Check for updates

A Non-stationary and Probabilistic Approach for Drought Characterization Using Trivariate and Pairwise Copula Construction (PCC) Model

Soumyashree Dixit[1] · K. V. Jayakumar[1]

Abstract

Under variable climatic conditions, the conventional Standardized Precipitation Index (SPI) and Reconnaissance Drought Index (RDI) are inadequate for predicting extreme drought characteristics. Non-stationary Standardized Precipitation Index (NSPI) and Non-stationary Reconnaissance Drought Index (NRDI) are, therefore, developed by fitting non-stationary distributions. The Generalized Additive Model in Location, Scale and Shape (GAMLSS) framework, with time varying location parameters considering the external covariates, is used to fit the non-stationary distributions. Multivariate ENSO Index (MEI), Southern Oscillation Index (SOI), Sea Surface Temperature (SST), and Indian Ocean Dipole (IOD) are considered as external covariates for the non-stationary drought assessment. The performances of stationary and non-stationary models are compared. The study also concentrated on the trivariate and the Pairwise Copula Construction (PCC) models to estimate the drought return periods. The comparison of two copula models revealed that the PCC model performed better than the trivariate Student's t copula model. The recurrence intervals arrived at for the drought events are different for trivariate copula model and PCC model. This study showed that non-stationary drought indices will be helpful in the accurate estimate of the drought characteristics under the changing climatic scenario.

Keywords NSPI · NRDI · MEI · IOD · SST · SOI · GAMLSS · PCC

1 Introduction

Drought is defined as the abnormally prolonged dry weather condition that results in hydrological imbalance for a particular region. Irrespective of the deficit in the availability of water, it is one of the challenging issues which human society is confronting. The drought monitoring for a particular area can be handled through various indices. Many authors have developed univariate drought indices, viz., Standardized Precipitation Index, SPI (McKee et al. 1993), Standardized Runoff Index, SRI (Shukla and Wood 2008), and

---

✉ Soumyashree Dixit
dixitsoumya27@gmail.com

K. V. Jayakumar
kvj@nitw.ac.in

[1]   Present Address: Department of Civil Engineering, National Institute of Technology, Warangal, India

Standardized Soil moisture Index, SSI (Hao and AghaKouchak 2013). Various studies have also concentrated on the multivariate drought indices such as the Joint Deficit Index, JDI, (Kao and Govindaraju 2010), Standardized Precipitation Evapotranspiration Index, SPEI, (Vicente-Serrano et al. 2010); Integrated Drought Index, IDI (Shah and Mishra 2020b); Copula-based Joint Drought Index, CJDI (Won et al. 2020) and Reconnaissance Trivariate Drought Index, RTDI (Dixit and Jayakumar 2021a). Indian agronomy and cropping pattern are dependent on the onset and withdrawal of monsoon, distribution of rainfall and water storage capacities. Moreover, drought during the year 2014–15 caused severe inadequacy of water by affecting around 330 million people (Mishra et al. 2016). A study by Poonia et al. (2021a) indicated that drought vulnerability will increase throughout India. Hence, adequate attention must be paid to understand the multi-scalar drought assessment, especially an agrarian country like India.

Stationarity refers to the parameters of the climate that are invariant and free of trends (Wang et al. 2015). Nevertheless, the outcome of environmental changes exhibits nonstationary behaviour in climatic parameters and hence, indicators have been developed for identifying drought status in non-stationary conditions (NSPI, Russo et al. 2013; NRDI, Bazrafshan and Hejabi 2018). The large-scale climate indices (external covariates) based on El Niño Southern Oscillation (ENSO) should not, however, be neglected in the computation of nonstationary indices (Bazrafshan and Hejabi 2018). Some of the studies proved the linkage between the global climate indices and Indian climatic conditions (Tamaddun et al. 2019; Shah and Mishra 2020a; Kumar et al. 2021a, b). It can, hence, be concluded that covariates provide greater insight into the climatic factors which influence the climatic parameters over time in a greater dimension.

Drought characteristics must have a dependence structure and hence a multivariate analysis of drought is widely accepted in hydro-meteorological studies for drought risk assessment. Copula function could preserve the multivariate dependence structure of drought characteristics (Saghafian and Mehdikhani; 2014; Mortuza et al. 2019; Poonia et al. 2021b; Dixit et al. 2021). Nevertheless, the higher dimensional multivariate Student's t copulas are not effective enough to model the complex dependence structure of extreme event variables (Aas et al. 2009). Due of the limitation of the vine or pair-copulas, an efficient way of copula construction method to model the complex dependence pattern emerged (Aas et al. 2009). Many studies have been conducted in the past to define the vital role of pair copula model in hydrological analysis (Song and Kang 2011; Gräler et al. 2013; Xu et al. 2022). Daneshkhah et al. (2016) developed a multivariate pair copula model using the flood properties in River Beas of the Himalayan region indicating that the Himalayan rivers are highly affected by the monsoon fluctuations and stored snow cover. Therefore, an extensive study based on drought characteristics must be carried out considering the multivariate dependent pattern of the drought variables.

Many studies have been carried out in the past for the non-stationary drought frequency analysis (Wang et al. 2015; Rashid and Beecham 2019) and the non-stationary flood frequency analysis (Villarini et al. 2009; Lopez and Frances 2013; Debele et al. 2017; Chen et al. 2021) using GAMLSS model. To the best of the knowledge, study focused on non-stationary drought analysis for the Godavari River Basin (GRB) have not been carried out in the past. GRB is the second largest river basin in India which spreads from Western Ghats to the Eastern coastal region, having a wide variation in climatic pattern. Most of the areas in the river basin have severe impacts on the economical, agricultural and social factors due to drought events (Dixit and Jayakumar 2021b). Hence, the nonstationary assessment of drought considering large-scale climate indices is an appropriate approach for this basin. The objectives of this study are: (i) use of GAMLSS

package in R developed by Rigby and Stasinopoulos (2005) for the non-stationary modelling of SPI and RDI in the Godavari River basin; (ii) construction of dependence among drought characteristics is studied by using the trivariate Student's t copula and PCC models; and (iii) the return period analysis for a better understanding of the statistical behaviour of drought events. The awareness regarding the extreme events and their response in river basin scale with the changing climate will be useful in planning, managing and combating the impacts on water resources systems.

2 Study Area and Data

Godavari is the second largest river basin in India and flows through the states of Maharashtra, Telangana, Andhra Pradesh, Odisha, Chhattisgarh and some parts of Karnataka, Madhya Pradesh. The total area of the Godavari basin is approximately 3, 12,810 sq. km., which is about 10% of the total area of the country. The length and width are about 995 km and 585 km respectively. The basin has a tropical climate and as a whole receives on an average 85% of the annual rainfall during the South-west monsoon. The annual precipitation in the river basin varies from 600 to 1200 mm. Most part of the river basin is rain-fed and the distribution of rainfall is highly irregular. The average temperature in the basin varies from 15° C in the winter to 45° C in the peak of summer. High variability in temperature and precipitation in the river basin affects the evapotranspiration and relative humidity of the basin. Therefore, it is essential to estimate the non-stationary drought condition related to precipitation and evapotranspiration variability. In this study, two sub-basins of GRB namely Upper Godavari River Basin (UGRB) and Lower Godavari River Basin (LGRB) are selected for non-stationary assessment of drought. The location map of the study area is given in Fig. 1.

