Regional analysis of drought Severity-Duration-Frequency and Severity-Area-Frequency curves in the Godavari River Basin, India.

K Satish Kumar[1], P AnandRaj[2], K Sreelatha[3] and Venkataramana Sridhar[4*]

[1,3] Research Scholar, Department of Civil Engineering, National Institute of Technology, Warangal, India

[2]Professor, Department of Civil Engineering, National Institute of Technology, Warangal, India

[4]Deparment of Biological Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

[1]satish005@student.nitw.ac.in, [2]a_raj_p@yahoo.co.in, [3]koppalasreelatha@student.nitw.ac.in

[4*]vsri@vt.edu (corresponding author)

Abstract

India is one of the most drought-ravaged countries in the world and faces at least one drought in one region or another in every 3 years. There is no single reliable approach in characterizing future droughts. To understand future drought risk, potential changes of drought properties and characteristics are analyzed in this study. Using Fuzzy c-means clustering approach, homogeneous drought regions are identified in the Godavari river basin and therefore, optimum number of clusters were assigned as four. The 12-month standardized precipitation index (SPI) using precipitation data from India Meteorological Department (IMD) and Global Climate Model (GCM) - MIROC-ESM-CHEM is calculated for the homogeneous regions of the Godavari basin. The best fit copula for observed and simulated severity and duration are: region 1-Clayton, Region 2 and 3 -Gumbel, Region 4- Frank copula. Severity-duration-frequency (SDF) and severity-area-frequency (SAF) curves were developed and analyzed using the best fit copulas. The research findings conclude that moderate and severe droughts are frequently increasing for future periods (2006-2099) compared to the historic period (1962-2005). Droughts with high severity and high mean interarrival time are observed as expected in the future. For the Godavari basin, the SDF curves were concave upward indicating an increase in severity with an increase in duration. The rate of increase of severity is small for shorter durations compared to that of longer-duration drought. Thus, more prolonged drought events in the twentieth-first century are likely to occur. The SAF curves with steeper slopes and high variability in topographical and hydrological characteristics have been observed over the Godavari basin. From these curves, for a specified percentage of area and return period, the drought severity can be calculated and the information can be used for crop management and agricultural water demands. Overall, the findings of this research offer a view of likely scenarios of drought in the Godavari basin.

Key words: Severity, Duration, Copula, SPI, Global Climate Model

1. Introduction

Drought is a natural threat considered when precipitation deficits are encountered for extended periods. Drought is the temporary deviation from normal weather conditions that may occur in a variety of climates. Whereas, aridity is a permanent climate feature based on long-term climatic conditions in an area (Karamouz et al., 2013, Araghi et al., 2018). As a result of climate change, aridity and drought are expected to increase, that may have significant implications on agriculture and ecosystem (Dai, 2013; Lobell et al., 2008). Water shortages arise when drought occurs over a season or longer periods, and the water that is available becomes insufficient to meet the demands of human activities and the environment. Owing to the negative impacts on the cultural, environmental and social sectors, this natural and recurrent climate pattern has adversely affected society throughout history. The characteristics of dryness are thus recognized as an important factor in water resource planning and management (Shiau, 2006). Drought can be characterized by various factors such as severity, intensity, and duration, and is therefore described as a complex natural disaster (Mishra and Singh, 2010; Kang et al., 2019). In contrast to other natural disasters, droughts also impact over a large region (Wilhite et al., 2014; Sehgal et al., 2017; Kang and Sridhar, 2017; Sehgal and Sridhar, 2019). Thus, during the analysis of drought condition, multivariate complexity and spatial variations should be considered.

Almost every three years, India faces drought in different parts of the country (Goyal et al., 2017). Over the last few decades, India has experienced long and severe droughts, and similarly, in recent times, frequency has also increased in many parts of the country (Bisht et al., 2019). During June to September, a significant amount (about 70-90 %) of annual rainfall occurs due to the southwest monsoon in India (Bisht et al., 2017a, b). Since most of the country depends on monsoon rainfall, drought may be caused when monsoon fails in India (Kumar et al., 2013). The study and distribution of drought conditions in India has been studied by several researchers from the literature (Gupta et al., 2020; Das et al., 2016; Jain et al., 2015; Reddy and Ganguli, 2012; Guhathakurta and Rajeevan, 2008). The vulnerability of drought is amplified by climate change that adversely affects India's water supplies. Therefore, the regional study of near-future spatial and temporal characteristics of drought will provide immense knowledge for decision-makers and planners to frame policies to mitigate the impacts of drought hazards.

For the assessment of drought conditions both globally and regionally, climate forecasts are widely applied in the 21st century (Kang and Sridhar, 2018; Thilakarathne and Sridhar, 2017; Masud et al., 2017; Xu et al., 2015; Wang and Chen, 2014; Dai, 2013). Yet uncertainty must

be kept in mind when contemplating possible climate scenarios. Dai (2013) found that global warming causes the intensification of drought conditions. GCMs are the only models available on a global scale to investigate climate change, but they are not suitable on a regional scale because of their grid size. Downscaling is also the only tool that can be used to measure climate change on a regional scale (Sehgal et al., 2018). The support vector machine is the most widely used machine learning technique used for downscaling among different downscaling methods.

Few studies have been published on drought climate predictions in Indian conditions. Numerous researches on the spatial and temporal distribution of monsoon rainfall have been carried out. Few researchers find a rise in average precipitation over the monsoon season along with inter-annual variability (Fan et al., 2012; Chaturvedi et al., 2012)

Globally, various regional studies have been carried out to understand the characteristics of spatial and temporal drought. Via multivariate modeling, such as Severity-Duration-Frequency (SDF) and Severity-Area-Frequency (SAF) curves, regional drought risk can be assessed (Reddy and Ganguli, 2013, Mishra and Singh, 2010). These drought curves are evaluated for different return periods based on mathematical relationships between severity, intensity, interarrival time, duration, and area coverage at various return periods. To denote all characteristics of droughts, these curves quantitatively relate the severity, duration/area and return period (Rajasekhar et al., 2015; Kang and Sridhar, 2020).

Currently, the concept of copula has been widely used to measure bivariate and multivariate joint probability distributions in hydrology and water management engineering (Thilakarathne and Sridhar, 2017; Liu et al., 2016; Tosunoglu and Can, 2016; Salvadori and De Michele, 2015; Ganguli, 2014; Zhang et al., 2012). In modeling the characteristics of two or more dependent variables such as severity, duration, and area of drought, the Copula functions are highly effective and efficient. This is because of the function that maintains a strong correlation between the variables considered and is not bound to the identical distribution of marginal probability with regard to long-term prediction.

The objective of this study is to evaluate the SDF and SAF regional drought curves under the conditions of drought in the Godavari basin, as it is one of the largest river basins with relatively high-water potential in India. A sudden and vital drop in the water level are observed in the Godavari basin. It is worth noting that this study is the first study to regionally examine the temporal and spatial variations in drought events, emphasizing as a multivariate analysis the significance of drought SDF and SAF curves under drought. Our analysis includes-

identifying homogeneous drought regions in the Godavari river basin using Fuzzy c-means clustering approach; quantifying the future changes in drought climatology with standardized precipitation index (12 month) over Godavari river basin using MIROC-ESM-CHEM global climate model (RCP 8.5 scenario) data for the three future simulations i.e., 2006-2039, 2040-2069, and 2070-2099; employing copula-based methodology to derive SDF curves by analyzing the changes in the joint return period of severity and duration; and deriving the SAF curves which examine the changes in drought return periods covering a specific percentage of areal extent and the corresponding severity values.

2. Methodology

2.1 Study area

The second biggest river basin in India is the Godavari River Basin that covers a total area of 3,12,812 Sq.km. With a total length of 1465 km, the Godavari basin is situated between 16°19' to 22°34' N latitude and 73°24' to 83°4' E longitude. Annual precipitation ranges from 755 mm to 1531 mm, with almost 84% of the total precipitation occurring during the monsoon season. About 60 % of the area in the Godavari basin is covered by agricultural land and 30 % by forest areas.

2.2 Fuzzy c-means clustering

Dunn (1973) proposed the concept of fuzzy c-means clustering (FCM) and Bezdek (1981) expanded it further. Consider a cluster c with M objects where Y_k is the data vector for k_{th} $(k = 1,2, \dots, M)$ object. The technique of FCM minimizes the objective function,

$$J(U, C) = \sum_{j=1}^{M} \sum_{i=1}^{c} u_{ik}^{\theta} \|Y_k - C_i\|^2 \quad (1)$$

where u_{ik} denotes the k_{th} point membership value in the i_{th} cluster, C_i is the centre of the i_{th} cluster $(i = 1,2, \dots, c)$, $\|Y_k - C_i\|^2$ is the Euclidean distance squared and C_i and θ is the fuzziness index or the fuzzifier that can have any value > 1. The number of clusters and the data vector of the cluster center are randomly selected in the FCM algorithm, and the membership matrix is then calculated as

$$u_{i<k}^{t} = \left[\sum_{j=1}^{c} \left(\frac{\|y_k - c_i\|}{\|y_k - c_j\|} \right)^{\frac{2}{\theta - 1}} \right]^{-1} \quad (2)$$

With the updated membership values, new cluster centers are calculated as follows

$$C_i = \frac{\sum_{k=1}^{M} u_{ik}^{\theta} y_k}{\sum_{k=1}^{M} u_{ik}^{\theta}} \quad (3)$$

Using new cluster centers, the membership matrix is updated as

$$u_{i<k}^{t+1} = \left[\sum_{j=1}^{c}\left(\frac{\|y_k - c_i\|}{\|y_k - c_j\|}\right)^{\frac{2}{\theta-1}}\right]^{-1} \quad (4)$$

If $\|U^{t+1} - U^t\| < \varepsilon$ the algorithm will stop. If not, it goes back to step 1.

In the present analysis, three validity indices were used to check the efficiency of clustering.

2.2.1 Separation index (Si)

The compactness and separation function (Si) proposed by Xie and Beni (1991) was defined as the ratio of variance within clusters to the variance between clusters. The Si value is lower, the clustering is better.

$$Si(U, V: X) = \frac{\sum_{i=1}^{c}\sum_{k}^{M}(u_{ik}^{a})\|c_i - y_k\|^2}{M \min\limits_{i\neq k}\|v_i - y_k\|^2} \quad (5)$$

2.2.2 Fuzziness partition index (Fpi)

The membership (fuzziness) that various classes share (Bezdek, 1974) is calculated by the Fuzziness partition index.

$$Fpi(U) = \frac{1}{m}\sum_{i=1}^{c}\sum_{k=1}^{M}u_{ik}^2 \quad (6)$$

2.2.3 Partition entropy (Pe)

Pe is represented as

$$Pe(U) = \frac{1}{m}\sum_{i=1}^{c}\sum_{k=1}^{M}u_{ik}\log_a(u_{ik}) \quad (7)$$

Lower the Pe value, better the clustering and it varies between 0 and $\log(c)$.

The Fuzzy c means clustering analysis is performed using R software (package: ppclust; version 3.5.1; https://www.r-project.org/)

2.3 Copula-based joint distribution function

Copula is a function developed by Sklar (1959) that joins or couples two or more random variables. It transforms two or more marginal distributions into bivariate or multivariate distributions of random variables.

Let X and Y be a pair of random variables with marginal cumulative distribution functions $F_X(x)$ and $F_Y(y)$. From Sklar's theorem, using Copula function C, the joint distribution function $F_{XY}(x, y)$ of the two dependent random variables can be expressed as

$$P(X \leq x, Y \leq y) = F_{XY}(x, y) = C(F_X(x), F_Y(y)) \quad (8)$$

Copulas are established from different families and further information about different copula families and their properties are found in Nelsen (2006). The significant copula classes are Archimedean, Extreme, Elliptical and Plackett copulas. In the present study, Gumbel, Clayton and Frank from Archimedean class, Galambos from Extreme class and Plackett copulas are used to model joint dependence of drought variables. Table 1 displays the cumulative distribution functions (CDF) expressions, the corresponding probability density functions (PDF) and the associated copula family set parameters.

2.3.1 Copula parameter estimation

The most frequently used methods for estimating copula parameters are (i) method of moments (MOM; Genest and Rivest, 1993), (ii) inference from margins (IFM; Joe, 1997) (iii) exact maximum likelihood method (EML), and (iv) maximum pseudo-likelihood method (MPL; Genest et al., 1995). The MPL method is used in the present study for copula parameter estimation. First using the maximum likelihood estimation, the marginal distribution parameters of drought variables are calculated. Then the optimal copula parameters are calculated with the maximization of the pseudo log-likelihood function (Shiau, 2006). Maximum Likelihood Estimation (MLE), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Mean Square Error (MSE) measurements are used to assess the performance of fitted copula models. The best copula model represents the lowest MLE and the highest AIC, BIC and MSE values (Reddy and Ganguly, 2012, Amirataee et al., 2018).

In the present study, drought characteristics (severity, duration) were determined from the run theory principles. High correlation between severity and duration was found, so both were modeled together for further analysis. To find a best fit copula, univariate distribution was fitted for all the variables for which dependence is being modeled. To find the best marginal distribution, namely exponential, normal, log-normal, gamma, Weibull, and Gumbel, Kolmogorov-Smirnov (K-S) and Anderson-Darling (A-D) are calculated. The distribution with minimum K-S and A-D values were selected for the frequency analysis. In this analysis, the maximum pseudo-likelihood approach was adopted to estimate copula parameters. The copula analysis is performed using R software (package: copula; version 3.5.1; https://www.r-project.org/)

2.4 Model Selection and Bias Correction

Global Climate Models (GCMs) are the mathematical tools used in global and regional studies to identify physical processes of the land, ocean and atmosphere to simulate impacts. When

selecting appropriate GCMs, care should be taken as it creates significant uncertainty leading to either overestimation or underestimation of the projected climatic variables (Knutti and Sedláček. 2013; Shivam et al., 2019). In India, numerous studies have been carried out to select appropriate GCMs for climatic parameters such as maximum temperature, minimum temperature, sea surface temperature, precipitation, and spatially and temporally mean surface temperature (Sreelatha and AnandRaj, 2019, 2020; Perkins et al. 2007; Maxino et al. 2008).

