Mishra, V., Smoliak, B.V., Lettenmaier, D.P., Wallace, J.M., 2012. A prominent pattern of year-to-year variability in Indian summer monsoon rainfall. Proc. Natl. Acad. Sci. USA . http://dx.doi.org/10.1073/pnas.1119150109.

Mo, K.C., 2008. Model-based drought indices over the United States. J. Hydrometeorol. 9, 1212–1230. http://dx.doi.org/10.1175/2008JHM1002.1.

Niranjan Kumar, K., Rajeevan, M., Pai, D.S., Srivastava, A.K., Preethi, B., 2013. On the observed variability of monsoon droughts over India. Weather Clim. Extrem. 1, 42–50. http://dx.doi.org/10.1016/j.wace.2013.07.006.

Niyogi, D., Kishtawal, C., Tripathi, S., Govindaraju, R.S., 2010. Observational evidence that agricultural intensification and land use change may be reducing the Indian summer monsoon rainfall. Water Resour. Res. 46, 17, doi: 201010.1029/2008WR007082.

Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass, in: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. The Royal Society, pp. 120–145.

Priestley, C., Taylor, R., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Monthly Weather Rev. 100, 81–92.

Rajeevan, M., 2006. High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells. Curr. Sci. 91, 296.

Rajeevan, M., Bhate, J., Jaswal, A.K., 2008. Analysis of variability and trends of extreme rainfall events over India using 104 years of gridded daily rainfall data. Geophys. Res. Lett. 35, 6, doi: 200810.1029/2008GL035143.

Reynolds, D.A., Rose, R.C., 1995. Robust text-independent speaker identification using Gaussian mixture speaker models. IEEE Trans. Speech Audio Process. 3, 72–83. http://dx.doi.org/10.1109/89.365379.

Roxy, M.K., Ritika, K., Terray, P., Murtugudde, R., Ashok, K., Goswami, B., 2015. Drying of Indian subcontinent by rapid Indian Ocean warming and a weakening land-sea thermal gradient. Nat. Commun., 6.

Rupa Kumar, K., Sahai, A.K., Krishna Kumar, K., Patwardhan, S.K., Mishra, P.K., Revadekar, J.V., Kamala, K., Pant, G.B., 2006. High-resolution climate change scenarios for India for the 21st century. Curr. Sci. 90, 334–345.

Sheffield, J., Wood, E.F., Roderick, M.L., 2012. Little change in global drought over

the past 60 years. Nature 491, 435–438. http://dx.doi.org/10.1038/nature11575.
Singh, D., Tsiang, M., Rajaratnam, B., Diffenbaugh, N.S., 2014. Observed changes in extreme wet and dry spells during the South Asian summer monsoon season. Nat. Clim. Change 4, 456–461. http://dx.doi.org/10.1038/nclimate2208.
Stephenson, D., 2001. Searching for a fingerprint of Global Warming in the Asian summer monsoon. MAUSAM 52, 213–213–220.
Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. Geogr. Rev., 55–94.
Trenberth, K.E., Dai, A., van der Schrier, G., Jones, P.D., Barichivich, J., Briffa, K.R., Sheffield, J., 2014. Global warming and changes in drought. Nat. Clim. Change 4, 17–22. http://dx.doi.org/10.1038/nclimate2067.
Ventura, V., Paciorek, C.J., Risbey, J.S., 2004. Controlling the proportion of falsely rejected hypotheses when conducting multiple tests with climatological data. J. Clim. 17, 4343–4356. http://dx.doi.org/10.1175/3199.1.
Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. J. Clim. 23, 1696–1718.
Wilks, D.S., 2006. On field significance and the false discovery rate. J. Appl. Meteorol. Climatol. 45, 1181–1189. http://dx.doi.org/10.1175/JAM2404.1.
Yue, S., Wang, C.Y., 2002. Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test. Water Resour. Res. 38, 1068. http://dx.doi.org/10.1029/2001WR000861.