In the present study, the Indian Meteorology Department (IMD) daily precipitation datasets of resolution 0.25×0.25 for the period of 1950 to 2017 are obtained from the IMD website. The Non-stationary Reconnaissance Drought Index (NRDI) requires both precipitation and evapotranspiration data to estimate the meteorological drought. Hence, the monthly 0.5×0.5 resolution evapotranspiration data has been downloaded from the
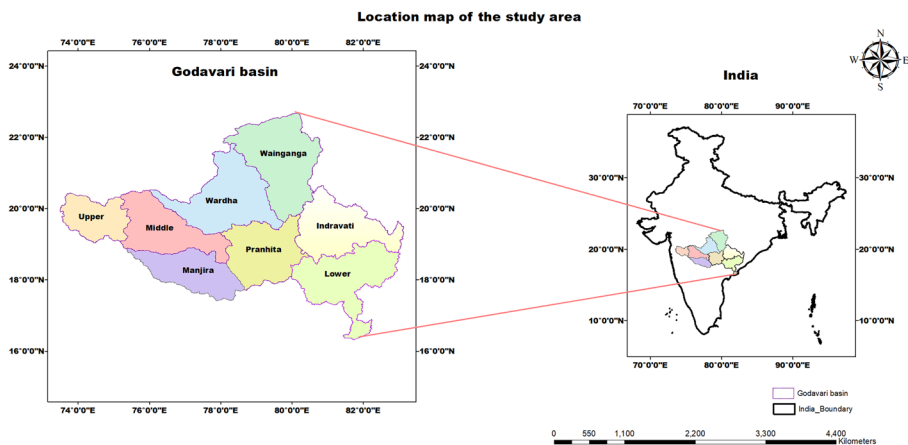

Fig. 1 Location map of the study area

Climate Research Unit (CRU) 4.03TS data sites for the same period. Recently, many studies have been conducted considering the CRU 4.03TS dataset (Zarch et al. 2015; Harris et al. 2020). Then the data is extracted and regridded to the IMD grids. There is an effect of large-scale climate indices on a non-stationarity evaluation of drought phenomenon. So, the non-stationary drought indices were computed based on the association with four large-scale climate indices such as Indian Ocean Dipole (IOD), Southern Oscillation Index (SOI), Sea Surface Temperature (SST) and Multivariate ENSO Index (MEI) and are presented in Table 1.

3 Methodology

3.1 Non-stationary Reconnaissance Drought Index (NRDI) and Non-stationary Standardized Precipitation Index (NSPI) Development

The stationary RDI is calculated based on the assumption that the parameters related to initial values (δ_0) are constant with time (Tsakiris et al. 2007). Under non-stationarity condition, some parameters of the distribution function of δ_0, the initial value, can get changed. δ_0 is taken as the ratio of precipitation series $(P_j(t))$ to evapotranspiration series $(PET_j(t))$ in different time windows (3-, 6- and 12-month), computed using the Eq. (1).

$$\delta_0(t) = \frac{\sum_{j=1}^{n} P_j(t)}{\sum_{j=1}^{n} PET_j(t)} \tag{1}$$

For the computation of NSPI and NRDI, correlation analysis between the precipitation and large-scale climate indices and the computed δ_0 values and the large-scale climate indices of different lags (1–12) are performed using Kendall's method (Kendall 1955) at a significance level of 0.05. The significant covariates from potential large-scale climate indices are evaluated from the best lags based on the minimum p-value. NSPI and NRDI can be developed using the filtered covariates by fitting the GAMLSS model.

3.2 GAMLSS Model Development

In the present study, the random variable y_i (precipitation series and the initial value series) was assumed to have a parametric cumulative distribution function. The related time varying parameters can be modelled as a function of selected covariates by using the GAMLSS model. The selected covariates based on the Kendall's τ lag correlation method for different time windows of precipitation and the series of initial values (δ_0) were varied linearly

Table 1 Large-scale climate indices and their data site

Large-scale climate indices	Downloading links
SOI	http://www.bom.gov.au/climate/current/%20soihtm1.shtml
SST	http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/nino34.long.anom.data
MEI	https://psl.noaa.gov/enso/mei.ext/table.ext.html
IOD	http://www.jamstec.go.jp/frcgc/research/d1/iod/DATA/dmi.monthly.txt

with the parameters (μ_t) is a time varying parameter) of the distribution. The step by step details on the use of GAMLSS analysis is available in Rigby and Stasinopoulos (2005).

In the present work, for computing the NSPI for a particular location, a two parametric Gamma distribution was fitted to the precipitation series with linearly varying location and constant scale parameter considering the explanatory variables given in Eq. (4), using Eqs. (2) and (3).

$$y_t \sim \text{Gamma}(\mu_t, \sigma) \tag{2}$$

$$\mu_t = a_o + a_1 I_1(t) + a_2 I_2(t) + \cdots + a_n I_n(t) \tag{3}$$

where $a_o, \ldots \ldots .., a_n$ are estimated mean coefficients for the linear variability for a particular location for fitted Gamma distribution and $I_1, \ldots \ldots, I_n$ represent the explanatory climate variables at time t.

$$f(y_t, \mu_t, \sigma) = \frac{1}{(\sigma^2 \mu_t)^{1/\sigma^2}} \frac{y^{\frac{1}{\sigma^2}-1} e^{-y/(\sigma^2 \mu_t)}}{(1/\sigma^2)} \tag{4}$$

The cumulative distribution functions of the aggregated rainfall series were computed by fitting the non-stationary model and transformed into standard normal values using Eq. (5).

$$\text{NSPI} = \varphi^{-1}(f(y_t, \mu_t, \sigma)) \tag{5}$$

where $f(y_t, \mu_t, \sigma)$ are the CDFs of the aggregated precipitation series, y_t is the aggregated precipitation at any time t, and φ^{-1} is the inverse CDF values.

Further, non-stationary Log-normal distribution with linearly varying location parameter (μ_t) with time considering the respective covariates and with the constant scale parameter (σ), was fitted to the δ_0 values as shown in Eqs. (6) and (7).

$$y_t \sim \text{LogNormal}(\mu_t, \sigma) \tag{6}$$

$$\mu_t = b_o + b_1 I_1(t) + b_2 I_2(t) + \cdots + b_n I_n(t) \tag{7}$$

where $b_o, \ldots, b_n$ are estimated mean coefficients for the linear variability for a particular location for the fitted non-stationary Log-normal distribution and $I_1, \ldots, I_n$ are the explanatory climate variables at time t. Then, using the time variant location parameter and the invariant scale parameter, the NRDI is estimated using the Eq. (8).

$$\text{NRDI}(t) = \frac{y_t - \mu_t}{\sigma} \tag{8}$$

Here, $y_t = \ln(\delta_0)$; μ_t is the time varying arithmetic mean and σ is the standard deviation of the observational variable. The parameters for Log-normal and Gamma distribution are estimated by Rigby and Stasinopoulos (RS) algorithm in the GAMLSS framework. The selection of both the models was evaluated in terms of lowest Akaike information criterion (AIC) values.

The classification of NSPI and NRDI is similar to the standard SPI (Mckee et al. 1993) and RDI (Tsakiris et al. 2007). The drought characteristics can be estimated using the widely

accepted method Run theory analysis (Ganguli and Reddy 2012; Dixit et al. 2021). A run is defined as the values below a preferred truncation level by considering the positive and negative run (Yevjevich 1967). In the present study, a value of -0.8 is considered as the threshold value below which all the values are taken to represent drought events. Duration (D), peak (P) and severity (S) are the three drought characteristics considered in the study.