Panjwani et al. (2019) selected four suitable GCMs for India based on the literature for precipitation, namely GFDL-ESM2M, IPSLLR, MIROC-ESM-CHEM, and FIO-ESM, as these models' performance is relatively better than other models compared to the correlation coefficient and root mean square error. Similarly, compared to other models, Chaturvedi et al. (2012) selected MIROC-ESM-CHEM as one of the better performers for precipitation. Also, it is commonly noted that projections generated by RCP 8.5 will show greater rainfall than those generated by the other RCPs (Ahmed et al. 2019; Ouyang et al. 2015). Therefore, for the historical period 1962-2005, the output of the MIROC-ESM-CHEM model (RCP 8.5 scenario) and the three future simulations, i.e., 2006-2039, 2040-2069 and 2070-2099, were used for this study. For the period 1962-2017 with a spatial resolution of $1° \times 1°$ for the study area, the monthly grid precipitation data from the India Meteorological Department (IMD) is considered (Rajeevan et al., 2008). The GCM dataset was converted to the IMD grid precipitation dataset resolution, i.e., $1° \times 1°$, using bilinear interpolation. The linear bias correction method was carried out to bias the precipitation dataset (Lenderink et al. 2007). For IMD and GCM datasets, 12-month SPI is evaluated and drought frequency analysis is performed with drought properties obtained from SPI-12. For detailed procedure regarding SPI, refer McKee et al., 1993.

2.5 Drought characterization using standardized precipitation index (SPI)

In meteorological terms, drought is defined as the deficiency of precipitation over an extended period. In this study, droughts are characterized using SPI, which is a significant indicator of meteorological drought that require only precipitation amounts to compute drought events. It is a normalized score that denotes an event departure from the mean and is represented in standard deviation units. SPI is simple, spatially invariant, and probabilistic in nature that can be applied to meteorological, hydrological and agricultural drought phenomenon. Temporal analysis of drought events can be estimated at different accumulation time scales such as 3, 6, 9, 12, 24, 36, and 48 months (Wu et al., 2007). For a given year i, calendar month j, for a time scale of k the following steps are followed for computing SPI (McKee et al., 1993).

- Calculation of cumulative precipitation series X_{ij}^{k} $(i = 1,2, \dots, n)$ for a specified period j. Each term represents the sum of precipitation of $k-1$ past successive months.
- Fitting of a cumulative probability distribution function (for e.g., gamma distribution) on accumulated monthly precipitation series and the gamma distribution function is defined as $g(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$; where, α and β are shape and scale parameters. $\Gamma(\alpha)$ is the gamma function. Then, the corresponding cumulative distribution function (CDF) $G(x)$ is calculated.
- A mixed gamma distribution function with zeros and continuous precipitation amount is employed and the corresponding CDF is given as $F(x) = q + (1-q)G(x)$; where, q is the probability of zero precipitation and $G(x)$ is the distribution function for nonzero precipitation.
- Performing equiprobability transformation proposed between CDF of mixed distribution and the CDF of the standard normal distribution with zero mean and unit variance (Panofsky and Brier, 1958), we get $SPI = \psi^{-1}(F(x))$. This transformed probability is the SPI. A positive (negative) value of SPI series shows that the value is above (below) average precipitation.

A drought period is assumed as a consecutive number of months with SPI values below certain threshold value (approximately a threshold value of -0.8). According to McKee et al., (1993), droughts are classified into four categories as presented in Table 2.

Droughts are often characterized by their severity, duration, and peak, and frequency of occurrences. Drought duration (D) is the period of time where the SPI remains below the fixed threshold value. The minimum duration of drought is considered as one month, as the drought event is defined at aggregation of monthly time scale. Drought severity (S) is the cumulative values of SPI within the drought duration. According to McKee et al., (1993), the severity of drought event i, S_i $(i = 1,2, \dots)$ is taken as positive and is given as

$$S_i = -\sum_{i=1}^{D} SPI_i \qquad (9)$$

In India, major portion of annual rainfall is received during monsoon season (4 months) and is followed by almost 8 months of the dry period. So, the hydrologic regime can be represented well by 12-month scale over India. Therefore, in this study SPI is evaluated at 12-month time scale. Many researchers have analyzed the drought conditions at 12-month timescale in India (Shah and Mishra 2015, and Gupta et al. 2020).

2.6 Drought Frequency Analysis

In the present research, the definition of runs was adopted to measure the characteristics of drought (severity and duration) (Yevjevich, 1967). Suppose, X_t is a drought variable with a time series t, then a run is a portion of the time series where all the values are either above or below a fixed threshold X_0. Therefore, value above (below) the threshold is denoted as positive run (negative run). As the threshold level may vary or constant with time, the drought properties mainly depend on the chosen threshold (Mishra and Singh, 2010). A threshold of 20 percentile (about a threshold value of -0.8) value of SPI in the grid is selected in this study so that a value below that threshold indicates an occurrence of drought (Reddy and Ganguli. 2013).

Several researchers have studied the drought properties using univariate frequency analysis (Tallaksen et al., 1997; Chung and Salas 2000; Cancelliere and Salas 2004). Therefore, researchers like Shiau and Shen (2001), González and Valdés (2003), Kim et al. (2006), Shiau (2006), Nadarajah (2009), Reddy and Ganguli (2012), Rajsekhar et al. (2015), and Gupta et al. (2020) have generalized univariate cases to bivariate frequency analysis since the joint behavior of multiple characteristics is not revealed in univariate cases. It is therefore important to examine the joint actions of drought characteristics for regional drought assessment and planning. This has led to the development of severity-duration-frequency curves and severity-area-frequency curves.

2.6.1 Severity-duration-frequency (SDF) analysis

SDF curves are useful multivariate tools for regional and global drought frequency analysis. The following are few studies from literature that performed SDF analysis for drought characteristics at different parts of the world. SDF analysis for wet periods and derived ISO-severity maps with return periods were performed for Greece by Dalezios et al. (2000). Similarly, for Iran, Saghafian et al., (2003) established SDF curves and ISO-severity maps. Using either empirical or semi-empirical formulations, these studies developed SDF curves. An analytical approach using copula for the derivation of SDF curves was therefore presented by Shiau and Modarres (2009). Reddy and Ganguli (2012), Halwatura et al., (2015), Rad et al., (2017), Adarsh et al., (2018), Samantaray et al., (2019) and Gupta et al., (2020) are other notable studies that have derived SDF curves based on copula. The following steps are involved in the derivation of the SDF curves using copula in this study.

- Standard goodness of fit statistics is selected to assess the best fitting marginal distribution for severity and duration.

- A joint and conditional marginal distribution is constructed for severity and duration using a best fit copula method.
- Relationship among severity, duration and frequency is established in terms of return period for drought events using the conditional recurrence interval provided by Shiau et al., (2007).

$$T_{S|D}(s|d) = \frac{1}{\gamma(1-F_{S|D}(s|d))} \quad (10)$$

where, d = drought duration, s = drought severity, $\gamma =$ arrival rate, $F_{S|D}(s|d)$ and $T_{S|D}(s|d)$ are the conditional CDF and conditional recurrence interval of S given $D = d$ respectively. The expression for conditional CDF is given below

$$F_{S|D}(s|d) = \frac{\partial F_{S,D}(s,d)}{\partial F_D(d)} \quad (11)$$

where $F_{S,D}(s,d) =$ joint CDF and $F_D(d)$ represents the CDF of drought duration.

- SDF curves for different return periods is derived from equations (10) and (11).

2.6.2 Severity-area-frequency (SAF) analysis

The return period of a drought with severity covering a particular percentage of areal extent is defined by SAF curves and it indicates that drought has occurred (Burke and Brown, 2010). Few earlier studies that developed SAF curves are Tase (1976), Santos (1983), Kim et al. (2002), Hisdal and Tallaksen (2003), Loukas and Vasiliades (2004), Mishra and Singh (2009), Burke and Brown (2010), Bonaccorso et al. (2014), Rajsekhar et al. (2015), Amirataee et al. (2018), and Gupta et al. (2020). To evaluate SAF curves, the following procedure is adopted.

- Using run theory, annual severity is estimated at each grid point.
- Different thresholds of severity are computed for various areal extents and is expressed in terms of percentage of total area for each year.
- The parameters of different distributions are calculated using L-moments and the best fit distribution is identified for severity values for various areal extents.
- Finally, the frequency analysis is performed to calculate the return periods involved with different drought severity values for various percentage areal extents.

In this study, drought severity was estimated using the run theory and areal extents were calibrated using severity for each drought region. Various probability distributions were identified to find best fitted distribution using BIC. Further, return levels were calibrated for

severity at various spatial extents and thus SAF curves were developed for various drought regions.

3 Results

Precipitation data from the Global Climate Model, MIROC-ESM-CHEM was collected and converted to the IMD precipitation data resolution i.e., $1° \times 1°$ by using linear interpolation method. To eliminate the bias in the precipitation dataset, a linear bias correction technique was then carried out. In this study, FCM was applied for the evaluation of the homogeneous drought regions at the regional scale.

3.1 Formation of homogeneous regions

For the Godavari Basin, the FCM concept was applied to the matrix containing mean annual precipitation, standard deviation, skewness and kurtosis. The selected variables were normalized before the application of cluster analysis because of the variation in the units. The efficiency of FCM depends on the fuzzifier index and optimum number of clusters so, the clusters were varied from 2 to 6 (Urcid and Ritter, 2012) with a fuzzifier index of 2 (Pal and Bezdek, 1995). The optimum number of clusters were selected using validity indices Si, Fpi, and Pe as discussed in the methodology that were provided in Table 3. From Table 3, Si, and Pe values were minimum at four and Fpi value was minimum at five number of clusters. Therefore, optimum number of clusters were assigned as four as shown in Figure 1.

3.2 Characterization of drought using SPI 12

The SPI-12 values were computed for each homogenous region of the Godavari basin using IMD (1962-2016) and GCM (MIROC-ESM-CHEM for RCP 8.5 scenario-2006-2099) monthly precipitation datasets. Figure 2 and Table 4 show the observed SPI time series (IMD:1962-2017) for four homogeneous drought regions with five major drought events. From Table 4, the most severe and longest drought was observed from August 2001 to June 2005 with a severity of -40.29 and a duration of 47 months for region 3. In region 1, with a severity of -39.23 and a duration of 43 months, the longest drought period was observed between October 1984 and April 1988. For region 2, the most severe drought was observed from July 1972 to May1975, with a severity of -32.21 and a duration of 35 months. Whereas for region 4, the most recent drought occurred from June 2008 to August 2010 with -26.32 severity and 27 months of drought duration. All regions experienced droughts during 1971-1975, 2002-2003 and 2008-2010 periods.

Figure 3 shows the historical (1962-2005) and future (2006-2099) periods of the GCM based SPI time series for four homogeneous regions. From Figure 3, it is noticed that all regions are projected to experience the maximum number of droughts between 2055 and 2085, with region 2 being more severe. Although the number of droughts in region 3 is expected to be lower compared to other regions, to better understand the drought situation in the twenty-first century, it is important to analyze the drought characteristics in detail. Further, the method of run (Yevjevich 1967) was applied to the SPI series with a threshold of −0.8 to evaluate drought characteristics for four homogeneous regions. Figure 4 shows the scatterplots and histograms for observed datasets of drought intensity versus duration in these defined regions.

Table 5 represents the drought properties for each homogeneous region based on IMD and considered GCM dataset. GCM historic period exhibited a greater number of droughts events compared to observed dataset (IMD). For the projected period, number of droughts decreased from 2006-2039 to 2040-2069 and then increased for the period 2070-2099 with the exception of region 3. The maximum severity was increased (decreased) for the time window 2070-2099 compared to 2006-2039 for regions 1 and 4 (regions 2 and 3). However, among the four regions, maximum severity of 51.1 (historic) and 41.8 (future 2070-2099) was observed for region 1 and 4 respectively. Region 2 and 3 has a longer drought interarrival time for historic period and hence, it experiences fewer droughts. Whereas in future projections, longer drought interarrival time and fewer droughts were observed for all regions.

Figure 5 summarizes the evolution of drought in time by examining drought characteristics for GCM and IMD time periods for each homogeneous drought region. From Figure 5, during the historical period (1962-2005), region 4 experienced the highest number of drought events, followed by region 2. However, the number of projected drought events were found to be less compared to the historical period. For the historic period, the maximum drought events were observed in region 4 compared to other three regions. The number of drought events were found to be decreasing in all regions from historic to future period 2006-2039, while, drought events were increased in region 1 for the future period 2040-2069 compared to 2006-2039 and then an increase of drought events were observed in region 2, 3, and 4 for 2070-2099 compared to 2040-2069.

For the observed, historical, and future time scales, region 3 experienced the maximum mean interval time between drought events, except in 2070-2099, from Figure 5. For future period, region 1, 2 reported the highest mean interval time for projected future periods, respectively. Whereas decrease in mean interarrival time was observed in 2070-2090 for region 3 and 4.

According to Figure 5, region 1 experienced the maximum severity for both observed and historical time period and therefore it is subsequently decreased in future period 2006-2039 and 2040-2069 and increased severity was observed in 2070-2099. Similar to the mean interarrival time, with the exception of 2070-2099, future periods in region 3 have experienced maximum severity. A decline in maximum severity for region 2 was observed for the predicted future period. However, for the remaining regions, maximum severity was fluctuating between the historic and future period.

