Analysis of ENSO-based climate variability in modulating drought risks over western Rajasthan in India

Poulomi Ganguli and M Janga Reddy*

Department of Civil Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India.
*Corresponding author. e-mail: mjreddy@civil.iitb.ac.in

This paper investigates the role of El Niño-Southern Oscillation (ENSO)-based climate variability in modulating multivariate drought risks in the drought-prone region of Western Rajasthan in India. Droughts are multivariate phenomenon, often characterized by severity, duration and peak. By using multivariate ENSO index, annual drought events are partitioned into three climatic states – El Niño, La Niña and neutral phases. For multivariate probabilistic representation of drought characteristics, trivariate copulas are employed, which have the ability to preserve the dependence structure of drought variables under uncertain environment. The first copula model is developed without accounting the climate state information to obtain joint and conditional return periods of drought characteristics. Then, copula-based models are developed for each climate state to estimate the joint and conditional probabilities of drought characteristics under each ENSO state. Results of the study suggest that the inclusion of ENSO-based climate variability is helpful in knowing the associated drought risks, and useful for management of water resources in the region.

1. Introduction

In hydrological studies, drought is a climatic anomaly, caused by either sub-normal rainfall, erratic rainfall distribution, higher water need or a combination of all these factors. In a large country like India where precipitation varies both in space and time, drought is one of the most frequently occurring natural calamities in various parts of the country. About two-thirds of the geographic area of India receives low rainfall (less than 1000 mm), which is also characterized by uneven and erratic distributions. Out of net sown area of 140 million hectares, about 68% is reported to be vulnerable to drought conditions and about 50% of such vulnerable area is classified as 'severe', where frequency of drought is almost regular (Murthy and Sesha Sai 2010).

As drought is a multivariate phenomenon characterizing severity, duration and peak; drought properties need to be modelled using effective probability models. Recently, copulas have been successfully applied in the field of hydrology for multivariate analysis of hydrological systems, viz., flood frequency analysis and drought frequency analysis (Kao and Govindaraju 2010; Mirakbari *et al.* 2010; Wong *et al.* 2010). By using copulas, some of the limitations of traditional multivariate distributions can be relaxed as suggested in recent studies (Favre *et al.* 2004; Genest and Favre 2007). Shiau (2006) investigated bivariate joint distribution of drought properties severity and duration in Southern Taiwan using standardized precipitation index (SPI) and theory of copulas. Shiau *et al.* (2007) performed bivariate frequency analysis of hydrological droughts using monthly stream flow data

of Yellow River basin employing theory of runs and Archimedean class of Clayton copula. Laux *et al.* (2009) analysed regional nature of droughts in Volta basin using effective drought index calculated from daily precipitation data in order to assess drought properties of five different rainfall regions within Volta basin. In their study, bivariate Archimedean class of Clayton copula family is chosen to construct joint distribution and associated joint return periods. Serinaldi *et al.* (2009) modelled four-dimensional joint distribution of drought properties – mean SPI, duration, minimum SPI, drought areal extent and computed joint return periods using elliptical class of Student's *t* copula for different degrees of freedom. Shiau and Modarres (2009) employed Archimedean class of Clayton copula to investigate the nature of S–D–F curves of two different climatic regions in Iran due to its simplified structure. Song and Singh (2010) modelled bivariate joint probability distribution of drought properties in Texas using meta-elliptical class of copulas and found that meta-Gaussian copula performed satisfactorily in modelling the dependence. Kao and Govindaraju (2010) proposed copula-based joint deficit index using precipitation and stream flow marginals with window sizes varying from 1 to 12 months in Indiana watershed. The standardized index is adopted for statistical analysis of hydrologic variables. The temporal dependence structure among hydrologic variables is modelled using elliptical class of Student's *t* copula due to large number of variables involved in the study and its ability to model tail dependence. Wong *et al.* (2010) investigated effect of El Niño-Southern Oscillation (ENSO) phenomenon on nature of multivariate drought frequencies using precipitation data from two districts in New South Wales, Australia. The performance of trivariate Gumbel–Hougaard and Student's *t* copulas were tested for each climatic phase and limitations of each copula model was discussed. The trivariate Gumbel–Hougaard copula has a limitation that weak dependences are averaged in asymmetric trivariate structure and there is a restriction in application as the correlation between two pairs should be identical and lower than the third pair. In contrast, trivariate *t* copula has no such restriction except that variance-covariance matrix has to be positive definite. A small difference was observed between Gumbel–Hougaard and Student's *t* copulas in distance-based goodness-of-fit measures and the latter emerged as a better model. The fitted models were then used to estimate annual recurrence intervals taking into account at least one of the three drought variables and all of the three variables exceeding critical values. Madadgar and Moradkhani (2011) investigated the effect of climate change on drought characteristics using the

copula method in Oregon's Upper Klamath River basin. The performance of two trivariate copulas – Gumbel–Hougaard and Student's *t* were tested in modelling joint distribution of drought properties – severity, duration and intensity for analysing streamflow-based drought indices. Results showed that both Gumbel–Hougaard and Student's *t* copulas performed similarly for historical (1920–2009) time period. To evaluate potential impact of climate change, five General Circulation Model (GCM), outputs under A1B emission scenario were used for multivariate drought analysis. The return period analysis based on bivariate and trivariate copulas showed that climate change causes an overall decline in drought severity and duration in the Upper Klamath River basin during projected time period 2020–2090. Lee *et al.* (2012) studied tail dependence of four different copula functions – Gumbel–Hougaard, Frank, Clayton and Gaussian copulas for bivariate drought frequency analysis in Canada and Iran. Their study showed that Clayton copula may not be suitable for modelling droughts as dependence between two variables in the upper tail of Calyton copula was found to be very weak and similar to the independent case, whereas Frank and Gumbel–Hougaard copula showed better performance for modelling bivariate drought characteristics.

The present study investigates the role of ENSO-based climate variability in influencing multivariate drought frequencies in the drought-prone region of Western Rajasthan in India. The state has the maximum probability of drought occurrence in India, with recurring droughts of about 3–4 years in a cycle of 5 years (Mall *et al.* 2006). Studies conducted by Rathore (2004), identified 48 drought years of varied intensity in the state during the period 1901–2002. During the year 2002, when about 29% of the total area of the country was affected by drought, seasonal departures of precipitation for west and east Rajasthan were −71% and −60%, respectively (RAPCC 2012). During the year 2001, the total monsoon rainfall in Rajasthan was 381 mm and 30,583 villages were affected by drought, whereas in 2002, the monsoon rainfall was only 173 mm and 41,000 villages faced drought situation (Khera 2004).

2. Large-scale climatic phenomena – ENSO

Drought is a normal part of natural climate variations. Past research studies based on tree-rings and other instrumental records have shown that many major droughts which occurred over different parts of the world are triggered by atmospheric teleconnection such as ENSO phenomenon (Dai 2011). ENSO refers to oscillation between a warm phase

(El Niño) characterized by abnormal warming of surface ocean waters of the central and eastern Pacific and enhanced convection in the atmosphere above; and a cold phase (La Niña) characterized by abnormal cooling of these waters and suppressed convection in the atmosphere above (Gadgil *et al.* 2007). Several researchers have explored the links between Indian monsoon rainfall and ENSO events (Webster and Yang 1992; Rajeevan *et al.* 2004). There have been many indices defined to represent ENSO events. Two well-known ENSO indices are southern oscillation index based upon sea surface pressure differences at Tahiti (18°S 150°W) and Darwin (12°S 131°E) from 1880 onwards; and the Niño index from 1950 onwards based upon sea-surface temperature (SST) anomalies (°C) of the central and eastern equatorial Pacific.

A recent index for describing ENSO phenomena is the multivariate ENSO index (MEI). Computation of MEI involves moving average of 2-months for the time series under consideration. The MEI represents the first unrotated principal component (PC) of all six observed climatic parameters over the tropical Pacific: sea-level pressure, north-south component of surface wind, east-west component of surface wind, SST, surface air temperature and total cloudiness fraction of the sky (Wolter and Timlin 2011). Inclusion of these climatic variables helps MEI to explain the ocean–atmospheric interactions better than the indices that rely on single climatic variable. Large positive values of the MEI represent El Niño, while negative values correspond to La Niña episodes.

3. Copula

Copulas are parametrically specified joint distribution functions obtained by linking marginal distribution of any form. If $X = (X_1, \ldots, X_d)$ is a random vector with continuous marginal cumulative distribution functions (CDF) $F_1, \ldots, F_d$, then joint distribution $H(\boldsymbol{X})$ can be expressed by Sklar's theorem (Sklar 1959) as:

$$H(\boldsymbol{X}) = C\{F_1(x_1), \ldots, F_d(x_d); \theta\}, \quad \boldsymbol{X} \in R^d, \quad (1)$$

where the function $C:[0, 1]^d \to [0, 1]$ is called a d-dimensional *copula*, with association parameter θ. There exists different class of copulas in the literature, viz., Archimedean, Plackett, elliptical copulas, etc. Though Archimedean class of copulas are simple and can be easily generated, extension to higher orders in symmetric form is limited (Grimaldi and Serinaldi 2006). In this study, elliptical class of Student's t copula is chosen due to its capability to model multivariate distribution with asymmetric dependence structure and can capture upper tail dependence quite well (Nelsen 2006).

3.1 *Student's t copula*

If $\sum \in R^d$ for $x \in R^d$ denotes a symmetric positive definite shape matrix, then multivariate Student's t copula for marginals $u = (u_1, \ldots, u_d) \in [0, 1]^d$ with ϑ degrees of freedom is defined as (Mashal and Zeevi 2002)

$$C(u_1, u_2, u_3; \vartheta, \Sigma) = t_{\vartheta, \Sigma}^d(t_\vartheta^{-1}(u_1), t_\vartheta^{-1}(u_2), t_\vartheta^{-1}(u_3))$$

$$= \int_{-\infty}^{t_\vartheta^{-1}(u_1)} \int_{-\infty}^{t_\vartheta^{-1}(u_2)} \int_{-\infty}^{t_\vartheta^{-1}(u_3)} \frac{\Gamma((\vartheta + d)/2)}{\Gamma(\vartheta/2)(\vartheta\pi)^{d/2} |\Sigma|^{1/2}}$$

$$\times (1 + y^T \Sigma^{-1} y/\vartheta)^{-(\vartheta+d)/2} dy_1 dy_2 dy_3, \quad (2)$$

where $d = 3$; $y = \{y_1, y_2, y_3\}$; ϑ and $\sum$ are parameters of Student's t copula.