3.3 Multivariate Dependence Modelling Using Copula Constructions

The multivariate drought frequency analysis based on copula theory can be used to overcome the limitations of univariate drought frequency analysis. A trivariate distribution function, fitted to three dependent drought characteristics P, D and S with marginal CDFs F_P, F_D and F_S can be represented based on copula function that guarantees the existence of a unique function c such that, P, D and S $\in$ R. The trivariate joint probability distribution resulting in the parameter associated with the specified copula function is presented by Eq. (9).

$$f(P, D, S) = f_P(P) \times f_D(D) \times f_S(S) \times c\big[F_P(P), F_D(D), F_S(S)\big] \tag{9}$$

Different types of copula families like Frank, Gaussian, Gumbel, Clayton, and Student's t copula are used in many hydro-meteorological studies. The dependency status between the interrelated drought variables is represented by the respective copula parameter. In this study, the parameters of copula families are estimated using the Maximum Likelihood Estimation (MLE) method based on its fitted marginal distribution. The best fitted copula was estimated based on the Goodness of Fit (GoF) measures such as Kolmogorov–Smirnov test (KS), Cramer-von Misses (CVM), and Chi-square (Chsq) tests (Genest and Favre 2007), AIC criteria and the maximum likelihood function. The copula parameters and their constructions carried out in this study are presented briefly (in Eqs. (10) to (14)). For more details regarding copula, Nelsen (2006) may be referred.

$$\text{Clayton} : C_\theta(\alpha, \beta) = \left(\alpha^{-\theta} + \beta^{-\theta} - 1\right)^{-\theta} \text{for all } \theta(0, \infty) \tag{10}$$

$$\text{Frank} : C_\theta(\alpha, \beta) = -\frac{1}{\theta}\log\left(1 + \frac{(\exp(-\theta\alpha) - 1)(\exp(-\theta\beta) - 1)}{\exp(-\theta) - 1}\right) \text{for } \theta \in R\{0\} \tag{11}$$

$$\text{Gaussian} : C_\theta(\alpha, \beta) = \int\limits_{-\infty}^{\varphi^{-1(\alpha)}} \int\limits_{-\infty}^{\varphi^{-1(\beta)}} \frac{1}{2\pi\sqrt{1-\theta^2}} e^{\frac{2\theta uv - u^2 - v^2}{2(1-\theta^2)}} \, dudv \tag{12}$$

$$\text{Gumbel} : C_\theta(\alpha, \beta) = \exp\left\{-[(-\log\alpha)^\theta + (-\log\beta)^\theta]^{1/\theta}\right\}, \text{ for } \theta \in [1, \infty) \tag{13}$$

$$\text{Student} - t : C_\theta(\alpha, \beta) = \int\limits_{-\infty}^{T_\gamma^{-1(\alpha)}} \int\limits_{-\infty}^{T_\gamma^{-1(\beta)}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\nu\pi\sqrt{1-\theta^2}} \left\{1 + \frac{u^2 + v^2 - 2\theta uv}{\nu(1-\theta^2)}\right\} \, dudv \tag{14}$$

Daneshkhah et al. ([2016]) and Brechmann and Schepsmeier ([2013]) showed the complex structure of the dependencies between multivariate data. For example, the parameters of Student's t copula showed a single degree of freedom which is the driving force for the dependency of all other pair of variables. Because of this limitation, a unique method of copula construction named as vine copula is introduced (Aas et al. [2009]). Regular vines are broadly categorised into two subsets i.e. D-vine and Canonical vine (Kurowicka and Cooke [2006]). Vine is a flexible graphically represented tree-like structure that computes the pairwise construction of variables that are mutually dependent, known as PCC model.

The D-vine structure that is used to model joint probability related to the drought characteristics P, D, and S with the given marginal densities F_P, F_D, F_S respectively. Here D-vine structure is selected for the joint density decomposition as shown in Eq. ([15]).

$$F(P, D, S) = f_P(P) \times f_D(D) \times f_S(S) c_{PD}\{F_P(P), F_D(D)\} \\ \times c_{DS}\{F_D(D), F_S(S)\} \times c_{PS|D}\{F_{PD}(P|D), F_{SD}(S|D)\} \tag{15}$$

$c_{PD}\{F_P(P), F_D(D)\}$ represents the bivariate copula fitted between $F_P(P)$ and $F_D(D)$. $c_{PS|D}\{F_{PD}(P|D), F_{SD}(S|D)\}$ represents the bivariate copula fitted to the second tree variables $F_{PD}(P|D), F_{SD}(S|D)$.

3.4 Estimation Methods for Pair-Copula Models

The dependency measurement between drought characteristics like P–D, D – S and S–P are estimated using Kendall's τ and Spearman's ρ and presented in supplementary Table S7. Then, an appropriate D-vine model was chosen in terms of the dependency of variables. Graphical tools like Kendall plot (K-plot) and Chi-plots are useful for the optimum choice of bivariate copula models (Genest and Favre [2007]). The GoF tests such as Vuong and Clarke tests, were also implemented to find the suitable copula family for this study (Vuong [1989]). Commonly used AIC criteria to distinguish between models were also used to find an optimum solution regarding the selection of copula family.

After deciding the pairwise copula models, the parameters were estimated based on the MLE method. For more details regarding the vine copula and process of parameter estimation for the distribution function, Aas et al. ([2009]) can be referred. Drought frequency analysis was carried out to identify the changes in the return period with respect to the trivariate copula model. The return period can be obtained by selecting the appropriate copula family based on the dependent parameters.

4 Results and Discussions

4.1 Non-stationary SPI and RDI Indices

The basis for this study to incorporate large-scale climate indices was that the global climatology is highly affected by the climatic phenomenon over the Pacific and Indian oceans (Rashid and Beecham [2019]). The large-scale climate indices are considered as external covariates for evaluating the NSPI and NRDI. For the LGRB and URGB,

Table 2 Significant lag for different large-scale climate oscillations

NRDI	SOI	MEI	IOD	SST	NSPI	SOI	MEI	IOD	SST
UGRB					**UGRB**				
3-month	2		0		3-month	2		5	
6-month	2				6-month	2			
12-month	0	0		0	12-month	0	0		0
LGRB					**LGRB**				
3-month	4				3-month	4			
6-month	3			2	6-month	3	2		
12-month	3	4		3	12-month	3	3		5

the monthly aggregation (3-, 6- and 12- month moving average) of precipitation and evapotranspiration data were prepared. In addition, the moving average of climate indices were obtained and then organised into 13 different lag values from 0 to 12. MLE method is used for the evaluation of parameters of Gamma and Log-normal distributions. Comparison between the stationary and non-stationary models was carried out by estimating the two stationary based indices namely SSPI and SRDI by keeping the parameters invariant with time.

Kendall correlation analysis was carried out at a significance level of 0.05 for the assessment of significant large-scale climate indices based on different lags for different time scales. Table 2 displays the dominant covariates for parameters of fitted distributions. It was observed that the 3-month cumulative precipitation and initial values (δ_0) showed a significant correlation with SOI and IOD for the UGRB and SOI for the LGRB for different lags respectively. Similarly, SOI was identified as the most influential covariate for 6-month aggregated precipitation and δ_0 series for the UGRB whereas SOI and SST showed a quantitative influence in the LGRB for different lag values. The MEI, SOI and SST are significantly associated with 12-month cumulative precipitation and δ_0 for both the sub-basins.

The stationary and non-stationary models are optimised by minimizing the AIC. Table 3 represents the AIC values for non-stationary and stationary models. The AIC values obtained from non-stationary models are consistently lesser than those obtained from the stationary models. Hence, it can be concluded that the non-stationary models achieved better performance than the stationary models.

Table 3 Comparison between stationary and non-stationary models using AIC values

Models	NRDI	SRDI	NSPI	SSPI
	UGRB		**UGRB**	
3-month	5272	5272	7119.57	7119.145
6-month	5615	5617	7591.586	7592.802
12-month	3759	3828	6175.181	6231.682
	LGRB		**LGRB**	
3-month	6923	6928	8935.202	8937.593
6-month	7268	7278	9266.781	9266.588
12-month	5061	5115	7343.492	7380.212