From Figure 5, the maximum duration was the highest in region 3 from the IMD dataset. The maximum duration was observed for region 4 in the historical period, and minimum for region 1 for both historical and future periods. In comparison with observed GCM dataset, an increase of maximum duration was observed for projected future periods in all regions except in 2070-2099.

Overall, region 1 and 2 has fewer droughts with higher severity and high mean arrival time in the future. For region 3, drought events are less but maximum severity and duration are expected to occur in the future. For region 4, more drought events of relatively high severity and duration were observed in the projected future. Overall, it is expected to experience a greater number of droughts with high mean arrival time, severity, and duration for most part of the Godavari Basin in the future. Gupta and Jain (2018) reported that, in most regions of the country, increase in the PET was observed at a higher rate compared to the rainfall. Increased dryness is therefore anticipated in the latter part of the 21st century, leading to a rise in the severity and duration of droughts.

Figure 6 represents the frequency of occurrence of different droughts classes (moderate, severe, and extreme) over four homogeneous regions for observed and projected future time scales. The occurrence of moderate droughts is higher in historic and projected future time scales for all regions. For region 4, the occurrence of moderate and severe droughts were higher compared to other regions. Moderate drought has been observed in all the regions for future projection periods. Significant increase in the occurrence of moderate and severe droughts were observed in almost all regions. Extreme droughts were highly expected in region 1 and 3 for future periods.

The availability of water in India depends primarily on the precipitation during the monsoon season (June-September) where 70% of the annual rainfall occurs; therefore existence, movement and distribution of droughts are highly dependent on monsoon rainfall. Evaporation

rate is expected to increase due to global warming which results in the drier conditions on the ground and increase of water vapor in the atmosphere over time.

3.3 Drought Frequency Analysis

3.3.1 Severity-Duration-Frequency analysis

From Table 6, the distribution with minimum K-S and A-D values were selected for the frequency analysis. From Table 7, Clayton-copula was the best-fit copula for region 1; Gumbel-copula was the best-copula for region 2 and 3; and Frank-copula suits best for region 4 according to L-L and AIC values. To analyze and visualize the results, scatter plots between observed and randomly simulated severity and durations from each best fitted copula class were plotted and presented in Figure 7. Figure 8 represents joint probability plots between severity and duration for four homogeneous regions.

Further, best parameter values were analyzed and probabilities of duration were calibrated using inverse cumulative distribution of different univariate distributions fitted. Then joint probability dependence of various return periods was calculated by using inverse h-function of best-fit copula. For each homogenous region SDF curves were developed as shown in Figure 9. Region 2,3 and 4 display higher severity for different return periods suggesting that a higher frequency of drought can be expected in the regions under consideration. Moreover, the SDF curves were concave upward for all the regions in this analysis, specifying an increase in severity with an increase in duration. So, the occurrence of longer drought events in the twentieth-first century is more likely to occur.

3.3.2 Severity-Area-Frequency analysis

For various spatial extents, gamma distribution was selected as the best fit. Distribution of the parameters were calculated using L-moments method. Figure 10 shows SAF curves at various return periods (5, 10, 25, 50, 75, 100) and for average drought extents (2021-2025, 2061-2065 and 2095-2099). For regions 1-4; 1971-1976, 1984-1989, 1995-2000, and 2000-2005 respectively, were considered as most severe drought periods in comparison with projected SAF curves.

From Figure 10 it is observed that region 3 has a steeper slope compared to other regions that specifies drought risk is expected for small spatial extent. In region 1, 2 and 4 severity values are falling below 5-year return period during 2021-2025 and 2061-2065. For the period 2095-2099, return period varies between 10 to 25 years with increase in areal extents over region 1, 2, and 4.

For region 1, the 2095-2099 return period varies from 10-25 years; for region 2, associated return period for 2095-2099 is approximately 5 years for 55% of the area and decreases with increase in areal extent. In region 3, the associated severity values in 2095-2099 and 2061-2065 with areal extents are higher than 5-year increase up to 25-year return period. In region 4, the associated severity values are approximately 10 year for 45% of the area and decrease with increase areal extent. The most severe historical drought in the considered regions was observed in region 3 relative to the severity values for the 2095-2099 period.

4 Discussion

Closer analysis of the results shows that rising severity and duration of drought were noticed in almost all regions over all time-progressive regions; however, in region 3, mean severity decreases over the period 2040-2060 and rises over subsequent decades. The study of the occurrence of various droughts (moderate, serious and extremely severe) reveals that the frequency of droughts is likely to increase in most regions, with the exception of region 2. The SDF curves show that for most of the regions considered, the severity rate increases with duration. However, critical drought events are likely to occur over multiple time scales in four homogeneous regions. In addition, the derived SAF curves suggest that droughts are likely to cover a greater areal scale than other periods for most of the regions during 2091-2095.

Previous studies on projected changes observed due to Indian monsoon precipitation indicates a positive connection between global warming and Indian summer monsoon, a chance of increase in monsoon precipitation in India is observed for Representative Concentration Pathways (RCPs) 8.5 scenario than other RCPs (Freychet et al., 2015; Kundu et al., 2014; Menon et al., 2013; Chaturvedi et al., 2012). Previously analyzed SAF study reports that greater areal extent of drought was expected in latter part of century (Gupta and Jain 2018). Whereas steeper SAF curve slopes with high variability in topographical and hydrological characteristics have been observed for regions 2, 3 and 4. The SAF curves that have been developed can help to compare past droughts with future droughts. Therefore, it is noted from the results that India is experiencing severe drought in the latter part of the twenty-first century; that drought has a major impact on the population in these regions. Lack of soil moisture can also cause prolonged droughts, so turbulent heat flux and boundary layer distribution should be regulated because of surface energy loss (Sehgal et al. 2017; Alapaty et al. 1997). The local hydrological system depends primarily on soil moisture availability, absence of soil moisture disrupts agricultural production. In India, over 50 per cent of people depend primarily on agriculture. Historically, the share of agriculture in India in the total gross domestic product (GDP) has gradually

decreased from 39% in 1983 to 14% in 2014 (Ministry of Agriculture and Farmers' Welfare, 2016). Agriculture, therefore is economically and socially important for India's well-being. People's migration can be seen in almost all parts of the country due to lower agricultural profits and a lack of water availability. As per the results of the study, severe droughts are expected un the twenty-first century, leading to the shortage of availability of water in basins. Therefore, it effects the long-term damage of plant species and leads to desertification.

Also, an increase in population results in the increase of energy demand. Presently after China and United States, India is the third largest user of energy. So, demand for electricity is increasing which directly affects the availability of water in the river basins over future periods. Therefore, better preparedness, monitoring and prediction of drought can be considered as the best adaptation technique for mitigating the hazard due to future droughts. Policies should be framed in case of local and regional vulnerabilities for successful mitigation of drought risk induced by climate change.

It is also important to include the limitations and uncertainty in this study. The primary uncertainty is due to GCMs since GCMs resolution is still coarse. Moreover, individual models have their own prediction of uncertainty, and it often arises from the techniques used for downscaling. Moreover, there is also ambiguity in the projection of drought in various drought indexes and their selected probability distributions. However, advancing process-based assessment of hydrological variables and climatology reduces uncertainty and helps to determine possible drought scenarios with better confidence.

5 Conclusion

A regional study of the meteorological drought over India was carried out using the SPI-12 drought index, taking into account four distinct homogeneous drought regions. The run theory approach was used to determine the characteristics of drought, such as severity and duration. Copula-based methodology was adopted to derive SDF curves by analyzing the changes in the joint return period of severity and duration. SAF curves were developed which examine the changes in drought return periods covering a specific percentage of areal extent. A detailed spatiotemporal study of the occurrence, distribution, and frequency of drought was performed over India in the twenty-first century to quantify drought risk.

The key research outcomes are briefly described below:

- In the predicted future scenarios, the occurrence of moderate and extreme drought conditions has increased gradually in almost all regions, while major increases in severe drought are expected in regions 1 and 4.
- For the Godavari Basin, a high number of droughts with a high mean arrival time, high severity and duration are likely to be noticed in the future.
- It is observed from the severity-duration-frequency curves that region 2, 3, and 4 show higher severity for different return periods, indicating higher drought frequencies. Almost all SDF curves are concave upwards, indicating an increase in severity and duration.
- It is observed from the severity-area-frequency curves that for region 1, the return period varies between 10-25 years for 2095-2099. For region 2, associated return period for 2095-2099 is approximately 5 years for 55% of the area and decreases with increase in areal extent. In region 3, the associated severity values with areal extents are higher than 5-year and increases up to 25-year return period (2095-2099 and 2061-2065). In region 4, severity values are approximately 10 year for 45% of the area and decrease with increasing areal extent.

Acknowledgement

We acknowledge the partial support received by the corresponding author from the Virginia Agricultural Experiment Station (Blacksburg) and the Hatch Program of the National Institute of Food and Agriculture, U.S Department of Agriculture (Washington, D.C.).

References

Adarsh, S., Karthik, S., Shyma, M., Das Prem, G., Shirin Parveen, A. T. and Sruthi, N. (2018) Developing short term drought severity-duration-frequency curves for Kerala meteorological subdivision, India using bivariate copulas. *KSCE J. Civ. Eng*. 22 (3), 962–973.

Alapaty, K., Raman, S. and Niyogi, D. S. (1997) Uncertainty in the specification of surface characteristics: A study of prediction errors in the boundary layer. *Boundary Layer Meteorol*. 82 (3), 475–502. https://doi.org/10.1023/A:1017166907476.

Amirataee, B., Montaseri, M. and Rezaie, H. (2018) Regional analysis and derivation of copula-based drought severity-area-frequency curve in Lake Urmia basin, Iran. *J. Environ. Manage.* 206 (Jan), 134–144. https://doi.org/10.1016/j.jenvman.2017.10.027.

Araghi, A., Martinez, C. J., Adamowski, J. and Olesen, J. E. (2018) Spatiotemporal variations of aridity in Iran using high-resolution gridded data. *Int. J. Climatol*, 38, 2701–2717.

Bezdek, J.C. (1974) Numerical taxonomy with fuzzy sets. *J Math Biol* 1:57–71.

Bezdek, J.C. (1981) Pattern recognition with fuzzy objective function algorithms." *Plenum Press*, New York.

Bisht, D.S., Chatterjeea, C., Raghuwanshi, N.S. and Sridhar, V. (2017a) An analysis of precipitation climatology over Indian urban agglomeration. *Theor. Appl. Climatol*. doi: 10.1007/s00704-017-2200-z

Bisht, D.S., Chatterjee, C., Raghuwanshi, N.S. and Sridhar. V. (2017b) Spatio-temporal Trends of Rainfall across Indian River Basins. *Theor. Appl. Climatol.* doi: 10.1007/s00704-017-2095-8

Bisht, D.S., Sridhar, V., Mishra, A., Chatterjee, C. and Raghuwanshi, N. S. (2019) Drought characterization over India under projected climate scenario, *Int. J. Climatol*, 39, 1889-1911, doi: 10.1002/joc.5922

Bonaccorso, B., Peres, D. J., Castano, A. and Cancelliere, A. (2014) SPI-based probabilistic analysis of drought areal extent in Sicily. *Water Res. Manage*, 29, 459–470.

Burke, E. J. and Brown, S. J. (2010) Regional drought over the UK and changes in the future. *J. Hydrol*, 394(3), 471–485.

Cancelliere, A. and Salas, J.D. (2004) Drought length properties for periodic stochastic hydrological data. *Water Resour. Res*, 40. DOI: 10.1029/2002WR001750.

Chaturvedi, R.K., Joshi, J., Jayaraman, M., Bala, G. and Ravindranath, N.H. (2012) Multi-model climate change projections for India under representative concentration pathways. *Curr Sci*, 791–802

Chung, C.H. and Salas, J.D. (2000) Drought occurrences probabilities and risks of dependent hydrologic processes. *J. Hydrol. Eng*, 5(3), 259–268.

Dai, A. (2013) Increasing drought under global warming in observations and models. *Nat. Clim. Change*, 3 (1), 52–58.

Dalezios, N. R., Loukas, A., Vasiliades, L. and Liakopoulos, E. (2000) Severity-duration-frequency analysis of droughts and wet periods in Greece. *Hydrol. Sci. J.* 45 (5), 751–769. https://doi.org/10.1080/02626660009492375.

Das, P. K., Dutta, D., Sharma, J. R. and Dadhwal, V. K. (2016) Trends and behaviour of meteorological drought (1901–2008) over Indian region using standardized precipitation–evapotranspiration index. *Int. J. Climatol*. 36 (2), 909–916.

Dunn, J.C. (1973) A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *J Cybern*, 3(3), 32–57.

Fan, K., Liu, Y. and Chen, H. (2012) Improving the prediction of the East Asian summer monsoon: new approaches. *Weather Forecasting*, 27 (4), 1017–1030.

Freychet, N., Hsu, H.H., Chou, C. and Wu, C.H. (2015) Asian summer monsoon in CMIP5 projections: A link between the change in extreme precipitation and monsoon dynamics. *J. Clim*. 28 (4), 1477–1493. https://doi.org/10.1175/JCLI-D-14-00449.1.

Ganguli, P. (2014) Probabilistic analysis of extreme droughts in Southern Maharashtra using bivariate copulas. *J. Hydrol. Eng*, 20 (1), 90e101.

Genest, C. and Rivest, L. P. (1993) Statistical inference procedures for bivariate Archimedean copulas. *J Am Stat Assoc*, 88(423), 1034–1043

Genest, C., Ghoudi, K. and Rivest, L. P. (1995) A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3), 543–552.

González, J. and Valdés, J. B. (2003) Bivariate drought recurrence analysis using tree ring reconstructions. *J. Hydrol. Eng*, 8(5), 247–258.