The parameter of $t-$copula is estimated using a two-step transformation procedure. The shape parameter matrix $\sum$ consists of elements $\hat{\sigma}_{i,j}$, which are functions of rank correlation coefficient. For trivariate case, the elements of shape parameter matrix $\sum \in \{\sigma_{11}\ \sigma_{12}\ \sigma_{13}; \sigma_{21}\ \sigma_{22}\ \sigma_{23}; \sigma_{31}\ \sigma_{32}\ \sigma_{33}\}$, and the elements $\hat{\sigma}_{i,j}$ are estimated using the relationship $\hat{\sigma}_{i,j} = \sin\left(\frac{\pi}{2}\hat{\tau}_{ij}\right)$, where $\hat{\tau}_{ij}$ is the pair-wise Kendall's dependence measure or rank correlation coefficient. Obtaining $\hat{\sum}$ in a three-dimensional case involves computing $\hat{\sigma}_{ij}$ for $i, j \in \{1, 2, 3\}$. Then, a numerical search technique is employed for estimating ϑ

$$l(\hat{\vartheta}) = \underset{\vartheta \in (2, \infty]}{\arg \max} \sum_{i=1}^{n} \ln\left[c_\theta\left\{(U_{i,1}, U_{i,2}, U_{i,3}) | \hat{\vartheta}, \hat{\Sigma}\right\}\right], \quad (3)$$

where $U_{i,d}$ denotes empirical CDF of dth random variable, and c_θ is the copula density.

3.2 *Testing suitability of the copula family*

After fitting copula, theoretical (fitted copula model) and observed probabilities (empirical observations) are compared. The empirical copula can be defined as (Genest and Favre 2007)

$$C_n(u_1, u_2, u_3)$$
$$= \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left(\hat{U}_{i,1} \le u_1, \hat{U}_{i,2} \le u_2, \hat{U}_{i,3} \le u_3\right),$$
$$u_1, u_2, u_3 \in [0, 1]. \quad (4)$$

Apart from distance-based test, the parametric and non-parametric tail dependence tests are performed for the fitted copula model. The tail dependence coefficient (TDC) of copulas provides a measure of strength of dependence in the tails of a bivariate distribution (McNeil *et al.* 2005). The parametric upper TDC of Student's

t copula with ϑ degrees of freedom and $\hat{\sum} \in \{\sigma_{11}\,\sigma_{12}\,\sigma_{13};\ \sigma_{21}\,\sigma_{22}\,\sigma_{23};\ \sigma_{31}\,\sigma_{32}\,\sigma_{33}\}$ is given by

$$\lambda_{U_{ij}}^{\text{param}} = 2 - 2t_{\vartheta+1}\left(\sqrt{(\vartheta+1)}\sqrt{\frac{(1-\sigma_{ij})}{(1+\sigma_{ij})}}\right),\quad (5)$$

where $t_{\vartheta+1}$ is CDF of Student's t distribution with $\vartheta+1$ degrees of freedom. For non-parametric tail dependence test, Capéraá-Fougéres-Genest (CFG) estimator (Capéraá *et al.* 1997; Frahm *et al.* 2005) is employed. If $\{(u_1, v_1), \ldots, (u_n, v_n)\}$ are random samples obtained from Copula $C\,(\bullet)$, the bivariate upper TDC using CFG estimator (λ_U^{CFG}) is given by

$$\hat{\lambda}_U^{\text{CFG}} = 2 - 2\exp\left[\frac{1}{n}\sum_{i=1}^{n}\log\left\{\sqrt{\log\left(\frac{1}{u_i}\right)\log\left(\frac{1}{v_i}\right)}\right.\right.$$
$$\left.\left./\log\left(\frac{1}{\max(u_i,v_i)^2}\right)\right\}\right]. \quad (6)$$

3.3 *Trivariate frequency analysis of droughts using copulas*

Multivariate frequency analysis is helpful in understanding critical behaviour of drought characteristics. As drought events may last for more than a year, drought characteristics can be analysed as a partial duration series (PDS) of independent events. The univariate return period of drought for a specific drought variable (say severity) can be computed by (Kim *et al.* 2003)

$$T_s = \frac{N}{n(1-F_s(s))} = \frac{\delta}{1-F_s(s)}, \quad (7)$$

where $\delta = N/n$ and N = total length of SPI time series (years), n = number of drought events, $F_s(s)$ is the univariate CDF of drought severity.

By following the above definition, the copula-based trivariate return period of drought for exceeding thresholds of any one of the drought variables (in OR-case) without accounting climatic state is given by

$$T_{SDI}^{\cup} = \frac{\delta}{1-F_{SDI}(s,d,i)} = \frac{\delta}{1-C_{SDI}(s,d,i)}$$
$$= \frac{\delta}{1-P_{SDI}^{\cup}}, \quad (8)$$

where $P_{SDI}^{\cup} = P\,(S \leq s \cup D \leq d \cup I \leq i)$ denotes joint probability of occurrence of any one of the drought variables, i.e., either severity (S) or duration (D) or peak (I).

Similarly, multivariate return period of drought for exceeding thresholds of all drought variables simultaneously (in AND-case) without accounting climate state is given by

$$T_{\text{SDI}}^{\cap} = \frac{\delta}{\begin{array}{c}1 - F_{\text{s}}(s) - F_{\text{D}}(d) - F_{\text{I}}(i) + F_{\text{SD}}(s,d)\\ + F_{\text{DI}}(d,i) + F_{\text{SI}}(s,i) - F_{\text{SDI}}(s,d,i)\end{array}}$$
$$= \frac{\delta}{\begin{array}{c}1 - F_{\text{s}}(s) - F_{\text{D}}(d) - F_{\text{I}}(i) + C_{\text{SD}}(s,d)\\ + C_{\text{DI}}(d,i) + C_{\text{SI}}(s,i) - C_{\text{SDI}}(s,d,i)\end{array}}$$
$$= \frac{\delta}{1-P_{\text{SDI}}^{\cap}}, \quad (9)$$

where $C_{SD}\,(s,\,d)$, $C_{DI}\,(d,\,i)$ and $C_{SI}\,(s,\,i)$ are the joint distributions obtained from bivariate copula for severity-duration, duration-peak and severity-peak combinations, respectively; $P_{SDI}^{\cap} = P(S \leq s \cap D \leq d \cap I \leq i)$ denotes joint probability of occurrence of all the drought variables simultaneously.

By using Bayes rule for obtaining total probability (from three ENSO state conditional probabilities) and concepts from Willems (2000), the mean annual multivariate return period of drought accounting climate state information (i.e., compound return period) in OR-case can be written as

$$T_{\text{SDI}}^{\cup,\text{Climate}} = \frac{\delta}{1 - \left(\begin{array}{l}p_{\text{El Niño}}P_{\text{SDI}\{\{\text{El Niño}\}}^{\cup}\\ + p_{\text{La Niña}}P_{\text{SDI}\{\text{La Niña}\}}^{\cup}\\ + p_{\text{Neutral}}P_{\text{SDI}\{\text{Neutral}\}}^{\cup}\end{array}\right)}, \quad (10)$$

where $P_{\text{SDI}\{\text{El Niño}\}}^{\cup}$, $P_{\text{SDI}\{\text{La Niña}\}}^{\cup}$ and $P_{\text{SDI}\{\text{Neutral}\}}^{\cup}$ denote joint probability of occurrence of any one of the drought variables at El Niño, La Niña and neutral climatic phases, respectively; $p_{\text{ENSO State}}$ denotes probability of occurrence of drought in a particular climate episode; ENSO State = {El Niño, La Niña, Neutral}. The $p_{\text{ENSO State}}$ can be obtained by dividing the number of drought events occurring at a particular climate state by the total number of drought events during the period of study. Similarly, mean annual recurrence interval of drought in AND- case incorporating climate episode can be written as

$$T_{\text{SDI}}^{\cap,\text{Climate}} = \frac{\delta}{1 - \left(\begin{array}{l}p_{\text{El Niño}}P_{\text{SDI}\{\text{El Niño}\}}^{\cap}\\ + p_{\text{La Niña}}P_{\text{SDI}\{\text{La Niña}\}}^{\cap}\\ + p_{\text{Neutral}}P_{\text{SDI}\{\text{Neutral}\}}^{\cap}\end{array}\right)}, \quad (11)$$

where $P_{\text{SDI}\{\text{El Niño}\}}^{\cap}$, $P_{\text{SDI}\{\text{La Niña}\}}^{\cap}$ and $P_{\text{SDI}\{\text{Neutral}\}}^{\cap}$ denote joint probability of occurrence of all variables

simultaneously at El Niño, La Niña and neutral climatic phases, respectively.

3.4 *Conditional return periods of drought characteristics*

The copula-based joint distribution can be used to obtain conditional return periods of drought characteristics. For example, joint return period of drought severity and duration conditional on drought peak $(T_{S,D|I})$ can be obtained from

$$T_{S,D|I} = \frac{\delta}{(1 - F_{S,D|I}(s,d|I \leq i'))}, \quad (12)$$

where

$$
\begin{aligned}
F_{S,D|I} &= P\left[S \leq s, D \leq d | I \leq i'\right] \\
&= \frac{F_{S,D,I}(s,d,i')}{F_I(i')} \\
&= \frac{C_{SDI}(s,d,i')}{F_I(i')}.
\end{aligned}
$$

Similarly, equivalent formula for conditional return period of drought severity and peak, given drought duration $(T_{S,I|D})$, and conditional return period of drought duration and peak, given severity $(T_{D,I|S})$ can also be obtained.