4.2 Comparison of Historical Drought Characteristics

The outputs in the form of rationales are the drought indices that are transformed from the cumulative probability of rainfall and initial values (δ_0). There are many differences of signals between the stationary and non-stationary time series, which are due to the inclusion of the external covariates. Large dissimilarities of signals were observed between NSPI and SSPI for 3-, 6- and 12-month aggregation levels for both the sub-basins as seen in Figs. 2 and 3. These dissimilarities are apparent in case of drought characteristics SPI and RDI and are fairly different from the NSPI and NRDI. For comparing stationary and non-stationary models, the box plots for 3-, 6-and 12-month time windows are shown in Figs. 4 and 5 for UGRB and LGRB to identify the variations of the drought characteristics. In these figures, P1, D1 and S1 refer to non-stationary drought characteristics while P2, D2 and S2 refer to stationary characteristics. After comparing 3-month NSPI and SSPI, significant differences are identified in the drought characteristics in the LGRB (Fig. 4b) compared to the UGRB (Fig. 4a). Larger differences were evident in the drought severity and peak between NSPI and SSPI for the LGRB for 3-month time scale. In the case of duration, the UGRB showed significant variations between NSPI and SSPI for 3-month time scale whereas the LGRB showed less variations. Figure 5 shows the comparison between NRDI and SRDI for the spatial and temporal variability of drought characteristics for the 3-month time scale. The estimated severities of SRDI series were lower than the severity from NRDI series in the case of UGRB whereas the estimated peaks of SRDI series were higher than those estimated from the NRDI series (Fig. 5a). In case of 3-month time series, the LGRB, the peak and severity estimated from SRDI were higher than that of the NRDI (Fig. 5b). The comparison between NRDI and SRDI showed some variations in drought duration. From Fig. 4c, it is observed that the UGRB displayed significant changes in peak, duration and severity of the 6-month NSPI and SSPI series. For LGRB, there were significant variations of peak and severity whereas lesser variations of NSPI and SSPI were detected for the drought duration (Fig. 4d). From Fig. 5c and d, no significant variations of the drought characteristics were observed for NRDI and SRDI for both the sub-basins for the 6-month time scale. Comparison of the non-stationary approach and stationary for the 12-month time scale for the drought events in both sub-basins showed significant deviations in drought characteristics (Fig. 5e and f). It can, hence, be concluded that substantial variations of drought characteristics were evident in the case of 12-month time window comparing to 3- and 6-month time window.

The differences observed in the drought characteristics between the stationary and non-stationary models have significant role in the implementation of sustainable water resources systems planning and management. In this study, non-stationary models have been considered for further analysis using trivariate and pairwise copula.

4.3 Trivariate Copula Models

The Archimedean copulas (Gumbel, Frank and Clayton) and elliptical copulas (Gaussian and Student's t) are used in this study to evaluate the dependence structure of drought characteristics. Different types of distributions namely Gumbel, Gamma, Log-normal, Weibull and Exponential distributions were fitted to the drought characteristics. The best fitted distributions for several time scales were obtained based on AIC criteria and the log-likelihood values to find the marginal distributions are provided in Supplementary Tables S1-S6.
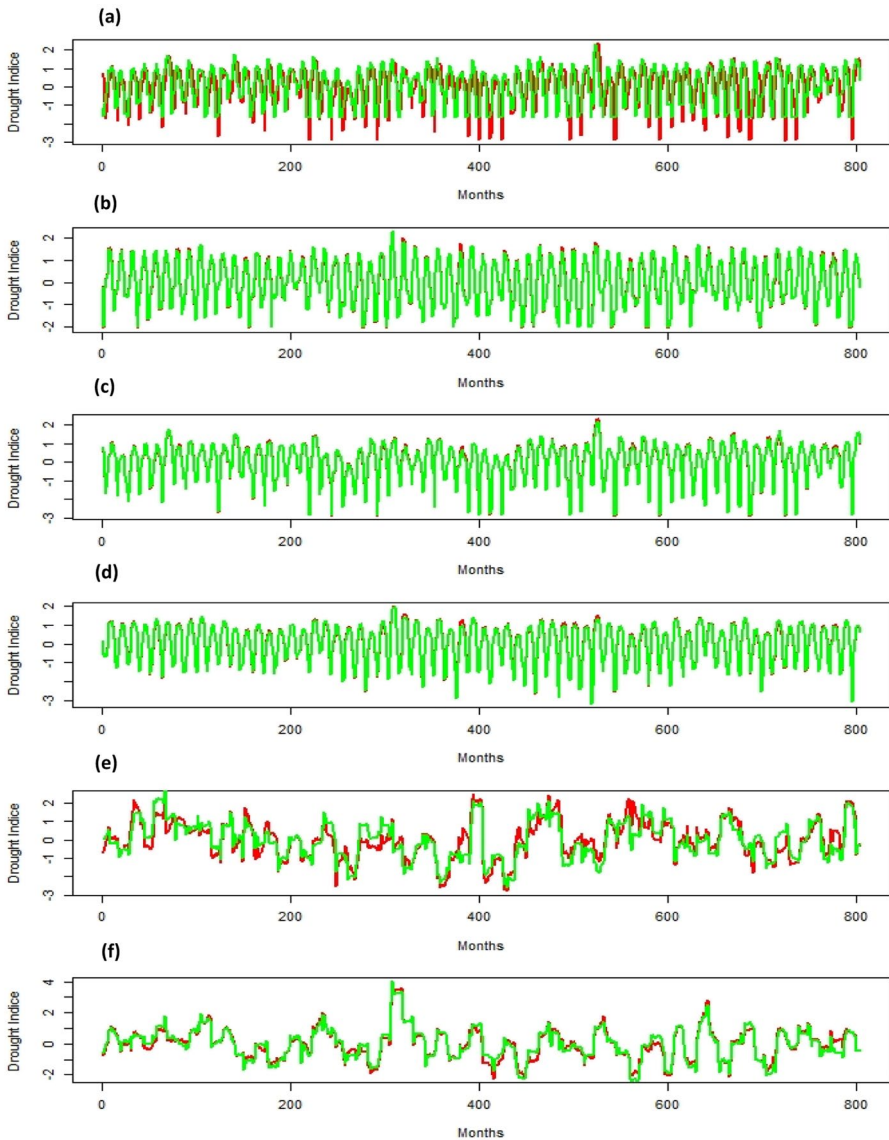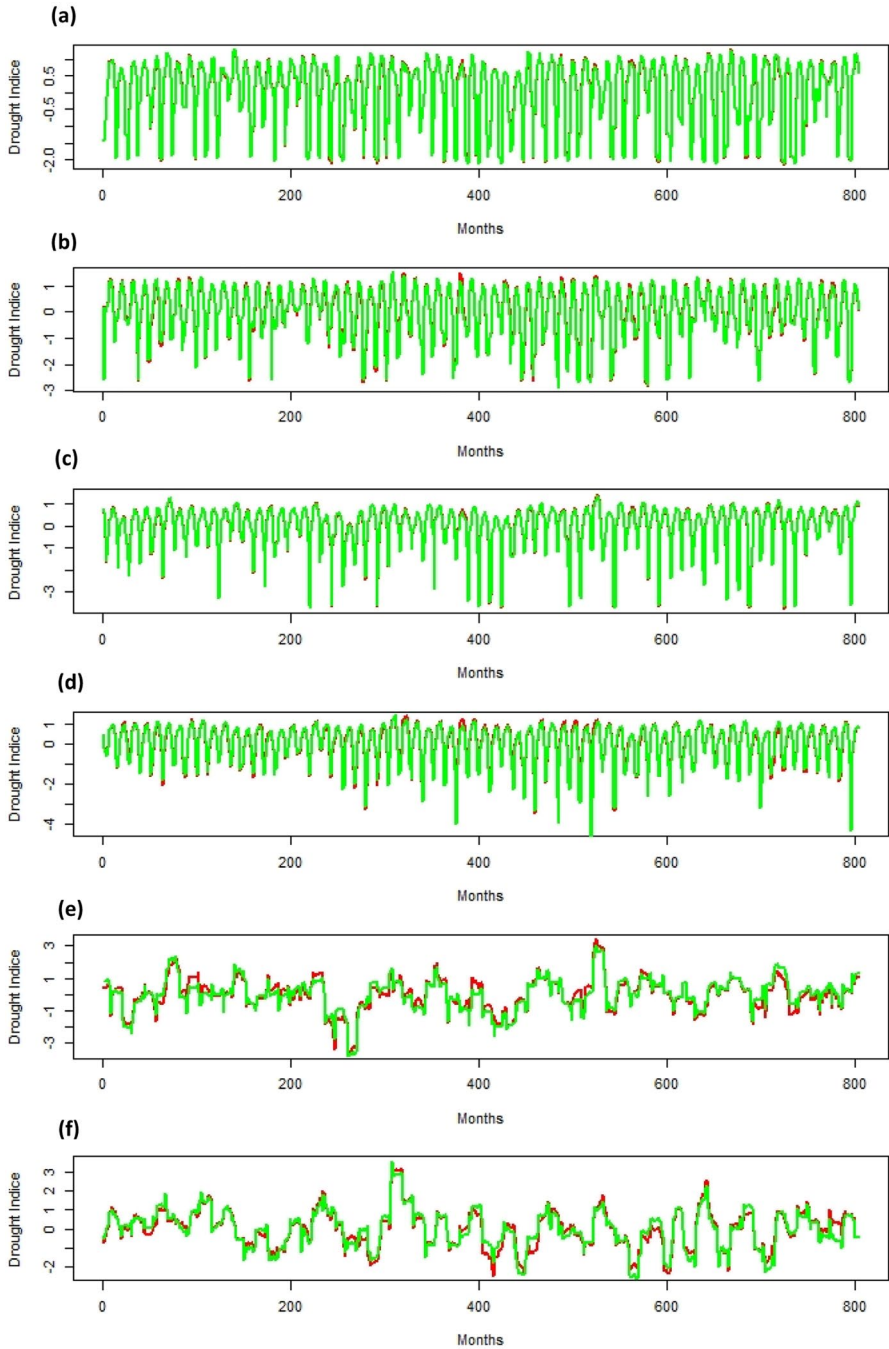
Fig. 2 NSPI and SSPI drought signals during 1951–2017 for 3-, 6- and 12-month time scales: (**a**) 3-month time scale for UGRB; (**b**) 3-month time scale for LGRB; (**c**) 6-month time scale for UGRB; (**d**) 6-month time scale for LGRB; (**e**) 12-month time scale for UGRB; (**f**) 12-month time scale for LGRB