Goyal, M.K., Gupta, V. and Eslamian, S. (2017) Hydrological Drought: Water Surface and Duration Curve Indices. Handbook of Drought and Water Scarcity: Principles of Drought and Water Scarcity.

Guhathakurta, P. and Rajeevan, M. (2008) Trends in the rainfall pattern over India. *Int. J. Climatol*, 28 (11), 1453–1470.

Gupta, V. and Jain, M. K. (2018) Investigation of multi-model spatiotemporal mesoscale drought projections over India under climate change scenario. J. *Hydrol.* 567 (Dec), 489–509. https://doi.org/10.1016/j.jhydrol.2018.10.012.

Gupta, V., Jain, M. K., Singh, V. P. (2020) Multivariate modeling of projected drought frequency and hazard over India. *J. Hydrol. Eng*, 25, 1–19.

Halwatura, D., Lechner, A. M. and Arnold, S. (2015) Drought severity-duration-frequency curves: A foundation for risk assessment and planning tool for ecosystem establishment in post-mining landscapes. *Hydrol. Earth Syst. Sci.* 19 (2), 1069–1091.

Hisdal, H. and Tallaksen, L. M. (2003) Estimation of regional meteorological and hydrological drought characteristics: A case study for Denmark. *J. Hydrol*, 281(3), 230–247.

Jain, V.K., Pandey, R.P., Jain, M.K. and Byun, H.R. (2015) Comparison of drought indices for appraisal of drought characteristics in the Ken River Basin. *Weather Clim. Extremes*, 8, 1–11.

Joe, H. (1997) Multivariate models and dependence concepts. *Monographs on statistics and applied probability*. Chapman and Hall, London, vol 73. p 399.

Kang, H. and Sridhar, V. (2017) Assessment of future drought conditions in the Chesapeake Bay watershed. *J. Am. Water Resour. Assoc*, doi: 10.1111/1752-1688.12600.

Kang, H. and Sridhar, V. (2018) Improved drought prediction using near real-time climate forecasts and simulated hydrologic conditions. *Sustainability*, 10, 1799, doi: 10.3390/su10061799.

Kang, H. and Sridhar, V. (2020) A novel approach combining simulated soil moisture and stochastic techniques to forecasting drought in the Mekong River basin. *Int. J. Climatol*, doi: 10.1002/joc.6860.

Kang, H., Sridhar, V., Mills, B., Hession, W. C. and Ogejo, J. A. (2019) Economy-wide climate change impacts on green water droughts based on the hydrologic simulations. *Agric. Syst*, Elsevier, 171, 76-88.

Karamouz, M., Nazif, S. and Falahi, M. (2013). Hydrology and hydroclimatology, principles and applications (1st ed.). Boca Raton, FL: CRC Press, 731 pp.

Kim, T. W., Valdés, J.B. and Yoo, C. (2006) Nonparametric approach for bivariate drought characterization using Palmer drought index. *J. Hydrol. Eng*, 11(2), 134–143.

Kim, T. W., Valdés, J. B. and Aparicio, J. (2002) Frequency and spatial characteristics of droughts in the Conchos River Basin, Mexico. *Water Int*, 27(3), 420–430.

Knutti, R. and Sedláček, J. (2013) Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Clim. Change*, 3 (4), 369–373. https://doi.org/10.1038/nclimate1716.

Kumar, P., Wiltshire, A., Mathison, C., Asharaf, S., Ahrens, B., Lucas-Picher, P. and Jacob, D. (2013) Downscaled climate change projections with uncertainty assessment over India using a high resolution multi-model approach. *Sci. Total Environ*, 468, S18–S30.

Kundu, A., Dwivedi, S. and Chandra, V. (2014) Precipitation trend analysis over eastern region of India using CMIP5 based climatic models. *Int. Arch. Photogramm. Remote Sens*. Spatial Inf. Sci. XL-8 (8), 1437–1442. https://doi.org/10.5194/isprsarchives-XL-8-1437-2014.

Lenderink, G., Buishand, A. and Van Deursen, W. (2007) Estimates of future discharges of the river Rhine using two scenario methodologies: Direct versus delta approach. *Hydrol. Earth Syst. Sci*, 11 (3), 1145–1159. https://doi.org/10.5194/hess-11-1145-2007.

Liu, X., Wang, S., Zhou, Y., Wang, F., Yang, G. and Liu, W. (2016) Spatial analysis of meteorological drought return periods in China using copulas. *Nat. Hazards*, 80 (1), 367e388.

Lobell, D. B., Burke, M. B., Tebaldi, C., Mastrandrea, M. D., Falcon, W. P. and Naylor, R. L. (2008). Prioritizing climate change adaptation needs for food security in 2030. *Science*, 319, 607–610.

Loukas, A. and Vasiliades, L. (2004) Probabilistic analysis of drought spatiotemporal characteristics in Thessaly region, Greece. *Nat. Hazards Earth Syst. Sci*, 4(5-6), 719–731.

633 Masud, M.B., Khaliq, M.N. and Wheater, H.S. (2017) Future changes to drought characteristics
634 over the Canadian Prairie Provinces based on NARCCAP multi-RCM ensemble. *Clim. Dyn*.
635 48 (7–8), 2685–2705.

636 Maxino, C.C., McAvaney, B.J., Pitman, A.J. and Perkins, S.E. (2008) Ranking the AR4 climate
637 models over the Murray-Darling basin using simulated maximum temperature, minimum
638 temperature and precipitation. *Int. J. Climatol*, 28, 1097–1112.

639 Mckee, T.B., Doesken, N.J. and Kleist, J. (1993) The relationship of drought frequency and
640 duration to time scales. AMS 8th Conf. *Appl. Climatol*, pp. 179–184 (citeulikearticle-
641 id:10490403).

642 Menon, A., Levermann, A., Schewe, J., Lehmann, J. and Frieler, K. (2013) Consistent increase
643 in Indian monsoon rainfall and its variability across CMIP-5 models. *Earth Syst. Dyn*, 4 (2),
644 287–300. https://doi.org/10.5194/esd-4-287-2013.

645 Ministry of Agriculture and Farmers' Welfare. (2016) State of Indian agriculture 2015–16.
646 New Delhi, India: Ministry of Agriculture and Farmers' Welfare.

647 Mishra, A. K. and Singh, V. P. (2009) Analysis of drought severity-area-frequency curves
648 using a general circulation model and scenario uncertainty. *J. Geophys. Res*, 114, D06120,
649 doi:10.1029/2008JD010986.

650 Mishra, A. K. and Singh, V. P. (2010) A review of drought concepts. *J. Hydrol*, 391 (1), 202–
651 216. https://doi.org/10.1016/j.jhydrol.2010.07.012.

652 Nadarajah, S. (2009) A bivariate pareto model for drought. *Stoch Environ Res Risk Assess*,
653 23(6), 811–822.

654 Nelsen, R.B. (2006) An introduction to copulas. Springer, New York

655 Pal, N.R. and Bezdek, J. C. (1995) On cluster validity for the fuzzy c-means model. *IEEE Trans*
656 *Fuzzy Syst*, 3(3), 370–379

657 Panjwani, S., Naresh Kumar, S., Ahuja, L. and Islam, A. (2019) Prioritization of global climate
658 models using fuzzy analytic hierarchy process and reliability index. *Theor. Appl. Climatol.* 137,
659 2381–2392.

660 Panofsky, H.A. and Brier, G.W. (1958) Some applications of statistics to meteorology.
661 Pennsylvania State University Press: University Park, PA; 224.

662 Perkins, S.E., Pitman, A.J., Holbrook, N.J. and McAveney, J. (2007) Evaluation of the AR4
663 climate models simulated daily maximum temperature, minimum temperature and
664 precipitation over Australia using probability density functions. *J. Clim*, 20, 4356–4376. doi:
665 10.1175/JCLI4253.1.

666 Rad, A. M., Ghahraman, B., Khalili, D., Ghahremani, Z., Ardakani, S. A. and Khalili, D. (2017)
667 Integrated meteorological and hydrological drought model: A management tool for proactive
668 water resources planning of semi-arid regions. *Adv. Water Resour.* 107 (Sep), 336–353.
669 https://doi.org/10.1016/j.advwatres.2017.07.007.

Rajeevan, M., Bhate, J. and Jaswal, A.K. (2008) Analysis of variability and trends of extreme rainfall events over India using 104 years of gridded daily rainfall data. *Geophysical Res. Lttrs*. 35, L18707. https://doi.org/10.1029/2008GL035143.

Rajsekhar, D., Singh, V. P. and Mishra, A. K. (2015) Integrated drought causality, hazard, and vulnerability assessment for future socioeconomic scenarios: An information theory perspective. *J. Geophys. Res.* 120 (13), 6346–6378. https://doi.org/10.1002/2014JD022670.

Reddy, M. J. and Ganguli, P. (2012) Application of copulas for derivation of drought severity-duration-frequency curves. *Hydrol. Processes*, 26(11), 1672–1685.

Reddy, M. J. and Ganguli, P. (2013) Spatio-temporal analysis and derivation of copula-based intensity–area–frequency curves for droughts in western Rajasthan (India). *Stoch Environ Res Risk Assess*, 27, 1975–1989.

Saghafian, B., Shokoohi, A. and Raziei, T. (2003) Drought spatial analysis and development of severity-duration-frequency curves for an arid region. In Proc., Int. Conf. on Hydrology of the Mediterranean and Semiarid Regions, 305–311. Wallingford, UK: IAHS Press.

Salvadori, G. and De Michele, C. (2015) Multivariate real-time assessment of droughts via copula-based multi-site Hazard Trajectories and Fans. *J. Hydrol*. 526, 101e115.

Samantaray, A. K., Singh, G., Ramadas, M. and Panda, R. K. (2019) Drought hotspot analysis and risk assessment using probabilistic drought monitoring and severity–duration–frequency analysis. *Hydrol. Processes*, 33 (3), 432–449. https://doi.org/10.1002/hyp.13337.

Santos, M. A. (1983) Regional droughts: A stochastic characterization. *J. Hydrol*, 66(1), 183–211.

Sehgal, V., Sridhar, V. and Tyagi, A. (2017). Stratified drought analysis using a stochastic ensemble of simulated and in- situ soil moisture observations. *J. Hydrol*, 545, 226-250, doi: 10.1016/j.jhydrol.2016.12.033

Sehgal, V. and Sridhar, V. (2019) Watershed-scale retrospective drought analysis and seasonal forecasting using multi-layer, high-resolution simulated soil moisture for Southeastern U.S. *Weather Clim. Extremes*, doi: 10.1016/j.wace.2018.100191

Sehgal, V., Lakhanpal, A., Maheswaran, R., Khosa, R. and Sridhar, V. (2018) Application of multi-scale wavelet entropy and multi-resolution Volterra models for climatic downscaling. *J. Hydrol*, 555, 1078-1095.

Shah, R.D. and Mishra, V. (2015) Development of an experimental near-real-time drought monitor for India. *J. Hydrometeorol.* 16 (1), 327–345.

Shiau, J.T. (2006) Fitting drought duration and severity with two dimensional copulas. *Water Resour Manag*, 20:795–815

Shiau, J.T. and Modarres, R. (2009) Copula-based drought severity-duration-frequency analysis in Iran. *Meteorol. Appl.* 16 (4), 481–489. https://doi.org/10.1002/met.145.

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

Shiau, J.T. and Shen, H. W. (2001) Resource analysis of hydrologic droughts of different severity. *J. Water Resour. Plan. Manag*, 127(1), 30–40.

Shivam, G., Goyal, M. K. and Sarma, A. K. (2019) Index-based study of future precipitation changes over subansiri river catchment under changing climate. *J. Environ. Inf*, 34 (1), 1–14. https://doi.org/10.3808/jei.201700376.

Sklar, A. (1959) Fonctions de re´partition a´ n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Universite´de Paris 8, 229–231

Sreelatha, K. and AnandRaj, P. (2019) Ranking of CMIP5-based global climate models using standard performance metrics for Telangana region in the southern part of India. *ISH J. Hydraulic Eng*, DOI: 10.1080/09715010.2019.1634648.

Sreelatha, K. and AnandRaj, P. (2020) Regional evaluation of global climate models for precipitation, maximum and minimum temperature over southern-part of India. *ISH J. Hydraulic Eng*, DOI: 10.1080/09715010.2020.1779137.

Tallaksen, L.M., Madsen, H. and Clausen, B. (1997) On the definition and modeling of stream drought duration and deficit volume. *Hydrol Sci J*, 42(1): 15–33.

Tase, N. (1976) Area-deficit-intensity characteristics of droughts. Doctoral dissertation, Colorado State Univ., Fort Collins.

Thilakarathne, M. and Sridhar, V. (2017) Characterization of future drought conditions in the Lower Mekong Basin. *Weather. Clim. Extremes*, 17, 47–58.

Tosunoglu, F. and Can, I. (2016) Application of copulas for regional bivariate frequency analysis of meteorological droughts in Turkey. *Nat. Hazards*, 82 (3), 1457e1477.

Urcid, G. and Ritter, G.X. (2012) Advances in Knowledge-Based and Intelligent Information and Engineering Systems. In: Graña, Manuel, Toro, Carlos, Posada, Jorge, Howlett, Robert J., Jain, Lakhmi C. (Eds.), Advances in Knowledge-Based and Intelligent Information and Engineering Systems. IOS Press, pp. 2140–2149.

Wang, L. and Chen, W. (2014) A CMIP5 multimodel projection of future temperature, precipitation, and climatological drought in China. *Int. J. Climatol*, 34 (6), 2059–2078.

Wilhite, D.A., Sivakumar, M.V. and Pulwarty, R. (2014) Managing drought risk in a changing climate: the role of national drought policy. *Weather Clim. Extremes*, 3, 4–13.