3.5 *Conditional probability of drought characteristics under each ENSO state*

Conditional distribution of drought characteristics under a particular ENSO state can be derived from copula-based joint distribution function fitted for drought characteristics in that climate state, which can be helpful in studying the influence of one dependent variable on another. The conditional probability of drought severity, given drought duration and peaks exceeding certain thresholds d' and i', respectively (under a specific climate state), is expressed as

$$
\begin{aligned}
&P\left(S \leq s | D \geq d', I \geq i'\right) \\
&= \frac{F_S(s) - F_{SD}(s,d') - F_{SI}(s,i') + F_{SDI}(s,d',i')}{1 - F_D(d') - F_I(i') + F_{DI}(d',i')} \\
&= \frac{F_S(s) - C_{SD}(s,d') - C_{SI}(s,i') + C_{SDI}(s,d',i')}{1 - F_D(d') - F_I(i') + C_{DI}(d',i')}
\end{aligned}
$$
$$(13)$$

$$
\begin{aligned}
&P(S \leq s, I \leq i | D \geq d') \\
&= \frac{F_{SI}(s,i) - F_{SDI}(S,d',i)}{1 - F_D(d')} \\
&= \frac{C_{SI}(s,i) - C_{SDI}(s,d',i)}{1 - F_D(d')}.
\end{aligned}
$$
$$(14)$$

4. Application

4.1 *Study area and data*

Western Rajasthan with an area of 196,150 km² occupies 57.31% of India's total arid zone area. The climate is characterized by low, highly variable and ill-distributed rainfall, high wind speed, high evaporation losses, and extremes of seasonal temperatures. Rajasthan has only 1% of the country's total surface water resources. The monsoon period is short (about 2 to 3 months, July–September), resulting in annual rainfall ranging from 150–900 mm in different parts of the state (average annual precipitation 576 mm) and temperature varies from 5°–45°C in different seasons (RACP 2012).

The monthly area-weighted precipitation data of nine rainfall stations in western Rajasthan meteorological subdivision for about 110 years (from 1896 to 2005) is obtained from Indian Institute of Tropical Management, Pune (http://www.tropmet.res. in). Map of the study area showing locations of the rain gauge stations is presented in figure 1. Relevant data for MEI of about 135 years (from 1871 to 2005) are obtained from NOAA website (http://www.esrl.noaa.gov/psd/enso/mei.ext/index.html#data). MEI values before 1950 (extended back to 1871) are based on Hadley Centre sea-level pressure and SST (Wolter and Timlin 2011).

4.2 *Drought modelling using standardized precipitation index*

Drought events are identified using standardized precipitation index, which is a commonly used indicator for drought assessment. Calculation of SPI for any location and time scale (such as 3, 6, 9 and 12 months) involves fitting a probability distribution function (generally Gamma or Pearson Type III distribution) for long time series of aggregated precipitation which is then transformed into a standardized normal distribution so that the mean SPI for the location and desired period is zero. In this study, 6-month aggregated precipitation data is fitted with Gamma distribution function and used for SPI computation. A drought period is identified when SPI value reaches 20 percentile or below threshold level (Svoboda *et al.* 2002). Drought duration (D) is taken as the number of consecutive intervals (months), where SPI remains below the specified threshold value. Drought severity (S) is the cumulative value of SPI within the drought duration. For convenience, the severity of drought event i, is taken as positive and given as $S_i = -\sum_{t=1}^{D} SPI_{i,t}, \forall i = 1, \ldots, n$. Drought peak (I) is the absolute value of the minimum value taken by the SPI over the duration of the drought.
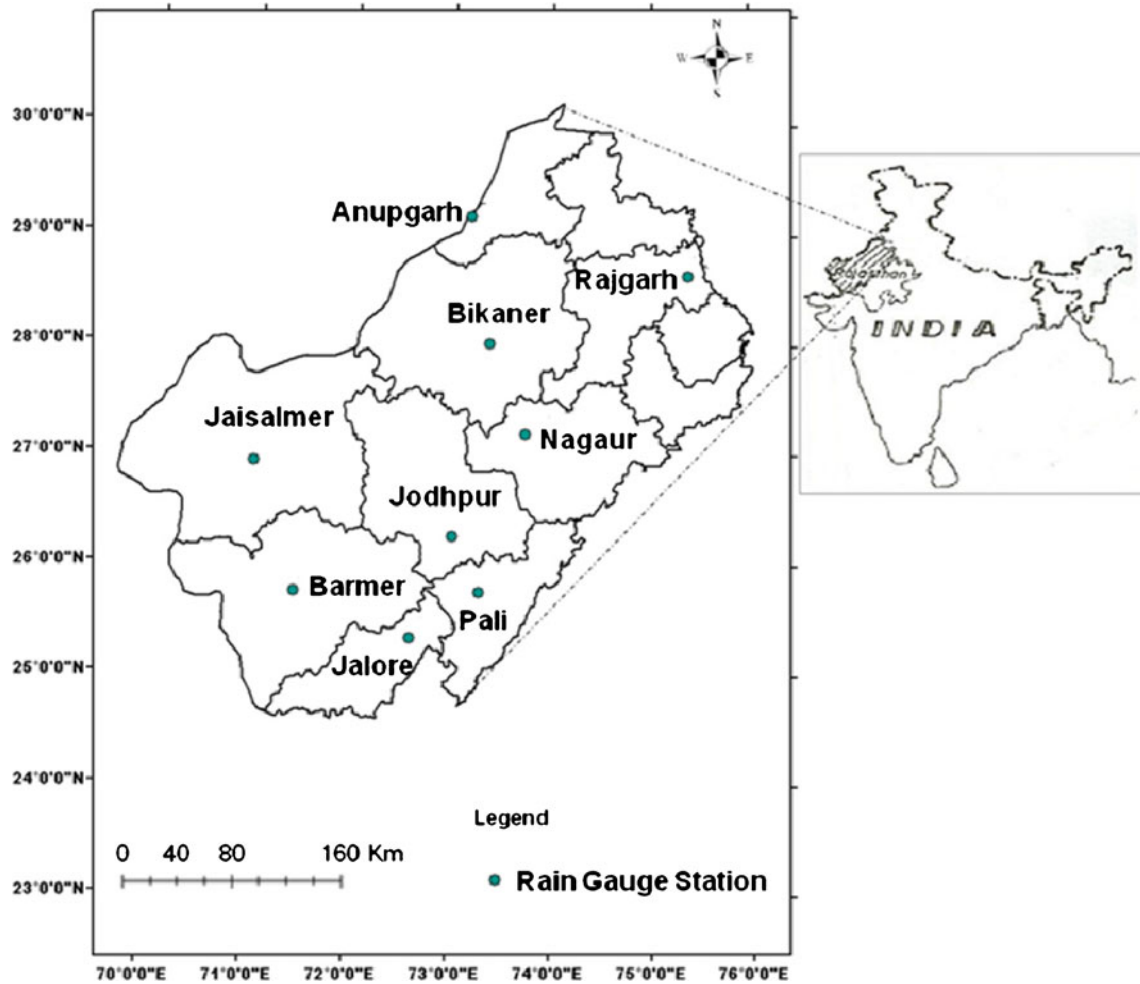
Figure 1. Location map of the study area showing rain gauge stations.

Figure 2 illustrates the definitions of drought characteristics. More details on SPI and definitions of the drought characteristics and their computation are well-documented by Janga Reddy and Ganguli (2012).

For data of the present study region, it is noticed that SPI-6 time series reaches 20 percentile limit at −0.8; hence it is taken as a threshold level for drought identification. Monthly SPI-6 series were calculated and 87 drought events were identified. The time series plots of SPI-6 and MEI time series are shown in figure 3. From figure 3a, it can be observed that the region had experienced major drought events during 1899–1900, 1901–1902, 1905–1906, 1911–1912, 1915–1916, 1918–1919, 1920–1921, 1939, 1946–1947, 1968–1970, 1973 and 2002–2003.

The three climate states (El Niño, La Niña and neutral phases) are assumed to be mutually exclusive and exhaustive during the study period so that each year belongs to one particular ENSO episode. Every year from 1896–2005 is given an ENSO classification based on May–November MEI index