The best fitted trivariate copula model among the copulas was decided by analysing the GoF measures–KS, CVM, Chsq considering 2000 sample runs including AIC, log-likelihood values and their estimated parameters are represented in Tables 4 and 5. The results from the tables show that the Student's t copula performed better rather than the other trivariate copula models. The parameter estimation of trivariate copula analysis was conducted using the MLE method. Higher dimensional Student's t copula is more

Fig. 3 NRDI and SRDI drought signals during 1951–2017 for 3-, 6- and 12-month time scales: (**a**) 3-month time scale for UGRB; (**b**) 3-month time scale for LGRB; (**c**) 6-month time scale for UGRB; (**d**) 6-month time scale for LGRB; (**e**) 12-month time scale for UGRB; (**f**) 12-month time scale for LGRB
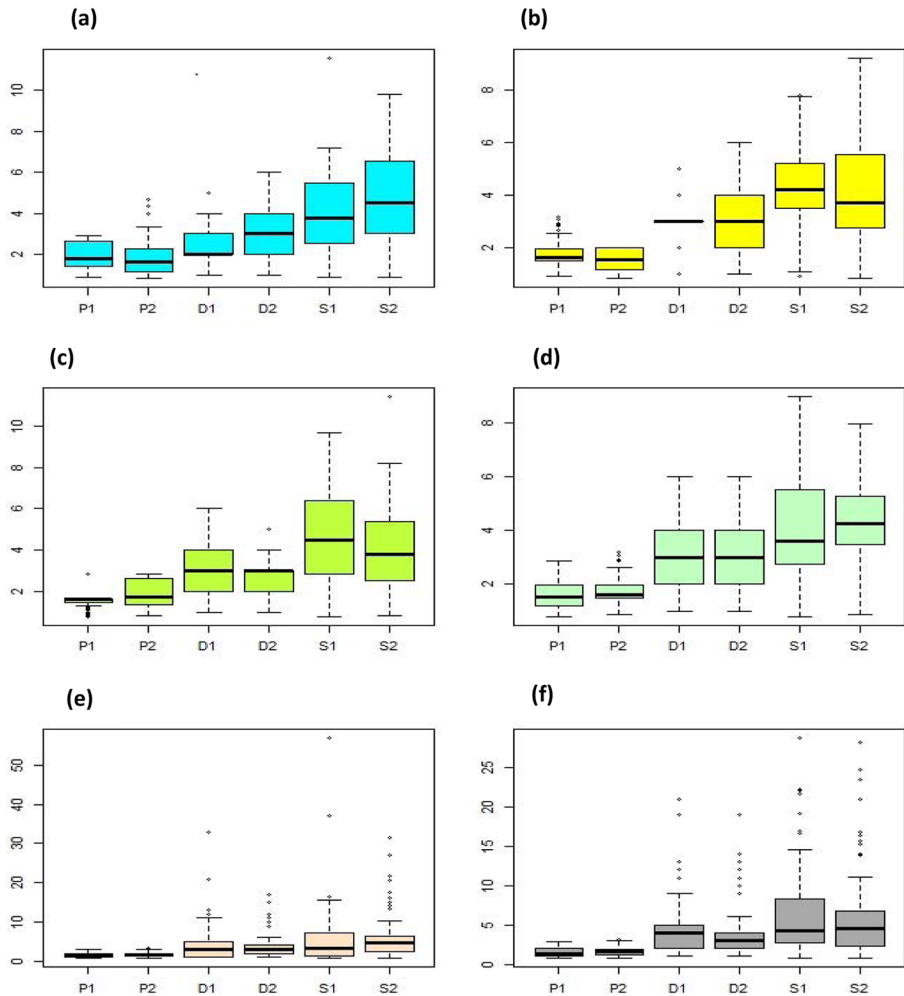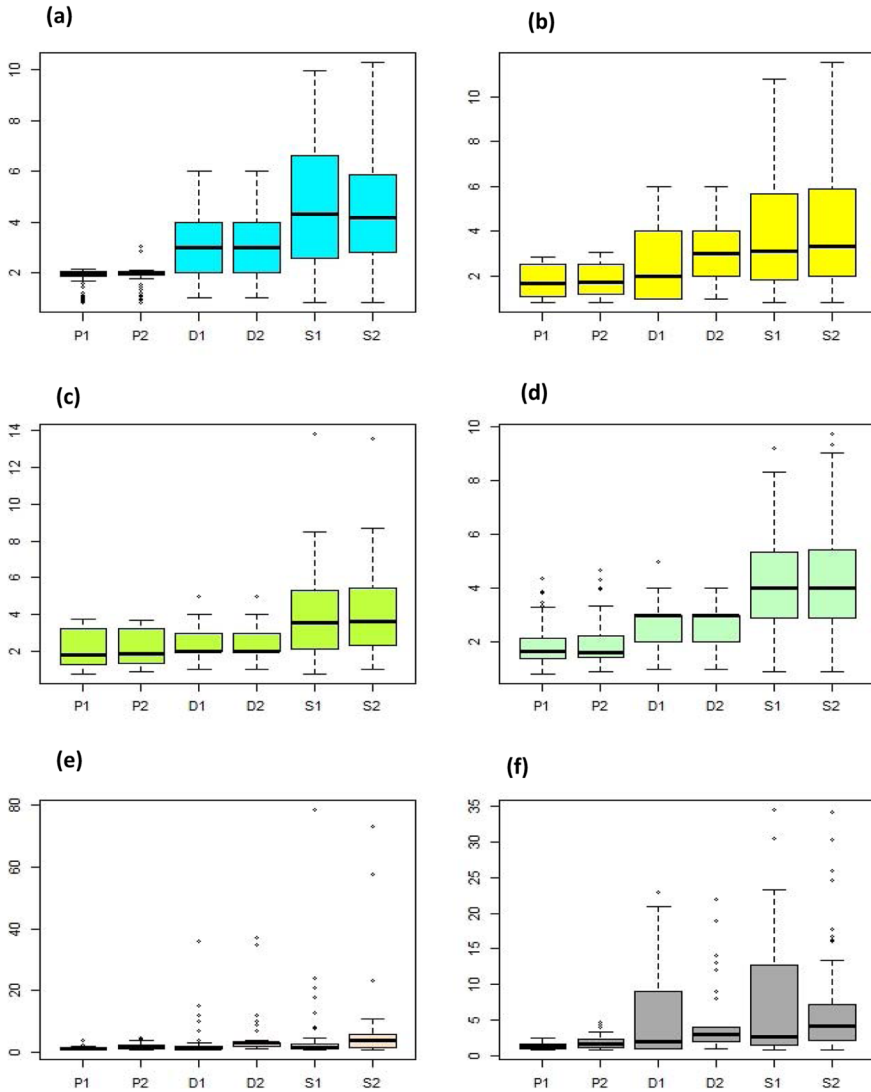
Fig. 4 Comparison by box plots for drought variables peak (P), duration (D) and severity (S) for NSPI and SSPI for different time scales: (**a**) 3-month time scale for UGRB (**b**) 3-month time scale for LGRB; (**c**) 6-month time scale for UGRB (**d**) 6-month time scale for LGRB; (**e**) 12-month time scale for UGRB (**f**) 12-month time scale for LGRB

popular in hydrological studies for modelling of different extreme events (Ganguli and Reddy 2013). The main limitation of modelling of the higher dimensional Student's t copula is that it has only a single degree of freedom for each pair of parameters (Aas et al. 2009). However, this limitation can be further efficiently modified by the PCC models (Tables 4 and 5).