Wu, H., Svoboda, M.D., Hays, M.J., Wilhite, D.A. and Wen, F. (2007) Appropriate application of the standardized precipitation index in arid locations and dry seasons. *Int. J. Climatol*, 27, 65–79.

Xie, X.L. and Beni, G. (1991) A validity measure for fuzzy clustering. *IEEE Trans Pattern Anal Mach Intell*, 138, 841–847

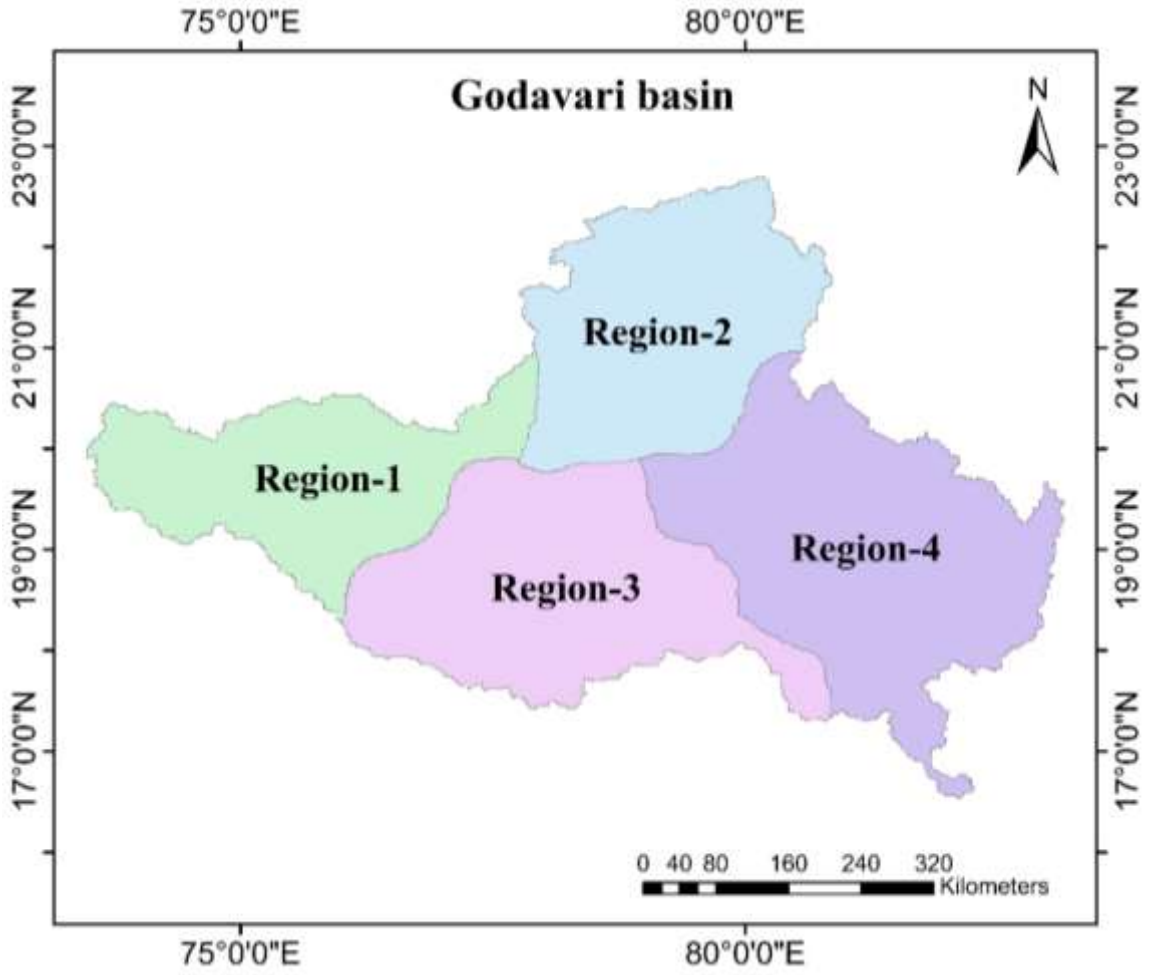
Xu, K., Yang, D., Yang, H., Li, Z., Qin, Y. and Shen, Y. (2015) Spatio-temporal variation of drought in China during 1961–2012: A climatic perspective. *J. Hydrol*. 526, 253–264.

Yevjevich, V. (1967) Objective approach to definitions and investigations of continental hydrologic droughts. *Hydrology Papers* No. 23. Rangely, CO: Colorado State Univ.

746 Zhang, Q., Singh, V.P., Li, J., Jiang, F. and Bai, Y. (2012) Spatio-temporal variations of
747 precipitation extremes in Xinjiang, China. *J. Hydrol*. 434, 7e18.

748

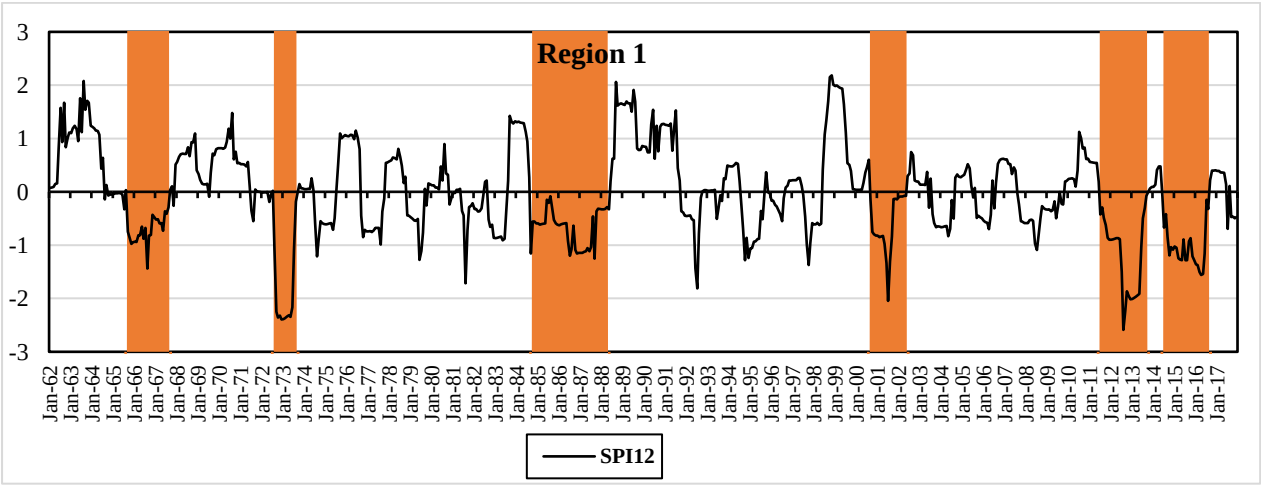
List of Figures



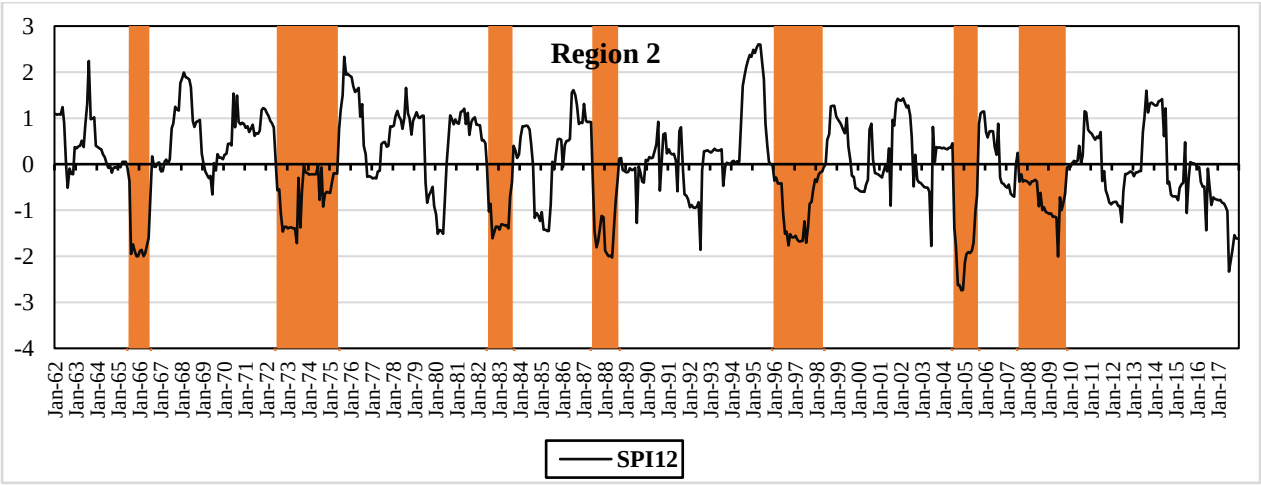
770

771 **Figure 1:** Clusters identified by fuzzy c-means clustering for precipitation series