(Wolter and Timlin 2011). The non-parametric Spearman's rank correlation is computed between MEI and SPI-6 time series at each ENSO phase year during 1896–2005. The rank correlations between SPI-6 and MEI time series at El Niño, La Niña and neutral phase years are found to be −0.14, −0.30 and −0.14 with corresponding p-values $5.11e^{-4}$, $2.3e^{-8}$ and 0.008, respectively. All correlations are statistically significant at 5% significance levels as tested by standard two-tailed t test, which indicates that drought in the study region is responsive to ENSO phenomenon. Figure 4 presents scatter plot of SPI-6 against MEI time series at each climate phase, which depicts the strength of association between the two time series. Using SPI-6 time series, a total of 87 drought events were identified during the study period. Among 87 drought events, 41 are classified under El Niño and 23 are classified under La Niña and the remaining under neutral episodes. Table 1 presents summary statistics of drought without and with accounting ENSO state. From table 1, it can be noticed that the droughts in
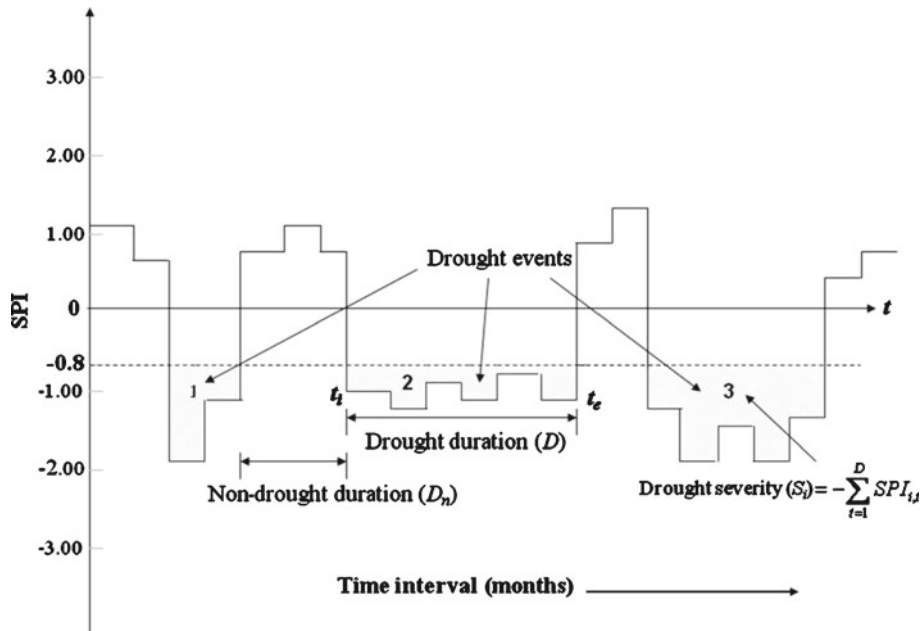
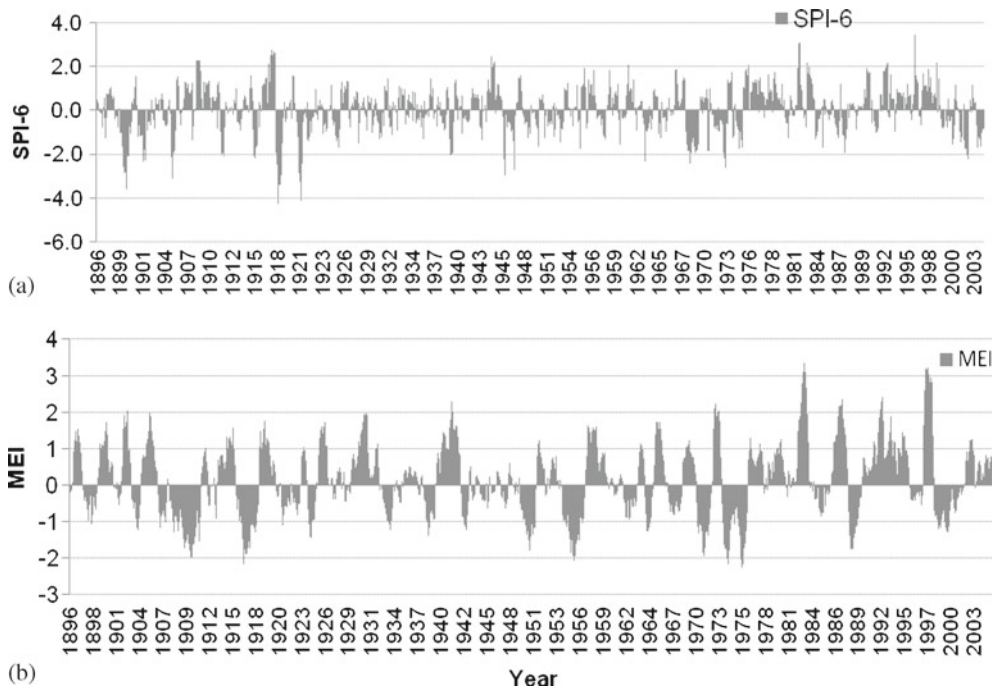Figure 2. Illustration of drought characteristics using SPI.

Figure 3. Time series plot of (**a**) SPI-6 and (**b**) MEI during study period (1896–2005).

El Niño episodes are longer and severe in nature, whereas the opposite is observed for La Niña year droughts. Neutral phase droughts lie in between El Niño and La Niña episode droughts in terms of severity and duration. Table 2 presents association between drought properties using Pearson's linear correlation and the two non-parametric dependence measures – Spearman's ρ and Kendall's τ with and without accounting ENSO phases. In all cases, corresponding p-value is less than 0.0001, indicating significant positive association between drought variables. Figure 5 presents scatter plots of pair-wise dependent drought variables at each ENSO state.
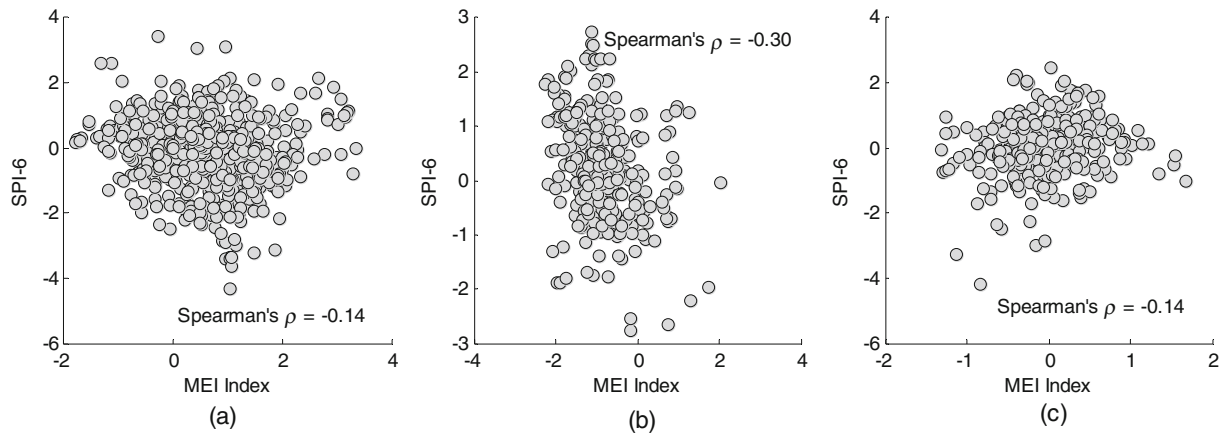
Figure 4. Scatter plot between SPI-6 and MEI time series at each ENSO phase during study period (1896–2005): (**a**) El Niño, (**b**) La Niña, and (**c**) neutral phases.

Table 1. *Drought characteristics without and with accounting ENSO climate phases.*

Climate state	Drought characteristics	Mean	Max	Standard deviation	Skewness coefficient
Without ENSO	Severity	4.1	30.7	5.9	2.8
	Duration	2.7	18	2.9	2.4
	Peak	1.4	4.3	0.7	2
El Niño	Severity	5.6	30.7	7.5	2.1
	Duration	3.6	18	3.7	1.8
	Peak	1.5	4.3	0.8	1.8
La Niña	Severity	2.3	9.5	2.2	1.8
	Duration	1.8	6	1.3	1.7
	Peak	1.3	2.8	0.6	1.3
Neutral	Severity	3.3	17.3	4.3	2.4
	Duration	2.2	10	2.2	2.4
	Peak	1.5	4.2	0.8	2.2

Table 2. *Strength of dependence between pair-wise drought characteristics without and with accounting ENSO climate phases.*

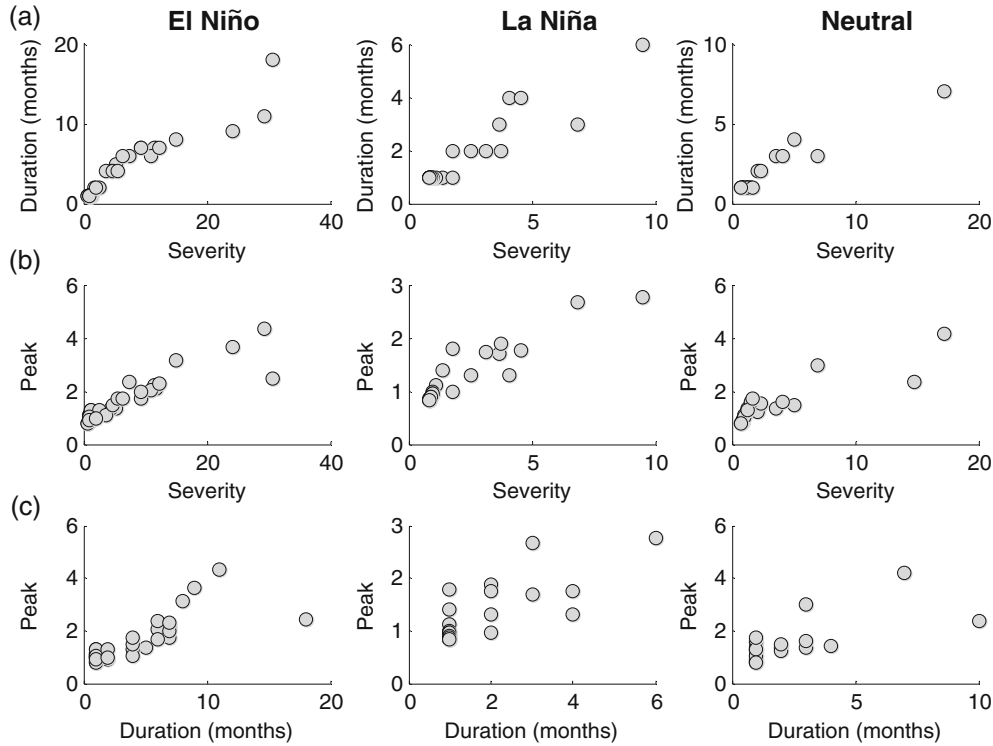| Climate state	Drought variables	Pearson's r	Kendall's τ	Spearman's ρ
Without ENSO	Severity-duration	0.95	0.79	0.90
	Duration-peak	0.74	0.61	0.76
	Severity-peak	0.85	0.85	0.94
El Niño	Severity-duration	0.95	0.84	0.93
	Duration-peak	0.81	0.70	0.83
	Severity-peak	0.91	0.87	0.95
La Niña	Severity-duration	0.93	0.75	0.87
	Duration-peak	0.76	0.59	0.71
	Severity-peak	0.92	0.87	0.95
Neutral	Severity-duration	0.94	0.76	0.87
	Duration-peak	0.69	0.51	0.63
	Severity-peak	0.87	0.80	0.90

Figure 5. Scatter plots of drought variables: (**a**) severity-duration (**b**) severity-peak, and (**c**) duration-peak pairs for three ENSO climate states – El Niño, La Niña and neutral phases.