Fig. 5 Comparison by box plots for drought variables peak (P), duration (D) and severity (S) for NRDI and SRDI for different time scales (**a**) 3-month time scale for UGRB (**b**) 3-month time scale for LGRB; (**c**) 6-month time scale for UGRB (**d**) 6-month time scale for LGRB; (**e**) 12-month time scale for UGRB (**f**) 12-month time scale for LGRB

4.4 Drought Characteristics Modelling Using Pair-Copula Models

Suitable vine structure between C-vine and D-vine models and the copula families must be selected for the dependent pair variables viz., peak (P)–duration (D), duration (D)–severity (S) and severity (S)–peak (P). The dependency measures between pair variables of drought events are given in Supplementary Table S7. This supporting document explains the stronger dependence between D and S. The next stronger dependency was observed

Table 4 Trivariate copula analysis for non-stationary reconnaissance drought characteristics

UGRB	3-month					6-month					12-month				
	Gumbel	Clayton	Frank	Normal	Student's t copula	Gumbel	Clayton	Frank	Normal	Student's t copula	Gumbel	Clayton	Frank	Normal	Student's t copula
Parameter	1.9	2.0	6.2	0.7	**0.7**	1.9	1.5	5.3	0.7	**0.7**	2.9	2.6	8.3	0.8	**0.7**
Loglikelihood	41.5	53.0	47.9	48.2	**97.5**	36.1	39.4	38.5	51.1	**81.3**	47.7	36.0	39.3	45.6	**70.6**
AIC	-81.0	-104.0	-93.8	-92.3	**-187**	-70.3	-76.7	-74.9	-98.2	**-154**	-93.4	-69.9	-76.6	-87.1	**-133**
KS	0.2	0.0	0.2	0.5	**0.4**	0.3	0.1	0.2	0.4	**0.3**	0.2	0.1	0.2	0.3	**0.8**
Chsq	0.4	0.1	0.3	0.4	**0.7**	0.1	0.0	0.1	0.1	**0.2**	0.1	0.0	0.0	0.0	**0.1**
CVM	0.2	0.1	0.1	0.2	**0.5**	0.1	0.1	0.1	0.4	**0.6**	0.0	0.0	0.0	0.0	**0.2**
LGRB															
Parameter	2.1	1.6	6.5	0.8	**0.8**	1.8	1.7	5.4	0.7	**0.7**	2.7	1.8	8.1	0.8	**0.8**
Loglikelihood	51.4	43.9	52.0	55.1	**96.3**	31.9	45.0	39.5	54.9	**86.9**	46.5	29.1	40.1	44.1	**68.7**
AIC	-100.8	-85.7	-102	-106.3	**-184**	-61.8	-87.9	-77.0	-105.7	**-165**	-91.1	-56.1	-78.1	-84.2	**-129**
KS	0.2	0.0	0.2	0.5	**0.1**	0.4	0.1	0.4	0.5	**0.8**	0.3	0.2	0.3	0.0	**0.0**
Chsq	0.1	0.0	0.2	0.3	**0.4**	0.1	0.1	0.2	0.0	**0.4**	0.2	0.0	0.1	0.1	**0.5**
CVM	0.1	0.0	0.0	0.0	**0.0**	0.0	0.0	0.0	0.4	**0.6**	0.2	0.1	0.1	0.2	**0.1**

Table 5 Trivariate copula analysis for non-stationary standardized precipitation drought characteristics

UGRB	3-month					6-month					12-month				
	Gumbel	Clayton	Frank	Normal	Student's t copula	Gumbel	Clayton	Frank	Normal	Student's t copula	Gumbel	Clayton	Frank	Normal	Student's t copula
Parameter	1.9	1.9	5.7	0.7	**0.7**	1.8	1.7	5.1	0.7	**0.7**	2.7	3.0	9.8	0.9	**0.9**
Loglikelihood	40.7	51.9	44.5	49.9	**88.2**	31.8	45.5	36.3	36.7	**96**	82.5	84.2	86.0	87.9	**138.8**
AIC	-79.5	-101.9	-87.1	-95.8	**-168**	-61.7	-89.1	-70.6	-69.4	**-184**	-163.1	-166.3	-169.9	-171.7	**-269**
KS	0.2	0.1	0.2	0.3	**0.6**	0.4	0.1	0.4	0.9	**0.6**	0.1	0.1	0.2	0.1	**0.6**
Chisq	0.0	0.2	0.2	0.1	**0.3**	0.2	0.1	0.1	0.0	**0.5**	0.0	0.0	0.0	0.0	**0.2**
CVM	0.1	0.0	0.0	0.8	**0.5**	0.5	0.2	0.3	0.6	**0.9**	0.4	0.3	0.3	0.1	**0.3**
LGRB															
Parameter	1.7	1.5	4.9	0.7	**0.7**	1.9	1.9	6.0	0.7	**0.7**	2.2	2.2	7.0	0.8	**0.8**
Loglikelihood	30.9	39.7	33.1	42.9	**87.5**	41.4	52.0	46.5	44.0	**106**	61.2	61.8	59.4	64.6	**137.5**
AIC	-59.7	-77.4	-64.2	-81.7	**-167**	-80.7	-102.0	-90.9	-84.0	**-205**	-120.3	-121.7	-116.8	-125.3	**-267**
KS	0.4	0.1	0.4	0.8	**0.3**	0.3	0.1	0.3	0.7	**0.9**	0.3	0.2	0.4	0.2	**0.1**
Chisq	0.0	0.0	0.0	0.0	**0.0**	0.0	0.0	0.0	0.0	**0.5**	0.0	0.0	0.1	0.0	**0.3**
CVM	0.0	0.0	0.0	1.0	**1.0**	0.3	0.1	0.2	0.2	**0.6**	0.2	0.1	0.1	0.1	**0.5**

between S and P. It can, hence, be concluded that S must be present between D and P for the D-vine structure. D-vine structure was selected for further analysis since it has more flexibility towards forming the pair copula rather than the C-vine structure because it creates a relationship of variables with a particular root variable that defines the key elements of the structure (Aas et al. 2009). Kendall's plots and Chi-plots of the pair variables P–D, D–S and S–P are given in the Supplementary Figs. S1-S12 signifying positive dependencies between the pair variables. The dependency measures and the Kendall's and Chi-plots of the pair variables signifies that D – S showed the strongest dependency than other pairs. The copula families for the pair variables were selected from various copulas based on Clarke and Vuong tests.

In the first phase of parameter estimation method, the sequential parameters were estimated using the MLE method for these pairs of drought variables D/S (θ_{DS}) and S/P (θ_{SP}). Then for the second phase of parameter estimation, the respective h functions (conditional distribution function) were computed between the pair variables D/S and S/P. The parameter, $\theta_{DP/S}$ was then estimated for drought variables (D/S, S/P). The copula families selected for the pair variables and the parameters estimated in the second phase ($\theta_{DP/S}$) considering the best fitted copula for the LGRB and UGRB for different time scales are given in Table 6.