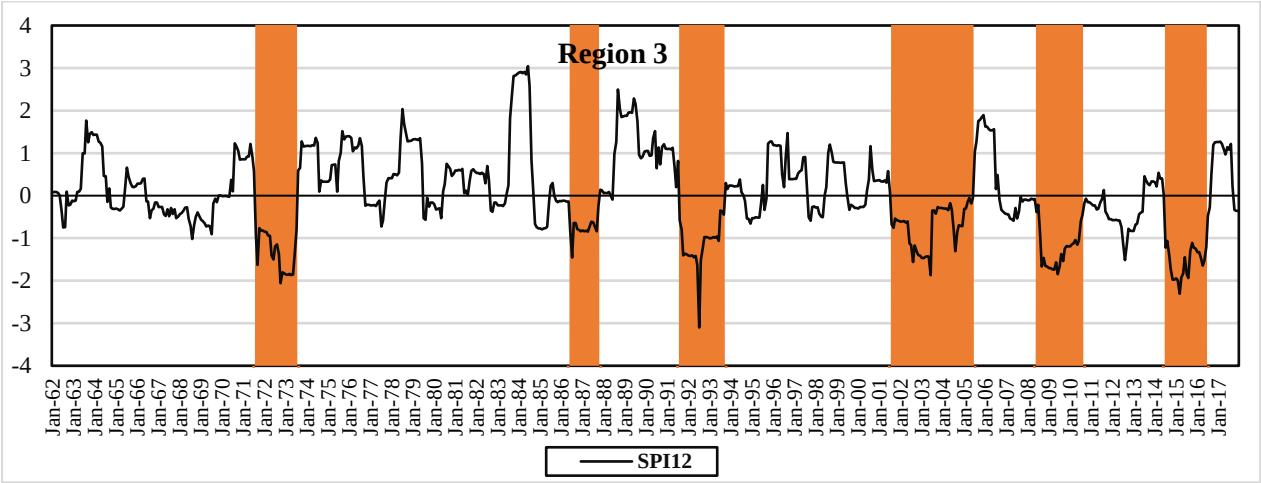
772



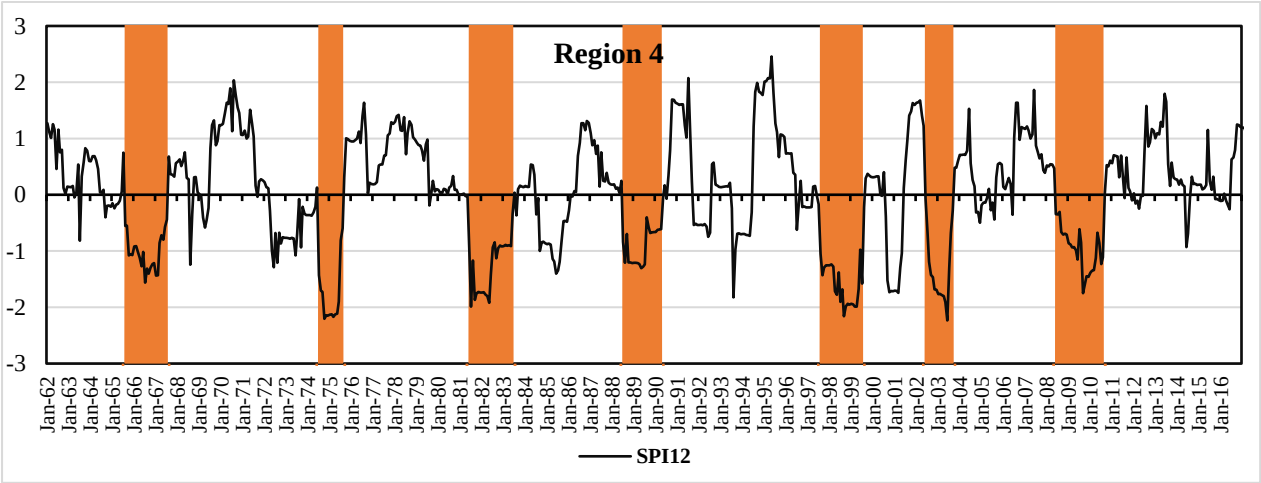
773



774

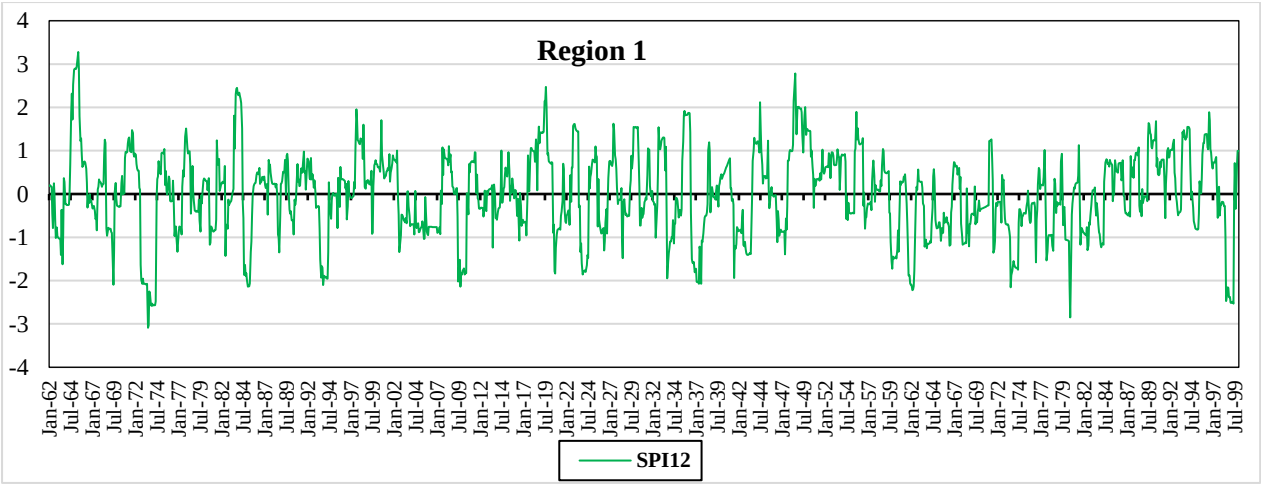


775



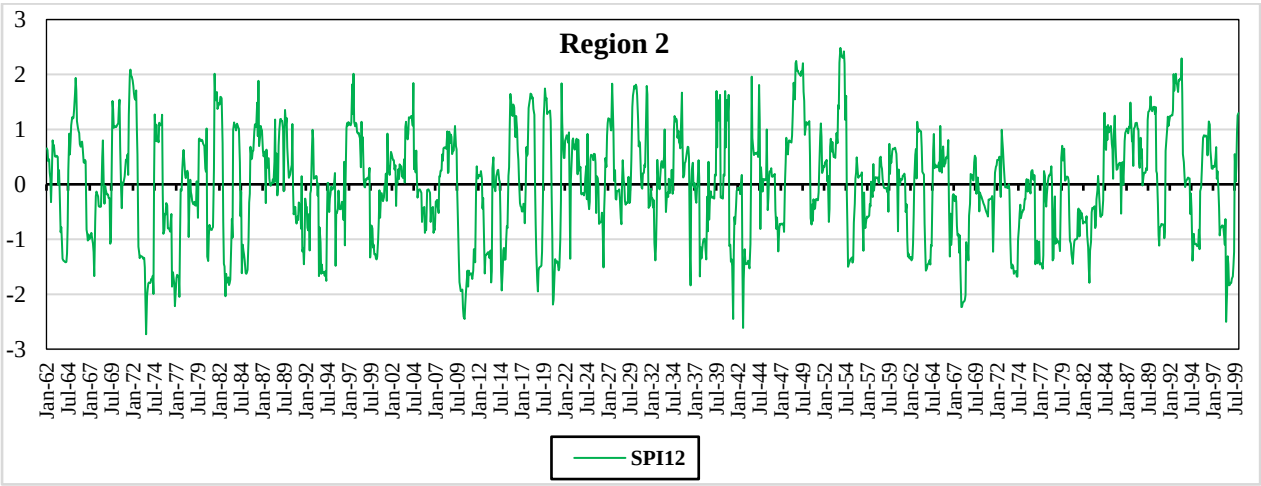
776

777 **Figure 2:** SPI-12 time series for four homogeneous drought regions over Godavari Basin for
778 the period of 1962–2017. The orange shaded area signifies the major drought events.



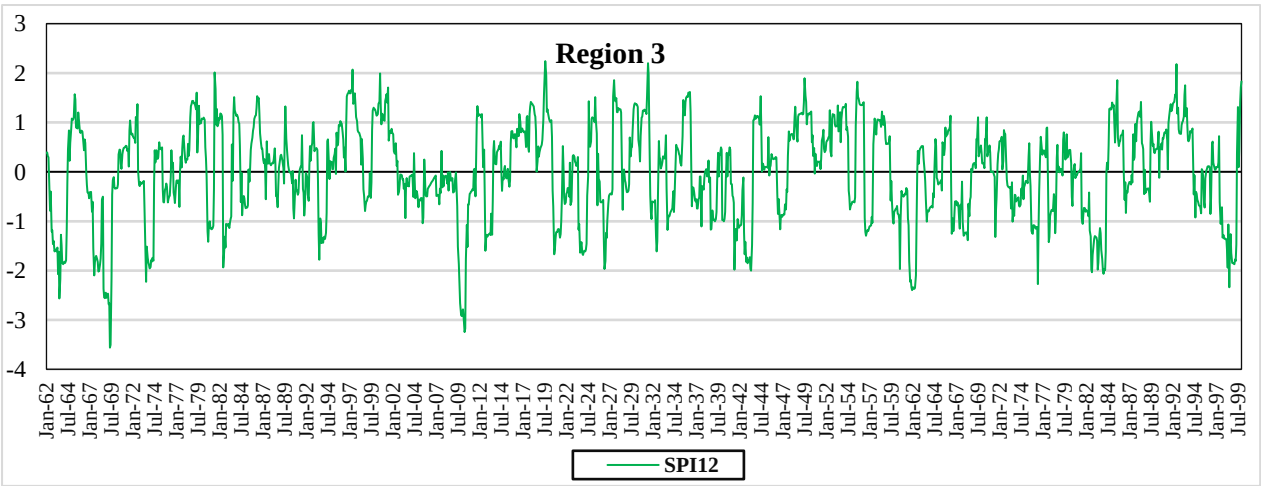
779

780



781

782



783

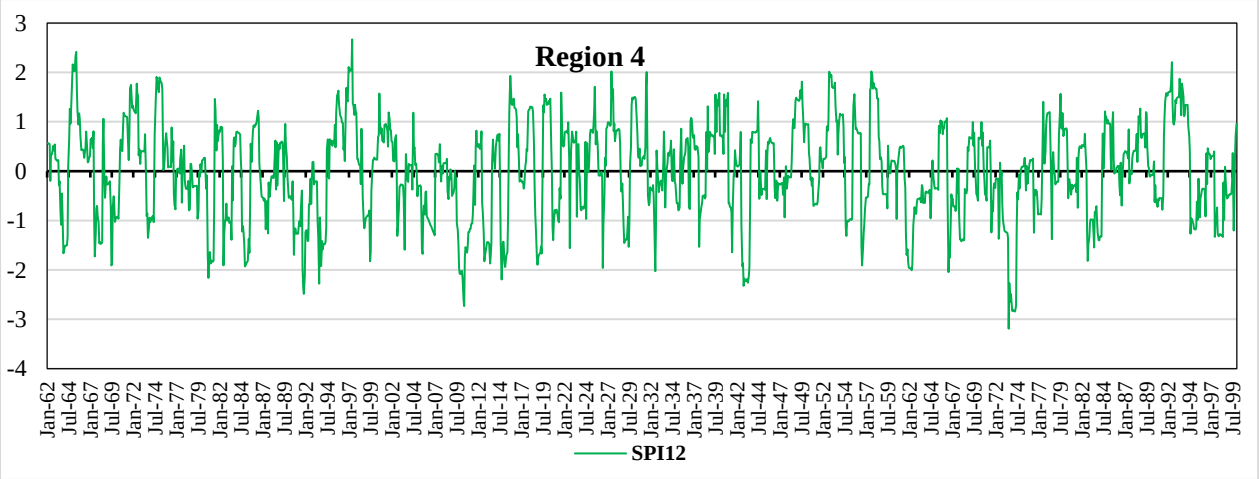


Figure 3: SPI-12 time series for four homogeneous drought regions over Godavari Basin for the GCM dataset (1962-2099).

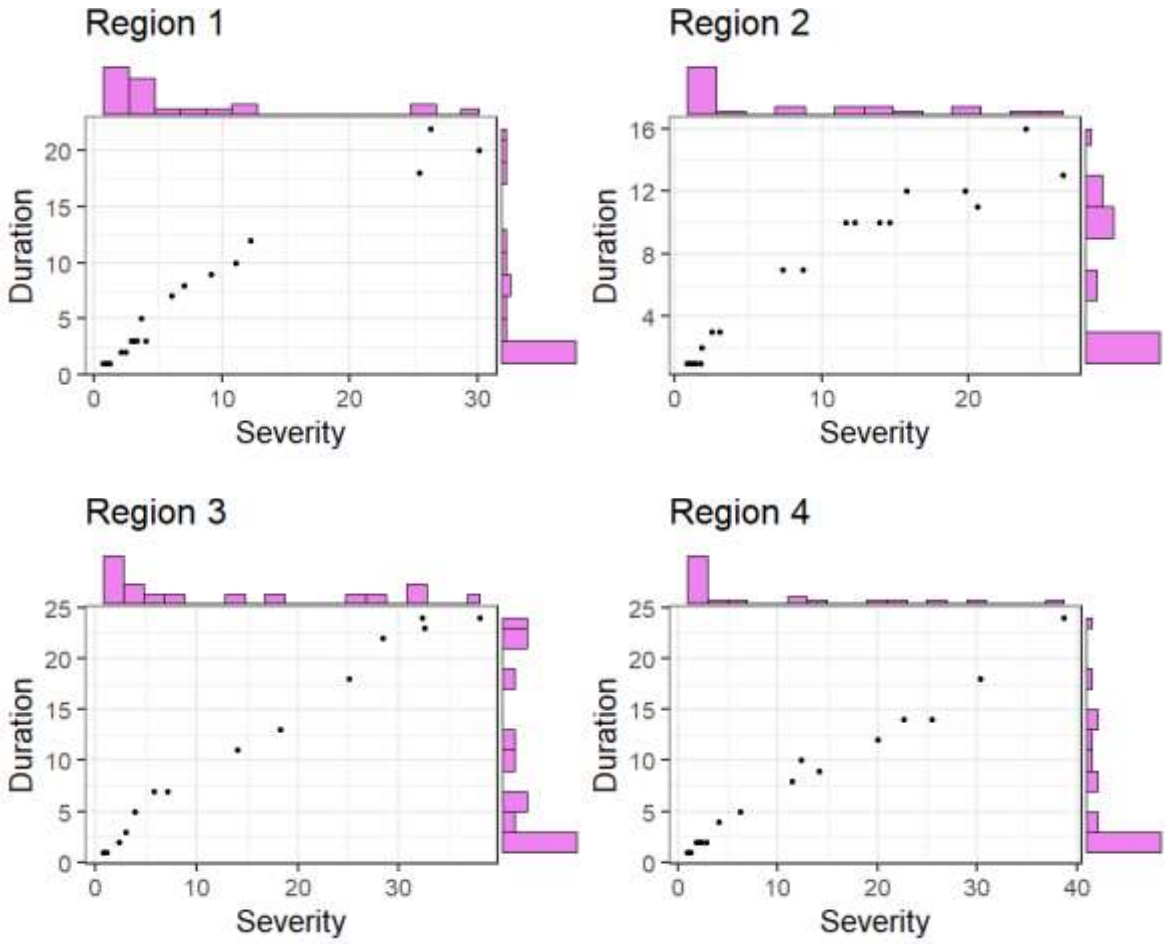


Figure 4: Scatterplot and histograms of drought severity and duration for four homogeneous drought regions over Godavari Basin.

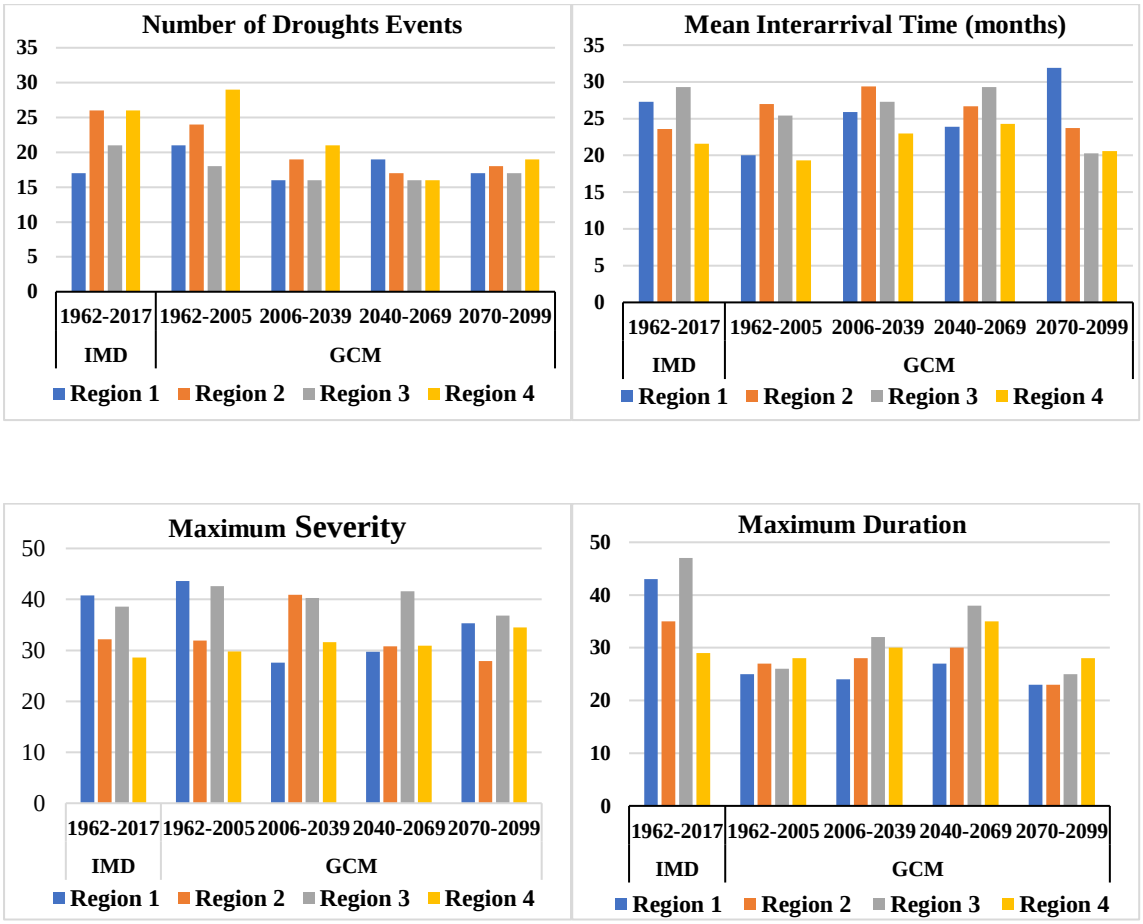


Figure 5: Drought characteristics for various homogeneous drought regions and time periods