Table 3. *Estimated parameters for marginal distribution of drought variables.*

Climate states	Drought characteristics	Probability distribution	Shape parameter	Scale parameter
Without ENSO	Severity	Weibull	0.88	3.8
	Duration	Exponential	–	2.74
	Peak	Log normal	0.41	0.27
El Niño	Severity	Weibull	0.83	4.97
	Duration	Exponential	–	3.58
	Peak	Log normal	0.43	0.29
La Niña	Severity	Log normal	0.78	0.52
	Duration	Exponential	–	1.83
	Peak	Log normal	0.39	0.19
Neutral	Severity	Log normal	0.89	0.69
	Duration	Exponential	–	2.17
	Peak	Log normal	0.41	0.31

4.3 *Marginal distribution of drought variables*

To fit marginal probability distributions for drought variables from each climatic phase, several parametric distributions are applied and their performance is evaluated using Akaike Information Criterion (AIC). It is found that the drought characteristics duration and peak are best represented by exponential and log-normal

distributions, respectively for all the cases (i.e., for droughts with and without ENSO phase considerations). Drought severity is best modelled by Weibull distribution for the case of without ENSO and El Niño phase data, log-normal distribution for La Niña and neutral phase data. Validity of each model is tested using Kolmogorov–Smirnov (KS) goodness-of-fit test. Table 3 lists the distributions and their associated parameters
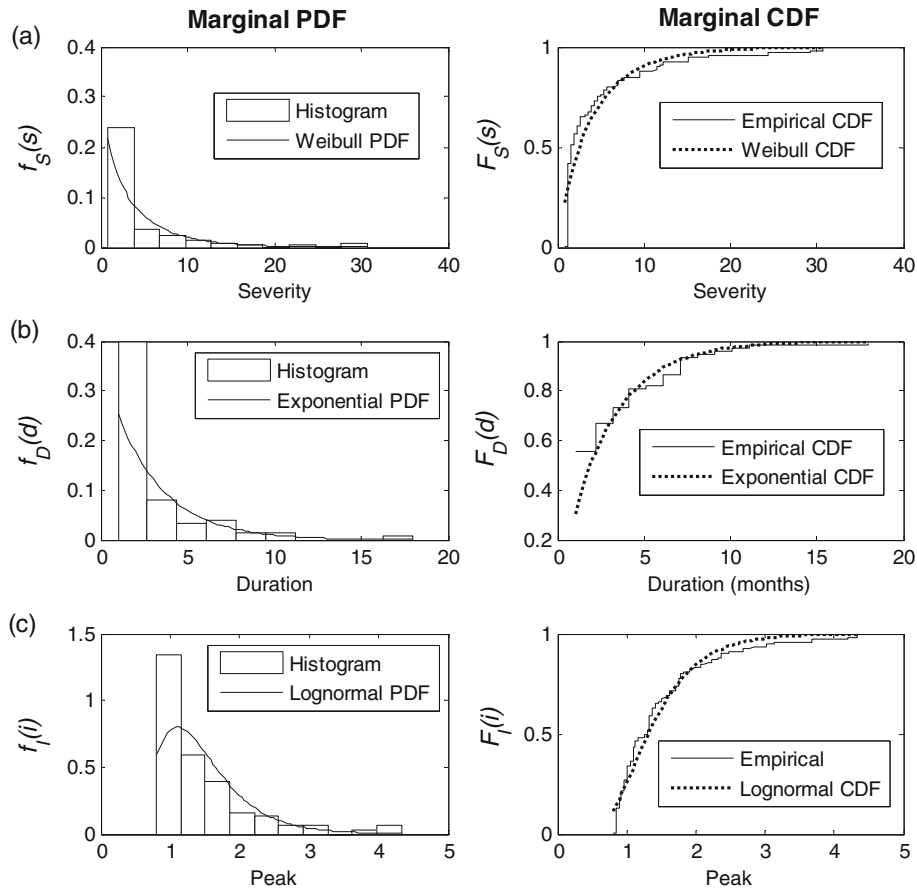
Figure 6. Marginal distribution fit of drought variables: (**a**) severity, (**b**) duration, and (**c**) peak without accounting ENSO phase information.

used for fitting marginal distributions of drought characteristics. Figure 6 presents marginal distribution fit of drought variables – severity, duration and peak without accounting ENSO phase information. Figures 7 and 8 present Probability Density Function (PDF) and CDF plots of marginal fit of drought variables under three ENSO states. The probability distribution plots of drought variables show good match between theoretical and empirical distributions.

4.4 *Dependence modelling using copulas*

Student's t copulas with 2 degrees of freedom are found to be adequate for fitting drought variables with and without accounting ENSO climate state information. The associated shape parameter matrix of Student's t copulas is listed in table 4. The root mean square error (RMSE) between parametric Student's t and empirical copulas for entire drought data (i.e., without considering ENSO stratification) is found to be 0.031; and the RMSE values for El

Niño, La Niña and neutral phase models are 0.035, 0.053 and 0.056, respectively. Figure 9 shows P–P plots of trivariate copulas fitted for joint distribution of drought characteristics considering with and without ENSO state information. The P–P plot gives the relation between empirical copula $C_n(\bullet)$ and parametric Student's t copulas $C_{\sum,\vartheta}$. From figure 9, it can be seen that the scatter plots between parametric and empirical copulas are close to the 45° line, which shows that a good correspondence exists between theoretical model and empirical distributions. To test the performance of the Student's t copula in modelling upper tail dependence, pair-wise TDC (for both parametric and non-parametric estimates) are computed and used for evaluation of copula model efficacy. The parametric upper TDC $\left(\hat{\lambda}_U^{\text{param}}\right)$ of Student's t copula is computed using equation (5) and the non-parametric upper TDC $\left(\hat{\lambda}_U^{\text{CFG}}\right)$ is computed using equation (6). The non-parametric TDC is computed both for observed samples (denoted by $\hat{\lambda}_U^{\text{obs}}$) and for copula-simulated data (denoted by $\hat{\lambda}_U^{\text{CFG}}$).
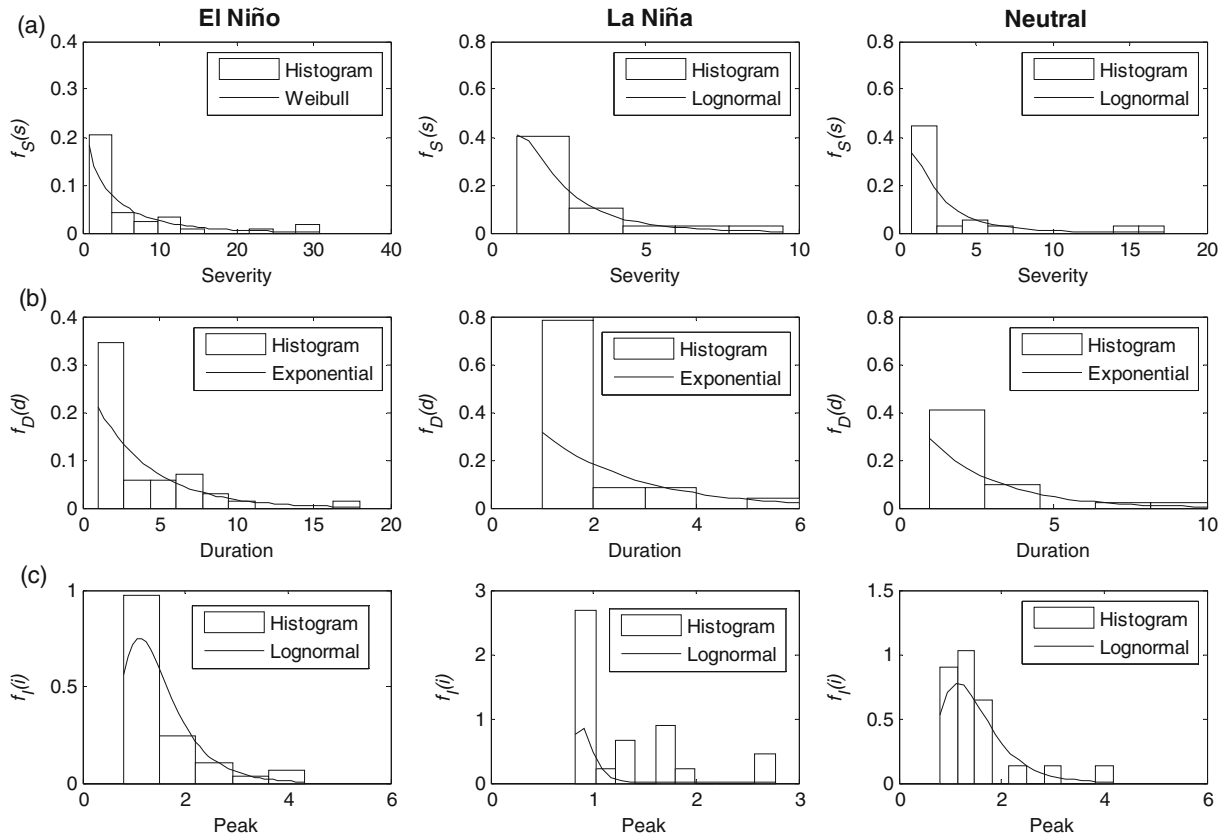
Figure 7. Marginal distribution fit of drought variables: (**a**) severity, (**b**) duration, and (**c**) peak, show PDF plots of fitted distributions for each phase of ENSO state.
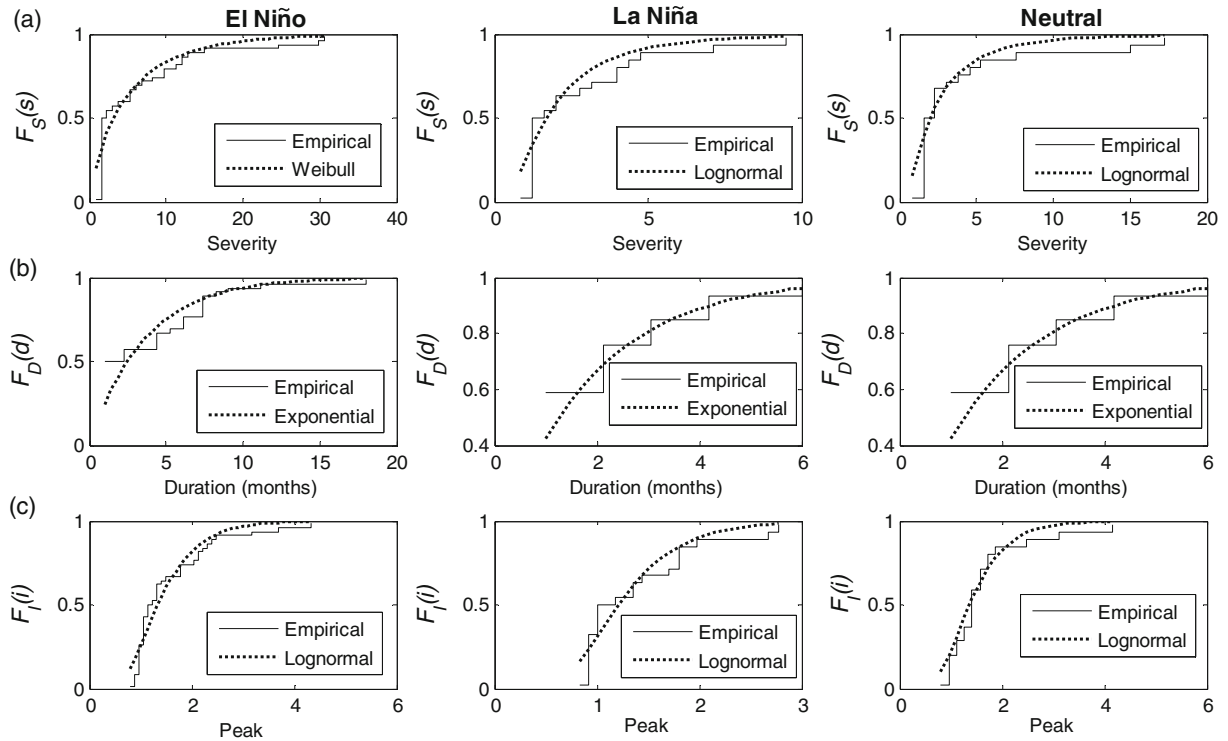