In the second phase of parameter estimation, the copula was selected from the first phase of copula families fitted for estimation of the sequential parameter. For example, in the case of the generated drought characteristics from 3- month NRDI, Gaussian copula was selected for estimation of the second phase of the parameter. Finally, the PCC model, accounting for the drought variables was compared with the derived trivariate copula using the AIC criteria. It can be observed from Supplementary Table S8 that the AIC of the trivariate model showed higher AIC value compared to PCC based model. It can, hence, be

Table 6 Pairwise Copula Construction (PCC) for NSPI and NRDI drought characteristics

NSPI UGRB	Clarke	Vuong	AIC	Selected Bivariate Copula			Parameter ($\theta_{DP/S}$)
3-month	0.62	0.914	-180.48	Clayton	Gaussian	Clayton	0.469
6-month	0.457	0.506	-197.83	Clayton	Frank	Gaussian	0.429
12-month	0.0248	0.984	-286.033	Frank	Gumbel	Frank	6.16
NSPI LGRB	Clarke	Vuong	AIC	Selected Bivariate Copula			Parameter ($\theta_{DP/S}$)
3-month	0.0026	0.0619	-169.575	Clayton	Clayton	Gumbel	0.244
6-month	0.62	0.692	-205.71	Clayton	Gaussian	Frank	0.46
12-month	0.024	0.0607	-273.98	Gaussian	Gumbel	Gaussian	0.623
NRDI UGRB	Clarke	Vuong	AIC	Selected Bivariate Copula			Parameter ($\theta_{DP/S}$)
3-month	0.703	0.6874	-199.82	Clayton	Frank	Gaussian	0.4855
6-month	0.526	0.358	-168.57	Gaussian	Gaussian	Clayton	0.291
12-month	0.511	0.556	-140.8	Gumbel	Gumbel	Gumbel	1.79
NRDI LGRB	Clarke	Vuong	AIC	Selected Bivariate Copula			Parameter ($\theta_{DP/S}$)
3-month	0.055	0.009	-224.66	Gaussian	Frank	Clayton	0.362
6-month	1	0.459	-171.19	Clayton	Clayton	Gaussian	0.338
12-month	1	0.418	-147.97	Gumbel	Gumbel	Gumbel	1.694

justified that the PCC model can flexibly model the drought characteristics by transforming the bivariate model to a higher dimensional model.

4.5 Return Period Analysis of Drought Variables

The frequency analysis of droughts can be related to the occurrence of extreme events and their probability distributions. Here, the values of peak (P), duration (D) and severity (S) that exceed their truncation level ($S \geq s, D \geq d, P \geq p$) were considered for analysing the multivariate drought frequency. The joint return period analysis was carried out by using the two probability cases i.e. (i) ''AND'' and ''OR'' return periods for the drought variable using the approach suggested by Salvadori and Michele (2004). For annual drought analysis, the mathematical expressions for return period considering the ''AND'' and ''OR'' case are shown in Eqs. (16) and (17).

$$T_{(SDP)}^{AND} = \frac{\varepsilon}{P(S \geq s, D \geq d, P \geq p)} = \frac{\varepsilon}{1 - F_S(s) - F_D(d) - F_P(p) + F_{S,D}(s,d) + F_{D,P}(d,p) + F_{P,S}(p,s) - F_{S,D,P}(s,d,p)}$$
$$= \frac{\varepsilon}{1 - F_S(s) - F_D(d) - F_P(p) + C(u_1,u_2) + C(u_2,u_3) + C(u_3,u_1) - C(u_1,u_2,u_3)} \quad (16)$$

$$T_{SDP}^{OR} = \frac{\varepsilon}{P(S \geq s, D \geq d, P \geq P)} = \frac{\varepsilon}{1 - P(S \geq s, D \geq d, P \geq P)}$$
$$= \frac{\varepsilon}{1 - F_{S,D,P}(s,d,p)} = \frac{\varepsilon}{1 - C(u_1,u_2,u_3)} \quad (17)$$

Supplementary Tables S9-S12 show the joint return periods (''AND'' and ''OR'' cases) obtained from the trivariate copula and PCC models using the drought characteristics for NSPI and NRDI for 3-, 6- and 12- month time scale. Here $T_{(SDP)}^{AND}$TC and $T_{(SDP)}^{AND}$PCC represent the joint return periods for trivariate Student's t copula and PCC model respectively. Similarly, the joint return period of "OR" case can be denoted as T_{SDP}^{OR}TC and $T_{(SDP)}^{OR}$PCC for trivariate Student's t copula and PCC models respectively. The ''OR'' joint return periods are less compared to ''AND'' return periods in both trivariate Student's t copula and PCC models. It can, hence, be concluded that the frequency of drought is more in the ''OR'' case when compared to the "AND" case.

5 Conclusions

The concept of non-stationarity was employed by aggregating the precipitation and initial value series with the large-scale climate indices with lag time of 0–12 months. It can be concluded that the non-stationary meteorological droughts considering the large-scale climate indices are capable of capturing the drought events in comparison with stationary drought indices. Both precipitation and evapotranspiration based non-stationary drought indices can be applied to identify more complex aspect of drought occurrence. The probabilistic estimation of drought characteristics must be carried out to estimate the recurrence intervals of droughts. The standard multivariate copulas are not flexible enough to model the higher dimensional copula for assessment of extreme events. The drawbacks of multivariate copulas can be removed using D-vine copula models. D-vine PCC model was also used to find the drought return periods. The conclusions derived in the present study are:

1. The non-stationary models performed better compared to the stationary models as the AIC values are lower in case of non-stationary models. UGRB, showed a significant

influence at various lags for 3-, 6- and 12- month time scales. SOI, MEI and SST are the most influential large oscillations at different time scales are the most influential covariates.

2. The non-stationary and stationary models showed variations in their time signals. The box plots between drought variables revealed that the drought properties significantly varied under stationary and non-stationary conditions in the both the basins for 12-month time scale.

3. The "AND" and "OR' joint return periods for PCC models are higher compared to those obtained from the trivariate Student's t copula model for both the non-stationary models. The drought frequencies for PCC model is higher as compared to the trivariate copula model.

4. After analysing the trivariate and PCC models, the return periods showed variations between "AND" and "OR" return periods for drought indices in different time scales. The variations of return periods between trivariate Student's t copula and PCC model are significant in case of 12-month time scale for both the non-stationary drought events. To summarize, the 'AND' and 'OR' return periods predicted by PCC model are more reliable compared to trivariate Student's t-copula model as PCC model performed better than the trivariate Student's t copula.

Declarations

References

Aas K, Czado KC, Frigessi A, Bakken H (2009) Pair-copula constructions of multiple dependence. Insur Math Econ 44:182–198

Bazrafshan J, Hejabi S (2018) A non-stationary reconnaissance drought index (NRDI) for drought monitoring in a changing climate. Water Resour Manag 32(8):2611–2624. https://doi.org/10.1007/s11269-018-1947-z

Brechmann EC, Schepsmeier (2013) Modeling dependence with C- and D-vine copulas: the R package CDVine. J Stat Softw 52 (3): 1–27.