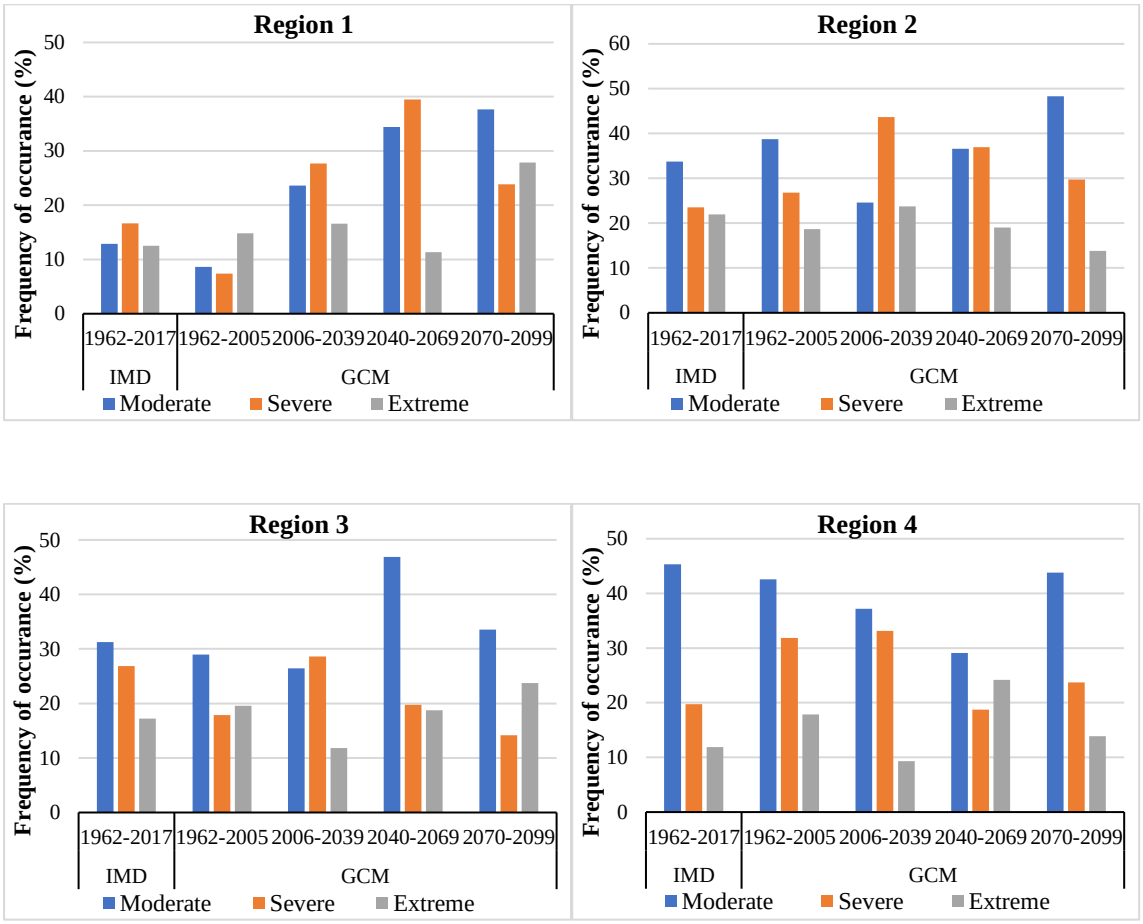


Figure 6: Frequency of occurrence of droughts over four homogeneous regions for IMD and GCM datasets

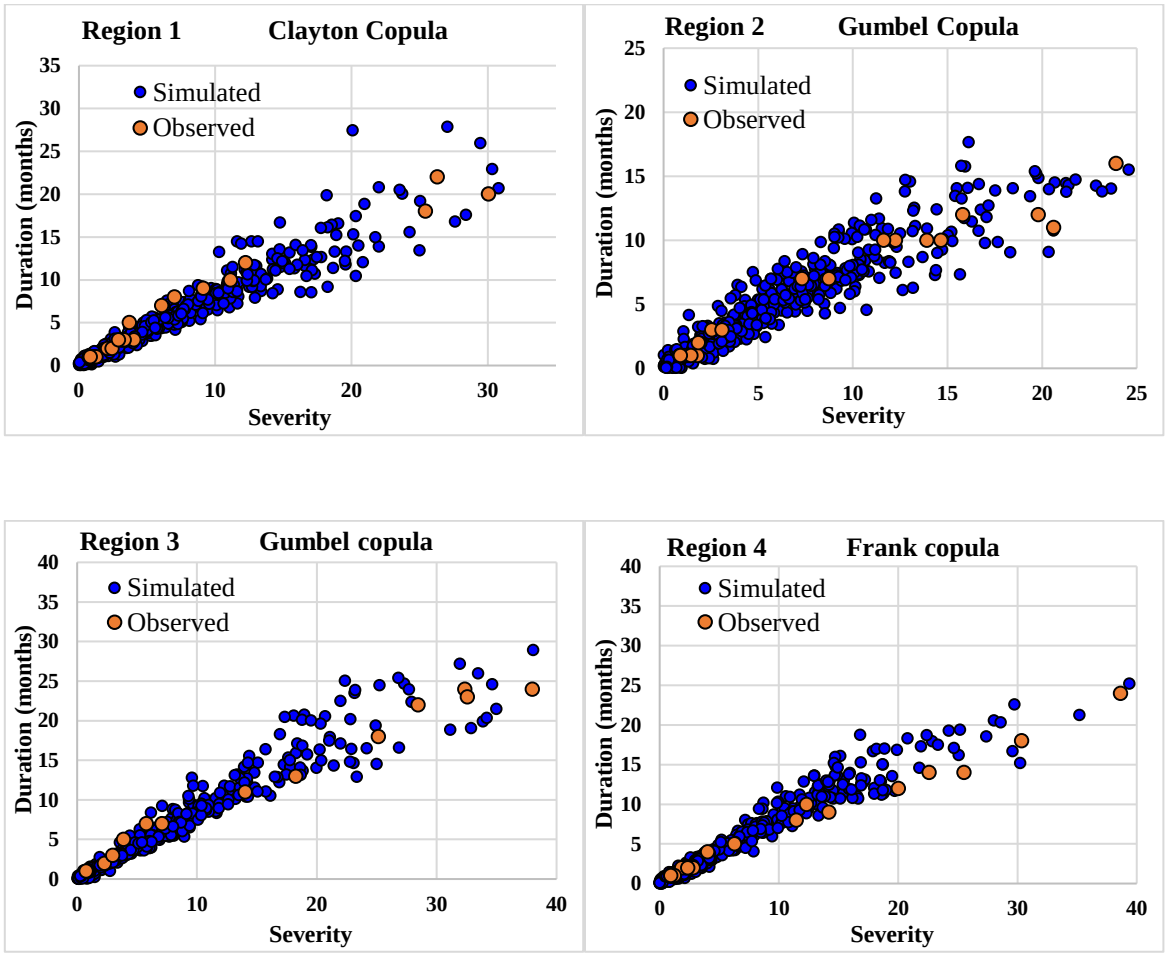


Figure 7: Scatterplots between observed and simulated severity duration values from best fit copula for each homogeneous region.

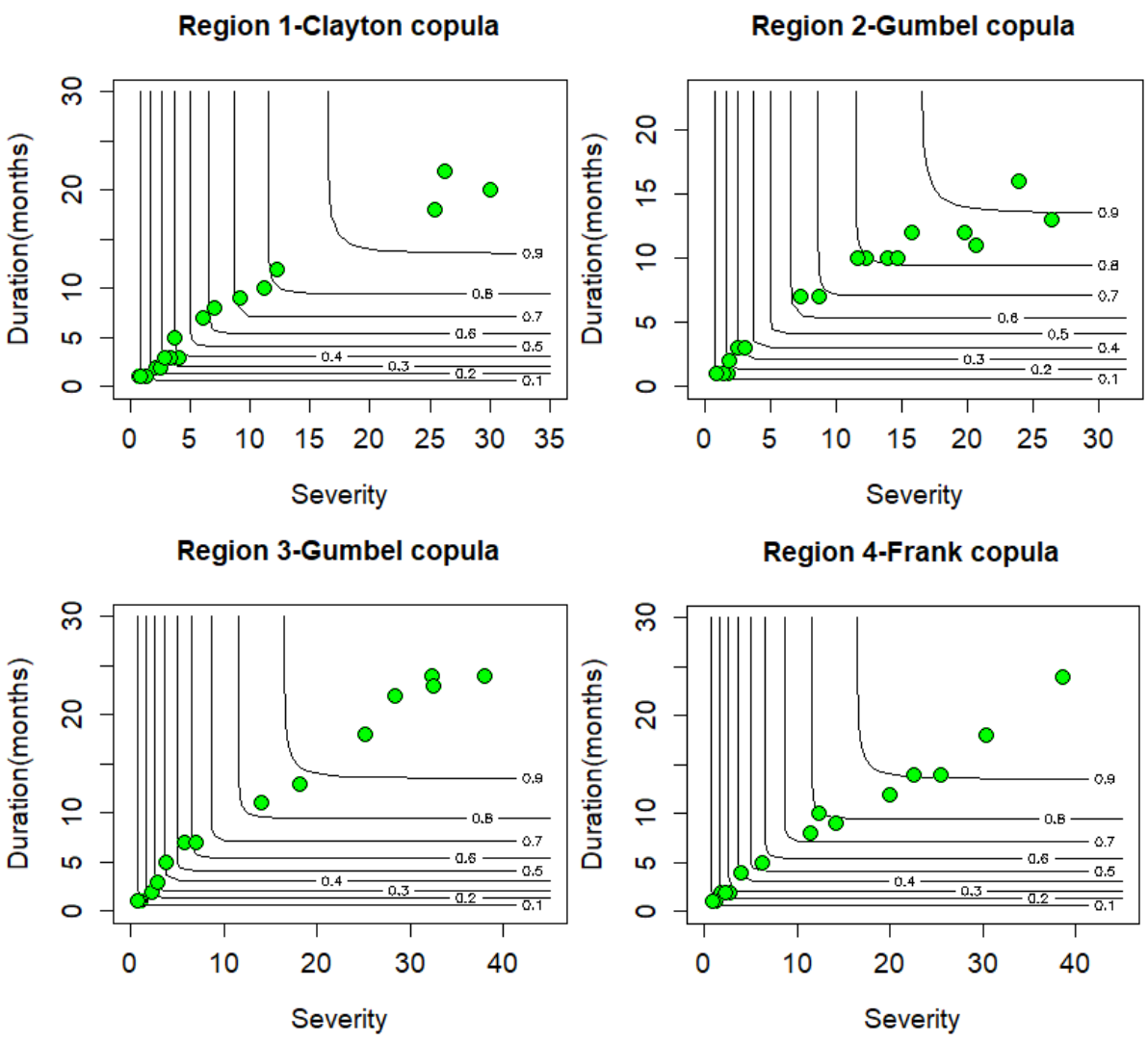
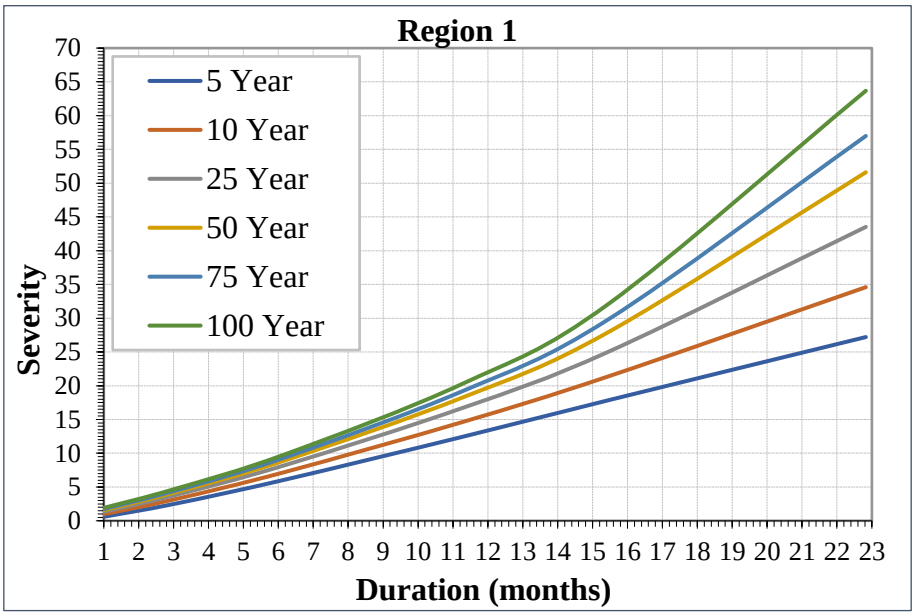


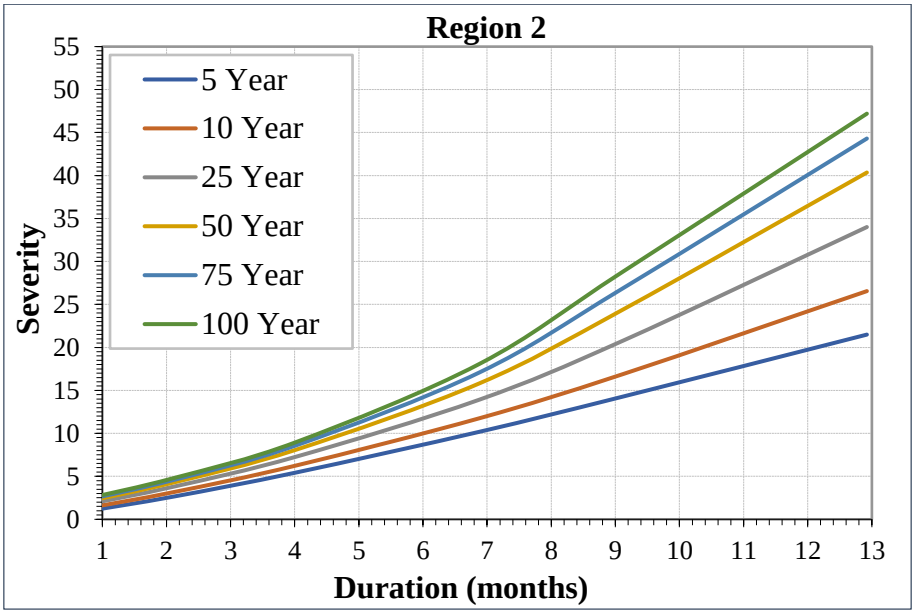
Figure 8: Joint probability plots between severity and duration for four homogeneous regions.