Figure 8. Marginal distribution fit of drought variables: (**a**) severity, (**b**) duration, and (**c**) peak, show CDF plots of fitted distributions for each phase of ENSO state.

Table 4. *The elements of shape parameter matrix* $(\hat{\sigma}_{i,j})$ *of Student's t copula for different climate states. In* $\hat{\sigma}_{i,j}$, *i and j takes values of* $\{1,2,3\}$, *which represents three drought characteristics severity, duration and peak, respectively.*

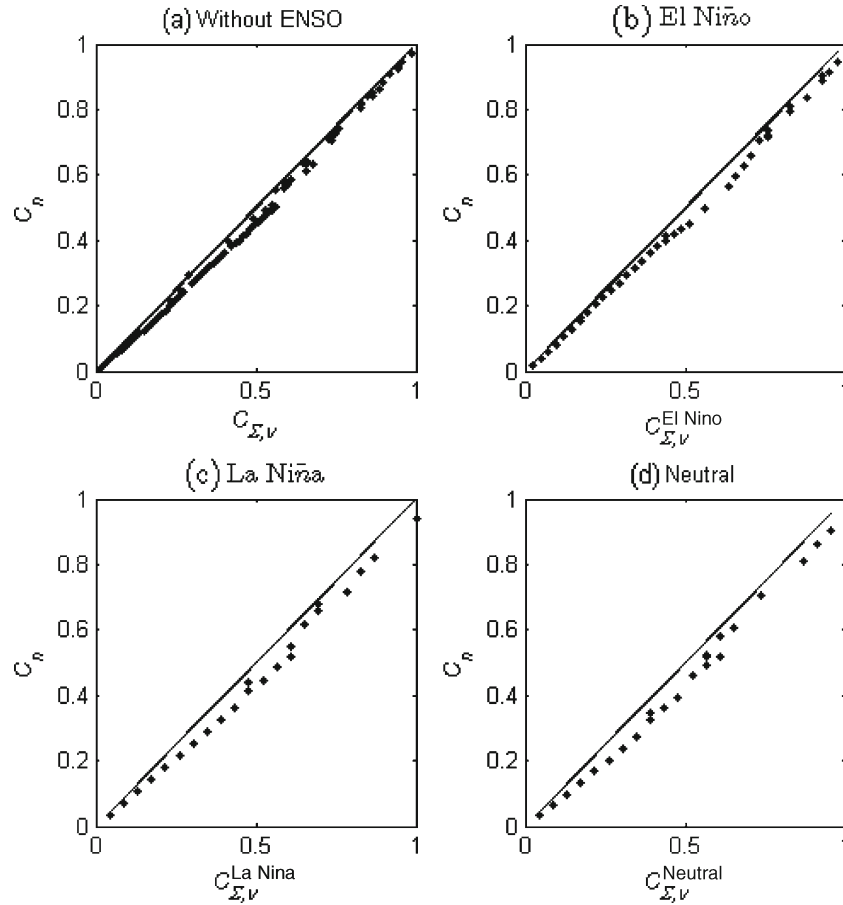| $i =$	$\hat{\sigma}_{i,j}$ (Without ENSO)			$\hat{\sigma}_{i,j}$ (El Niño)			$\hat{\sigma}_{i,j}$ (La Niña)			$\hat{\sigma}_{i,j}$ (Neutral)		
$j =$	1	2	3	1	2	3	1	2	3	1	2	3
1	1	0.94	0.96	1	0.97	0.98	1	0.92	0.97	1	0.92	0.94
2	0.94	1	0.82	0.97	1	0.89	0.92	1	0.80	0.92	1	0.72
3	0.96	0.82	1	0.98	0.89	1	0.97	0.80	1	0.94	0.72	1

Figure 9. The P–P plots of trivariate copulas, show the scatter plot between empirical copula (C_n) and Student's t copulas $(C_{\sum,\vartheta})$: (**a**) without ENSO-state, (**b**) at El Niño, (**c**) at La Niña, and (**d**) at neutral ENSO phase droughts.

The $\hat{\lambda}_U^{\text{obs}}$ is computed for data of observed samples, and the $\hat{\lambda}_U^{\text{CFG}}$ is computed for 1000 samples generated from copulas. The non-parametric TDC for copula models is repeated for hundred runs and the corresponding mean $\hat{\mu}\left(\hat{\lambda}_U^{\text{CFG}}\right)$ and standard deviation $\hat{\sigma}\left(\hat{\lambda}_U^{\text{CFG}}\right)$ values for various combinations of drought variables are presented in table 5. It can be seen from the table that Student's t copula is able to capture the observed upper tail dependence satisfactorily.

4.5 *Analysis of trivariate return periods conditional on ENSO-state*

The use of tele-connections between ENSO and regional droughts is of great importance for planning and management of water resources projects in the region, as ENSO is one of the major sources of atmospheric variability in global scale. Hence, this study uses the climate tele-connection between ENSO and SPI-6 for ENSO-conditioned meteorological drought-risk analysis for the Western Rajasthan region. The drought probabilities during

Table 5. *Upper tail dependence coefficient of Student's t copulas for pair-wise dependent drought variables at each ENSO-state drought condition.*

ENSO state	Drought variables	$\hat{\lambda}_U^{\text{param}}$	$\hat{\lambda}_U^{\text{Obs}}$	$\hat{\lambda}_U^{\text{CFG}}$	
				$\hat{\mu}\left(\hat{\lambda}_U^{\text{CFG}}\right)$	$\hat{\sigma}\left(\hat{\lambda}_U^{\text{CFG}}\right)$
Without ENSO	Severity-duration	0.78	0.82	0.84	0.006
	Duration-peak	0.62	0.65	0.69	0.012
	Severity-peak	0.83	0.84	0.87	0.005
El Niño	Severity-duration	0.83	0.86	0.87	0.005
	Duration-peak	0.70	0.74	0.77	0.009
	Severity-peak	0.87	0.87	0.90	0.004
La Niña	Severity-duration	0.74	0.79	0.80	0.008
	Duration-peak	0.60	0.65	0.67	0.014
	Severity-peak	0.85	0.89	0.89	0.005
Neutral	Severity-duration	0.74	0.78	0.80	0.008
	Duration-peak	0.54	0.59	0.60	0.016
	Severity-peak	0.77	0.83	0.82	0.007

Table 6. *Comparison of trivariate return periods (in years) of drought characteristics without and with accounting ENSO state conditions. The first column gives univariate return period (T) of drought characteristics.*

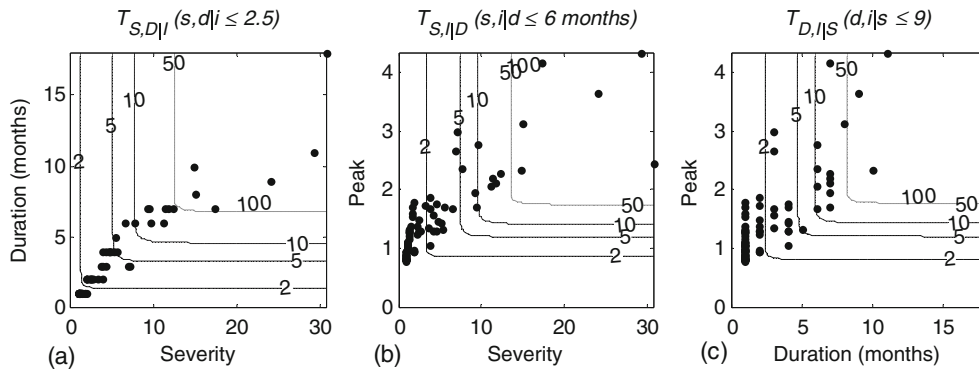|	Drought property quantile value at T			Without ENSO		With ENSO	
T	S	D (months)	I	$T_{\text{SDI}}^{\cup}$	$T_{\text{SDI}}^{\cap}$	$T_{\text{SDI}}^{\cup,\text{ Clim}}$	$T_{\text{SDI}}^{\cap,\text{ Clim}}$
2.5	2.48	1.88	1.31	2.10	3.13	2.10	3.24
5	5.47	3.78	1.73	3.89	7.30	3.88	7.94
10	8.69	5.69	2.11	7.46	16.31	7.12	17.15
20	12.05	7.59	2.47	14.67	35.05	12.88	34.41
25	13.15	8.21	2.59	18.35	45.27	15.53	42.63
30	14.07	8.71	2.68	21.94	55.07	18.07	50.66

Figure 10. Contour plots of conditional return periods (years) of droughts without accounting ENSO state information: (**a**) conditional return periods of drought severity and duration at given peak, $T_{S,\,D|I}$, (**b**) conditional return periods of drought severity and peak at given duration, $T_{S,\,I|D}$, and (**c**) conditional return periods of drought duration and peak at given severity, $T_{D,I|S}$. Historical droughts events are shown as black dots.