Chen M, Papadikis K, Jun C (2021) An investigation on the non-stationarity of flood frequency across the UK. J Hydrol 597:126309. https://doi.org/10.1016/j.jhydrol.2021.126309

Daneshkhah A, Remesan R, Chatrabgoun O, Holman IP (2016) Probabilistic modeling of flood characterizations with parametric and minimum information pair-copula model. J Hydrol 540:469–487. https://doi.org/10.1016/j.jhydrol.2016.06.044

Debele SE, Strupczewski WG, Bogdanowicz E (2017) A comparison of three approaches to non-stationary flood frequency analysis. Acta Geophys 65(4):863–883. https://doi.org/10.1007/s11600-017-0071-4

Dixit S, Jayakumar KV (2021a) A study on copula-based bivariate and trivariate drought assessment in Godavari River basin and the teleconnection of drought with large-scale climate indices. Theor Appl Climatol 146(3):1335–1353. https://doi.org/10.1007/s00704-021-03792-w

Dixit S, Jayakumar KV (2021b) Spatio-temporal analysis of copula based probabilistic Multivariate drought index using CMIP6 model. Int J Climatol. https://doi.org/10.1002/joc.7469

Dixit S, Tayyaba S, Jayakumar KV (2021) Spatio-temporal variation and future risk assessment of projected drought events in the Godavari River basin using regional climate models. J Water Clim Change 12(7):3240–3263. https://doi.org/10.2166/wcc.2021.093

Ganguli P, Reddy MJ (2012) Risk Assessment of Droughts in Gujarat Using Bivariate Copulas. Water Resour Manag 26:3301–3327. https://doi.org/10.1007/s11269-012-0073-6

Ganguli P, Reddy MJ (2013) Analysis of ENSO-based climate variability in modulating drought risks over western Rajasthan in India. J Earth Syst Sci 122:253–269. https://doi.org/10.1007/s12040-012-0247-x

Genest C, Favre AC (2007) Everything you always wanted to know about copula modelling but were afraid to ask. J Hydrol Eng 12(4):347–368. https://doi.org/10.1029/2006WR005275

Gräler B, van den Berg M, Vandenberghe S, Petroselli A, Grimaldi S, De Baets B, Verhoest N (2013) Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. Hydrol Earth Syst Sci 17(4):1281–1296. https://doi.org/10.5194/hess-17-1281-2013

Hao Z, AghaKouchak A (2013) Multivariate standardized drought index: a parametric multi-index model. Adv Water Resour 57:12–18. https://doi.org/10.1016/j.advwatres.2013.03.009

Harris I, Osborn TJ, Jones P, Lister D (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci Data 7(1):1–18. https://doi.org/10.1038/s41597-020-0453-3

Kao SC, Govindaraju RS (2010) A copula-based joint deficit index for droughts. J Hydrol 380(1–2):121–134. https://doi.org/10.1016/j.jhydrol.2009.10.029

Kendall MG (1955) Rank correlation methods. Hafner Publishing Co., Oxford, England

Kumar KS, Anand Raj P, Sreelatha K, Bisht DS, Sridhar V (2021a) Monthly and Seasonal Drought Characterization Using GRACE-Based Groundwater Drought Index and Its Link to Teleconnections across South Indian River Basins. Climate 9(4):56. https://doi.org/10.3390/cli9040056

Kumar N, Poonia V, Gupta BB, Goyal MK (2021b) A novel framework for risk assessment and resilience of critical infrastructure towards climate change. Technol Forecast Soc Change 165:120532. https://doi.org/10.1016/j.techfore.2020.120532

Kurowicka D, Cooke R (2006) Uncertainty Analysis with High Dimensional Dependence Modeling. John Wiley, Hoboken

López J, Francés F (2013) Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates. Hydrol Earth Syst Sci 17(8):3189–3203. https://doi.org/10.5194/hess-17-3189-2013

Mckee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scale. In: Proceedings of the Eighth Conference on Applied Climatology (vol. 17, p 179–184)

Mishra V, Aadhar S, Asoka A, Pai S, Kumar R (2016) On the frequency of the 2015 monsoon season drought in the Indo-Gangetic Plain. Geophys Res Lett 43:12102–12112. https://doi.org/10.1002/2016GL071407

Mortuza MR, Moges E, Demissie Y, Li HY (2019) Historical and future drought in Bangladesh using copula-based bivariate regional frequency analysis. Theor Appl Climatol 135(3):855–871. https://doi.org/10.1007/s00704-018-2407-7

Nelsen RB (2006) An Introduction to Copulas. Springer, New York

Poonia V, Goyal MK, Gupta BB, Gupta AK, Jha S, Das J (2021a) Drought occurrence in Different River Basins of India and block chain technology based framework for disaster management. J Clean Prod 312:127737. https://doi.org/10.1016/j.jclepro.2021.127737

Poonia V, Jha S, Goyal MK (2021b) Copula based analysis of meteorological, hydrological and agricultural drought characteristics across Indian river basins. Int J Climatol 41:4637–4652. https://doi.org/10.1002/joc.7091

Rashid MM, Beecham S (2019) Development of a non-stationary Standardized Precipitation Index and its application to a South Australian climate. Sci Total Environ 657:882–892. https://doi.org/10.1016/j.scitotenv.2018.12.052

Rigby RA, Stasinopoulos DM (2005) Generalized additive models for location, scale and shape. J R Stat Soc Ser C Appl Stat 54:507–554. https://doi.org/10.1177/0962280212473302

Russo S, Dosio A, Sterl A, Barbosa P, Vogt J (2013) Projection of occurrence of extreme dry-wet years and seasons in Europe with stationary and nonstationary Standardized Precipitation Indices. J Geophys Res Atmos 118:7628–7639. https://doi.org/10.1002/jgrd.50571

Saghafian B, Mehdikhani H (2014) Drought characterization using a new copula-based trivariate approach. Nat Hazards 72(3):1391–1407. https://doi.org/10.1007/s11069-013-0921-6

Salvadori G, De Michele C (2004) Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. Water Resour Res 40(12). https://doi.org/10.1029/2004WR003133

Shah D, Mishra V (2020a) Drought onset and termination in India. J Geophys Res Atmos 125(15):e2020JD032871. https://doi.org/10.1029/2020JD032871

Shah D, Mishra V (2020b) Integrated Drought Index (IDI) for drought monitoring and assessment in India. Water Resour Res 56(2):e2019WR026284. https://doi.org/10.1029/2019WR026284

Shukla S, Wood AW (2008) Use of a standardized runoff index for characterizing hydrologic drought. Geophys Res Lett 35:L02405. https://doi.org/10.1029/2007GL032487

Song SB, Kang Y (2011) Pair-copula decomposition constructions for multivariate hydrological drought frequency analysis. In: 2011 International Symposium on Water Resource and Environmental Protection (vol. 4, p 2635–2638). IEEE, Manhattan

Tamaddun KA, Kalra A, Bernardez M, Ahmad S (2019) Effects of ENSO on temperature, precipitation, and potential evapotranspiration of North India's monsoon: An analysis of trend and entropy. Water 11(2):189. https://doi.org/10.3390/w11020189

Tsakiris G, Pangalou D, Vangelis H (2007) Regional drought assessment based on the Reconnaissance Drought Index (RDI). Water Resour Manag 21(5):821–833

Vicente-Serrano SM, Beguería S, Lopez-Moreno JI (2010) A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. J Clim 23(7):1696–1718. https://doi.org/10.1175/2009JCLI2909.1

Villarini G, Smith JA, Serinaldi F, Bales J, Bates PD, Krajewski WF (2009) Flood frequency analysis for non-stationary annual peak records in an urban drainage basin. Adv Water Resour 32(8):1255–1266. https://doi.org/10.1016/j.advwatres.2009.05.003

Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57(2):307–333

Wang Y, Li J, Feng P, Hu R (2015) A Time-Dependent Drought Index for Non-Stationary Precipitation Series. Water Resour Manag 29:5631–5647. https://doi.org/10.1007/s11269-015-1138-0

Won J, Choi J, Lee O, Kim S (2020) Copula-based Joint Drought Index using SPI and EDDI and its application to climate change. Sci Total Environ 744:140701. https://doi.org/10.1016/j.scitotenv.2020.140701

Xu P, Wang D, Wang Y, Singh VP (2022) A Stepwise and Dynamic C-Vine Copula-Based Approach for Nonstationary Monthly Streamflow Forecasts. J Hydrol Eng 27(1):04021043. https://doi.org/10.1061/(ASCE)HE.1943-5584.0002145

Yevjevich V (1967) An objective approach to definitions and investigations of continental hydrologic drought. Fort Collins, Colorado (Hydrology Paper no. 23)

Zarch MAA, Sivakumar B, Sharma A (2015) Droughts in a warming climate: A global assessment of Standardized precipitation index (SPI) and Reconnaissance drought index (RDI). J Hydrol 52:183–195. https://doi.org/10.1016/j.jhydrol.2014.09.071