El Niño, La Niña and neutral years are found to be 0.47, 0.26 and 0.26, respectively. Table 6 presents univariate return periods and its associated drought quantiles, and joint return periods computed using OR as well as AND-case utilizing with and without ENSO state information. Due to limited data on La Niña and neutral years (each have 23 drought events), analysis is restricted to only shorter return periods. From table 6, it can be noticed that the return period in AND-case is always greater than that of the return period in OR-case. A small difference in return periods is observed on accounting without and with climatic state information, which shows that the developed models are consistent, and the uncertainty associated with models (and its associated parameters) are handled satisfactorily.

The joint return periods of drought variables conditional on third drought property – $T_{S,D|I}$, $T_{S,I|D}$ and $T_{D,I|S}$ are computed using equation (12) (without considering ENSO states) and the corresponding contour plots are presented in figure 10.

This figure also shows the superimposed historical (or observed) drought variables, which indicates that most of the historical drought events have shorter return periods. At the same time, there are few historical drought events, which have return periods of more than 100 years.

The conditional probability $P(S \leq s | D \geq d', I \geq i')$ of drought severity, given drought duration and peaks exceeding certain thresholds, d' and i' are computed using equation (13). These conditional probabilities are computed for each ENSO state, where d' takes values of duration at 50th, 75th and 90th percentile levels, and i' takes values of peak at 50th and 75th percentile levels and are shown in figure 11. From the plots in figure 11, it is possible to know the conditional non-exceedance probability of drought severity, given drought duration and peak exceeding certain values. For all three cases, the skewness of conditional probability curves shows increasing trends. These results can be helpful in assessing drought risks during
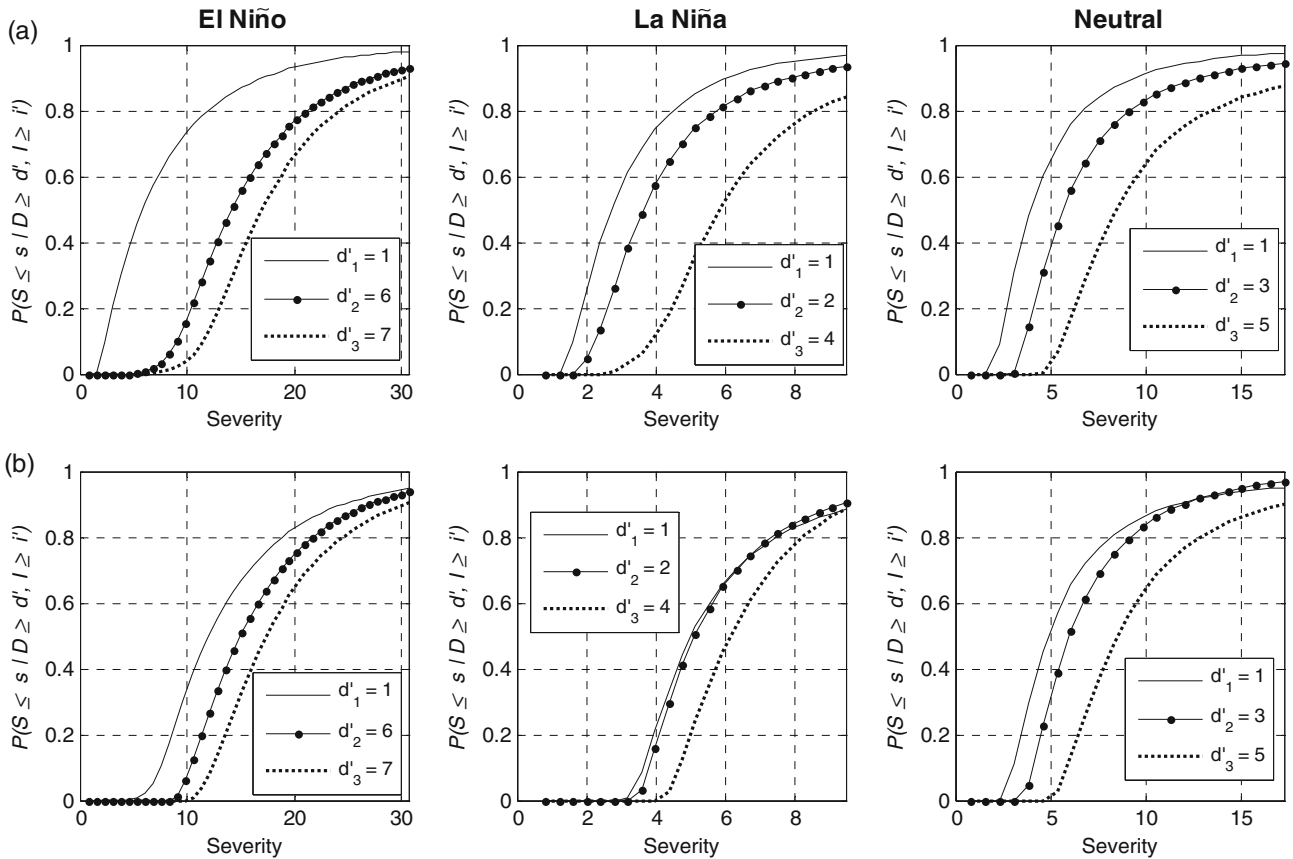

Figure 11. The conditional probabilities $P\left(S \leq s | D \geq d', I \geq i'\right)$ of drought severity, given drought duration and peak exceeding certain thresholds d' and i', respectively for each ENSO state (where d' takes values of duration at 75th, 90th and 95th percentile levels, and i' takes values of peak at 50th and 75th percentile levels). First row (**a**) shows the conditional probability curves for i' taking values of peak at 50th percentile level, and second row (**b**) shows the conditional probability curves for i' taking values of peak at 75th percentile level for El Niño, La Niña and neutral state droughts.
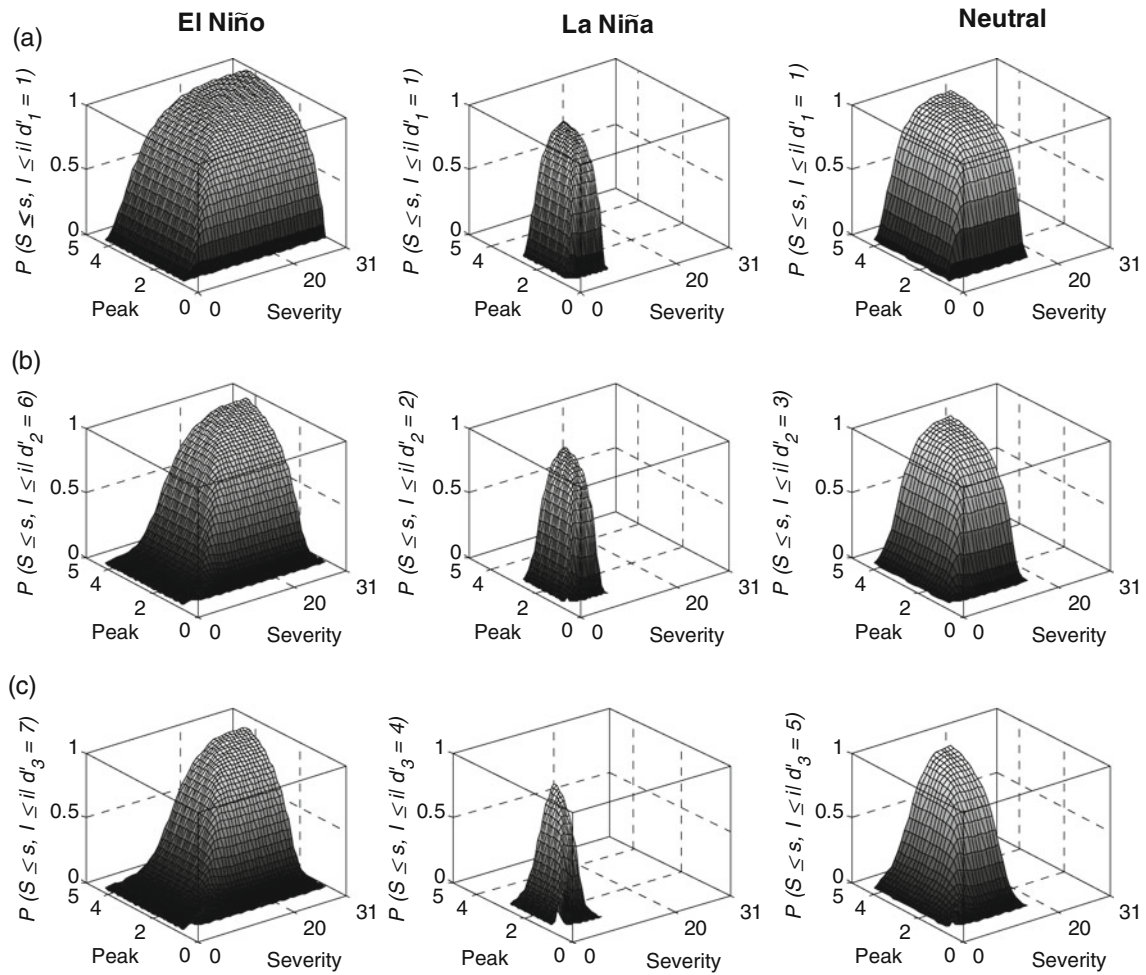
Figure 12. Conditional probabilities $P\left(S \leq s, I \leq i \,|\, D \geq d'\right)$ of drought severity and peak, given duration exceeding certain threshold d' for El Niño, La Niña and neutral state droughts. First row (**a**), second row (**b**), and third row (**c**) give the conditional probabilities at drought duration exceeding 50th, 75th and 90th percentile values. Here d'_1, d'_2 and d'_3 correspond to values of drought duration at 50th, 75th and 90th percentiles.

Table 7. *Conditional probabilities of drought characteristics for different ENSO state droughts. $P_{SI|D\geq d}$ denotes conditional probability of drought severity and peak for given duration exceeding certain threshold.*

| ENSO state | Percentile levels | Drought property quantiles at different percentiles | | | |
		S	D (months)	I	$P_{SI|D\geq d}$
El Niño	50th	1.31	1	1.11	0.073
	75th	7.95	6	1.78	0.080
	90th	13.29	7	2.39	0.251
	95th	26.42	10	3.35	0.702
La Niña	50th	1.11	1	0.98	0.109
	75th	3.54	2	1.74	0.697
	90th	5.02	4	2.05	0.760
	95th	7.76	5	2.70	0.907
Neutral	50th	1.55	1	1.34	0.215
	75th	3.26	3	1.58	0.344
	90th	8.53	5	2.48	0.763
	95th	15.64	8	3.40	0.883

different phases of ENSO. For example, for a given data of drought properties: severity $s = 7$, duration $d = 4$ months and peak $i = 1.45$, the corresponding conditional probability $P\left(S \leq s \mid D \geq d', I \geq i'\right)$ values estimated for El Niño, La Niña and neutral state droughts are 0.15, 0.66 and 0.44, respectively.

Similarly, the conditional non-exceedance probabilities $P\left(S \leq s, I \leq i \mid D \geq d'\right)$ of drought severity and peak for a given duration exceeding certain threshold at 50th, 75th and 90th percentiles levels are computed using equation (14) and the corresponding surface plots are presented in figure 12 for each ENSO state condition. Table 7 presents conditional probability $P\left(S \leq s, I \leq i \mid D \geq d'\right)$ of drought variables at different percentile values of drought properties for three ENSO state droughts. From table 7, it can be observed that the magnitude of drought properties are always higher for El Niño phase as compared to La Niña and neutral phases. Also, the conditional probability is increasing in nature for different percentile values of drought variables. Thus, the conditional probabilities of drought characteristics under different ENSO state conditions can be useful for analysing the associated drought risks. The copula-based trivariate modelling of drought characteristics is very useful in estimation of multivariate return periods and probabilistic assessment of drought risks, since the conventional univariate risk analysis of drought characteristics may lead to overestimation or underestimation of associated risks. Thus, by using the information on ENSO-based climate variability in modulating drought risks, it is possible to effectively plan agriculture and manage water resources in the region.

5. Conclusion

This study investigated the influence of ENSO-based climate variability on multivariate drought risks in drought-prone region of Western Rajasthan, India. Using multivariate ENSO index, the drought events are partitioned into three climatic states – El Niño, La Niña and neutral episodes. It is observed that during the study period (1896–2005), El Niño year droughts were more severe and longer than La Niña and neutral phase droughts. Different probability distributions are employed for fitting marginal distributions for drought variables and the best-fitted models are selected based on goodness-of-fit test and AIC criteria. To construct multivariate joint distribution of drought characteristics, the elliptical class of trivariate Student's t-copula family is adopted and used for deriving multivariate return periods

of drought properties in two cases, i.e., without and with accounting ENSO state information. Then, the conditional probabilities of drought characteristics are analysed under three ENSO states – El Niño, La Niña and neutral phases. It is found that the droughts during El Niño state are more severe than the other two states. The results of the study suggest that: (i) the conventional univariate risk analysis of drought characteristics may lead to overestimation or underestimation of associated risks, hence, the copula-based estimation of multivariate return periods can be used for effective risk assessment of droughts; (ii) the ENSO-based climate variability can be used for assessing drought risks and appropriate planning of agriculture and water management in the region.

References

Capéraá P, Fougéres A-L and Genest C 1997 A nonparametric estimation procedure for bivariate extreme value copulas; *Biometrika* **84(3)** 567–577.

Dai A 2011 Drought under global warming: A review; Wiley Interdisciplinary Reviews; *Clim. Change* **2(1)** 45–65.

Favre A-C, Adlouni S E l, Perreault L, Thiemonge N and Bobée B 2004 Multivariate hydrological frequency analysis using copulas; *Water Resour. Res.* **40** W01101, doi: 10.1029/2003WR002456.

Frahm G, Junker M and Schimdt R 2005 Estimating the tail-dependence coefficient: Properties and pitfalls; *Insur. Math. Econ.* **37** 80–100.

Gadgil S, Rajeevan M and Francis P A 2007 Monsoon variability: Links to major oscillations over the equatorial Pacific and Indian oceans; *Curr. Sci.* **93(2)** 182–194.

Genest C and Favre A-C 2007 Everything you always wanted to know about copula modeling but were afraid to ask; *J. Hydrol. Eng.* **12(4)** 347–368.

Grimaldi S and Serinaldi F 2006 Asymmetric copula in multivariate flood frequency analysis; *Adv. Water Resour.* **29(8)** 1155–1167.

Janga Reddy M and Ganguli P 2012 Application of copulas for derivation of drought severity–duration–frequency curves; *Hydrol. Process.* **26(11)** 1672–1685.

Kao S C and Govindaraju R S 2010 A copula-based joint deficit index for droughts; *J. Hydrol.* **380** 121–134.

Khera R 2004 *Drought proofing in Rajasthan: Imperatives, experience and prospects*; Discussion Paper Series-5, UNDP, India, 50p.

Kim T W, Valdes J B and Yoo C 2003 Nonparametric approach for estimating return periods of droughts in arid regions; *J. Hydrol Eng.* **8(5)** 237–246.

Laux P, Wagner S, Wagner A, Jacobeit J, Bárdossy A and Kunstmann H 2009 Modelling daily precipitation features in the Volta Basin of the West Africa; *Int. J. Climatol.* **29(7)** 937–954.

Lee T, Modarres R and Ouarda T B M J 2012 Data-based analysis of bivariate copula tail dependence for drought duration and severity; *Hydrol. Process.*, doi: 10.1002/hyp.9233.

Madadgar S and Moradkhani H 2011 Drought analysis under climate change using copula. *J. Hydrol. Eng.*, doi: 10.1061/(ASCE)HE.1943–5584.0000532.

Mall R K, Gupta A, Singh R, Singh R S and Rathore L S 2006 Water resources and climate change: An Indian perspective; *Curr. Sci.* **90** 1610–1626.

Mashal R and Zeevi A 2002 Beyond correlation: *Extreme co-movements between financial assets*; Technical Report, Columbia University, 48p.

McNeil A J, Frey R and Embrechts P 2005 *Quantitative risk management: Concepts, techniques and tools* (Princeton: Princeton University Press), 538p.

Mirakbari M, Ganji A and Fallah S R 2010 Regional bivariate frequency analysis of meteorological droughts; *J. Hydrol. Eng.* **15(12)** 985–1000.

Murthy C S and Sesha Sai M V R 2010 *Agricultural drought monitoring and assessment*; NRSC Report, National Remote Sensing Service Centre, ISRO, India, 29p.

Nelsen R B 2006 *An introduction to copulas*; Lecture Notes in Statistics, No. 139, Springer, New York, 216p.

RACP 2012 *Rajasthan Agriculture Competitiveness Project: Social Assessment and Management Framework*; Technical Report, Department of Agriculture, Government of Rajasthan, 218p.

Rajeevan M, Pai D S, Dikshit S K and Kelkar R R 2004 IMD's new operational models for long-range forecast of southwest monsoon rainfall over India and their verification for 2003; *Curr. Sci.* **86(3)** 422–431.

RAPCC 2012 *Rajasthan State Action Plan on Climate Change*; Technical Report, Government of Rajasthan, 218p.

Rathore M S 2004 *State level analysis of drought policies and impacts in Rajasthan, India*; Working Paper 93, Drought Series Paper No. 6, International Water Management Institute, 35p.

Serinaldi F, Bonaccorso B, Cancelliere A and Grimaldi S 2009 Probabilistic characterization of drought properties through copulas; *Phys. Chem. Earth* **34(10–12)** 596–605.

Shiau J T 2006 Fitting drought duration and severity with two-dimensional copulas; *Water Resour. Manag.* **20** 795–815.

Shiau J T, Feng S and Nadarajah S 2007 Assessment of hydrological droughts for the Yellow River China, using copulas; *Hydrol. Process.* **21(16)** 2157–2163.

Shiau J T and Modarres R 2009 Copula-based drought severity–duration–frequency analysis in Iran; *Meteor. Appl.* **16(4)** 481–489.

Sklar A 1959 *Fonctions de répartition á n dimensions et leurs marges*; Publications de l'Institut de Statistique de l'Université de Paris **8** 229–231.

Song S and Singh V P 2010 Meta-elliptical copulas for drought frequency analysis of periodic hydraulic data. *Stochast. Environ. Res. Risk Assess.* **24(3)** 425–444.

Svoboda M *et al.* 2002 The drought monitor; *Bull. Am. Meteorol. Soc.* **83** 1181–1190.

Webster P J and Yang S 1992 Monsoon and ENSO: Selectively interactive systems; *Quart. J. Roy. Meteorol. Soc.* **118** 877–926.

Willems P 2000 Compound intensity/duration/ frequency-relationships of extreme precipitation for two seasons and two storm types; *J. Hydrol.* **233** 189–205.

Wolter K and Timlin M S 2011 El Niño/Southern oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext); *Int. J. Climatol.* **31** 1074–1087.

Wong G, Lambert M F, Leonard M and Metcalfe A V 2010 Drought analysis using trivariate copulas conditional on climate states; *J. Hydrol. Eng.* **15(2)** 129–141.