887

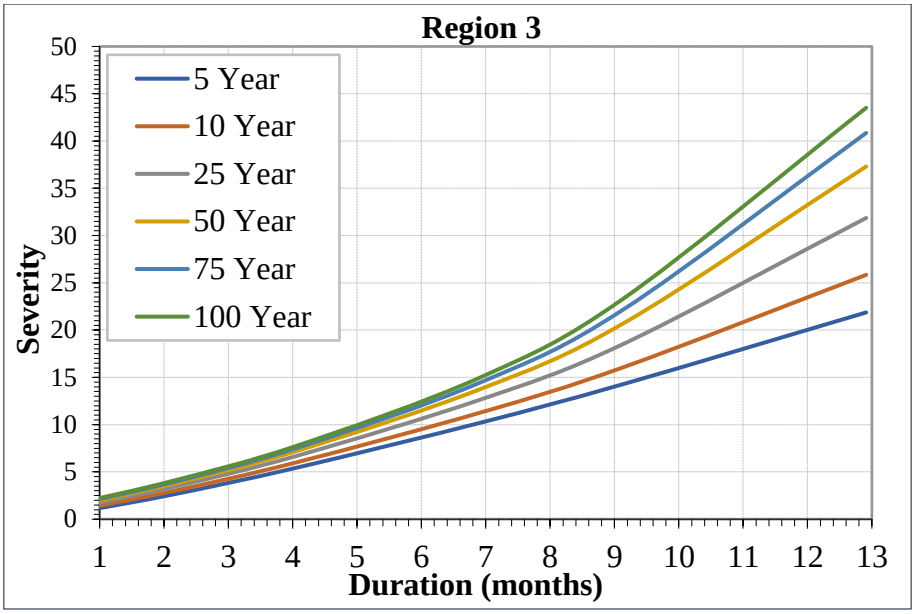


888

889



890



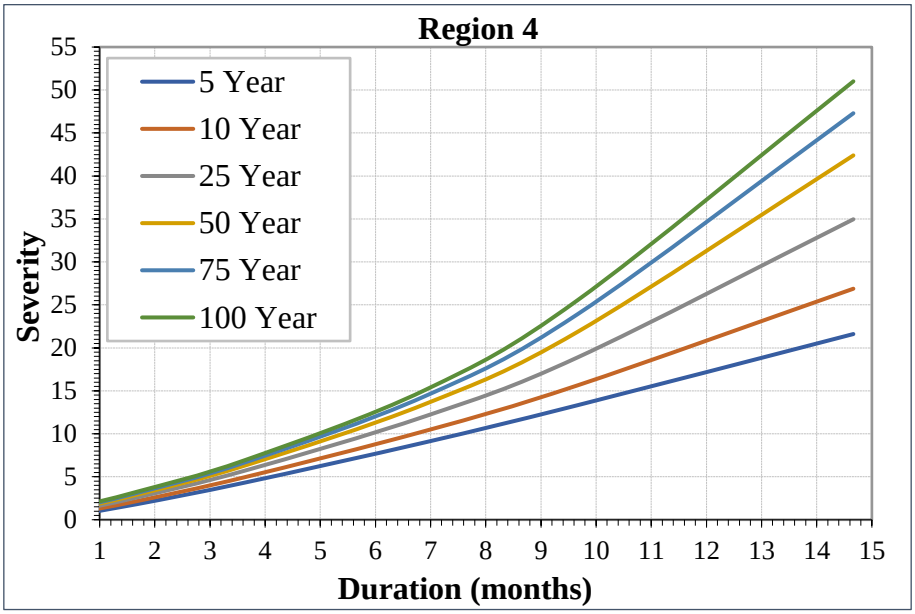
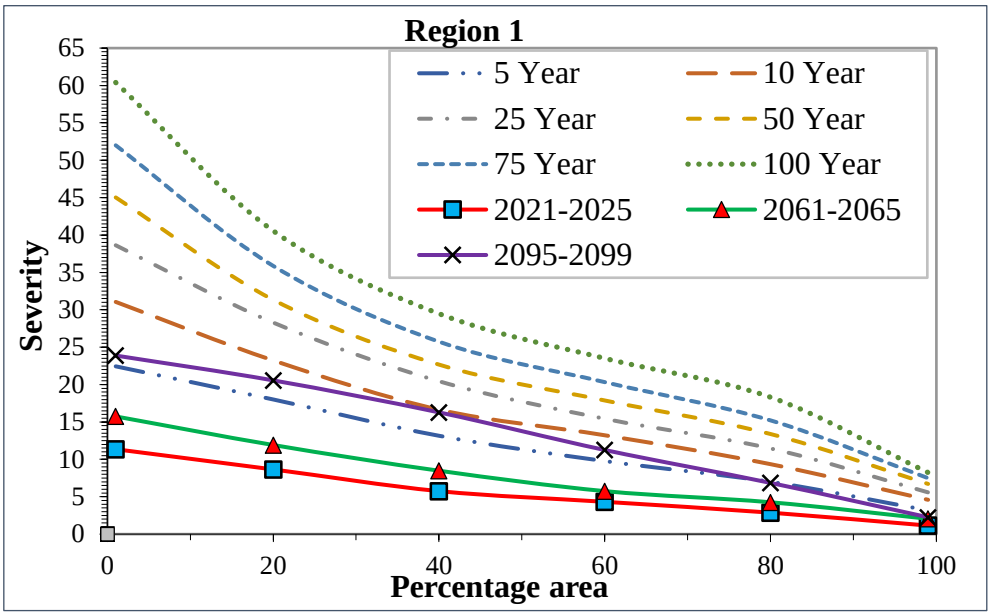


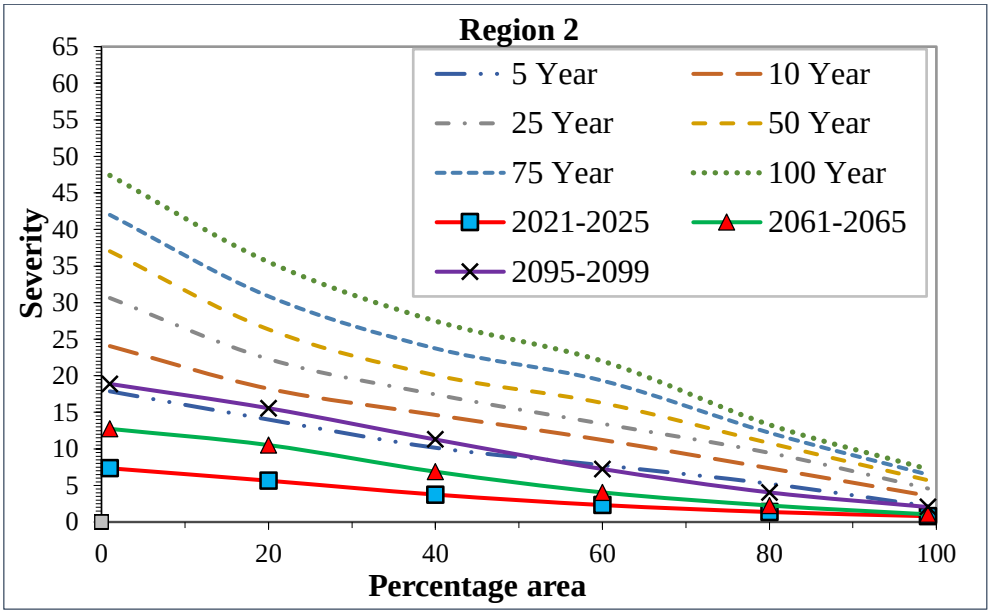
Figure 9: Drought SDF curves at various return periods for four homogeneous regions of the Godavari Basin

903

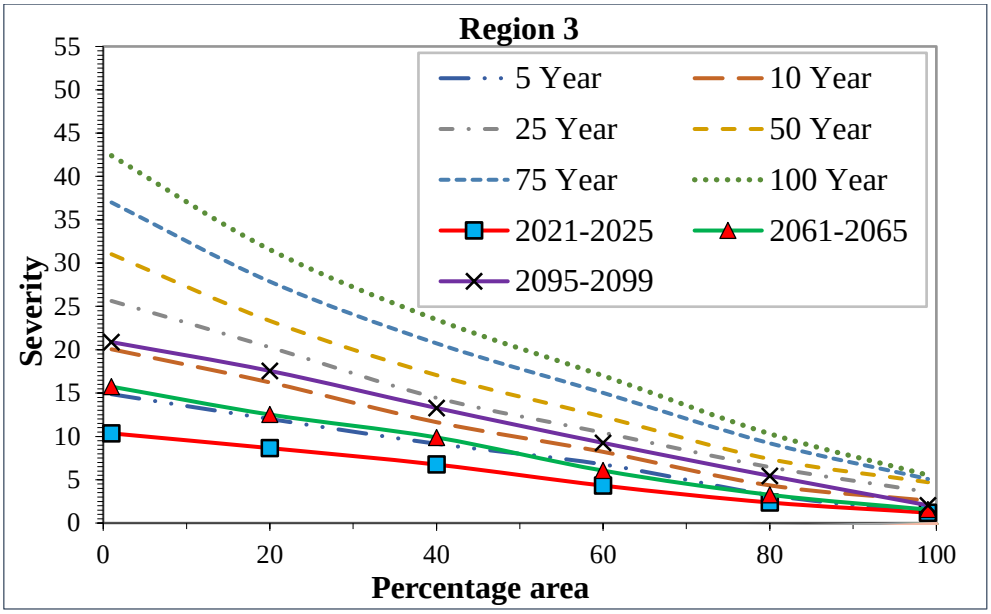


904

905



906



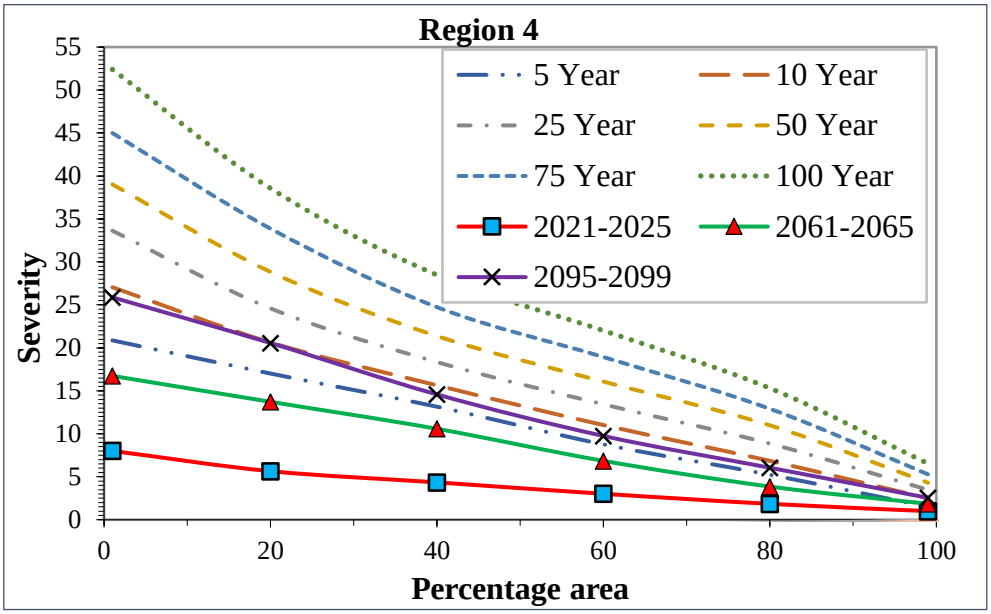


Figure 10 Drought SAF curves at various return periods for four homogeneous regions of the Godavari Basin

List of tables

Table 1: Expression for CDF ($C(u,v)$) and parameter space of the copula families.

Copula family	$C(u,v)$	Parameter space
Clayton	$(u^{-\theta}+v^{-\theta}-1)^{-\frac{1}{\theta}}$	$\theta \geq 0$
Frank	$-\frac{1}{\theta} ln\left[1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{(e^{-\theta}-1)}\right]$	$\theta \neq 1$
Gumbel	$\exp\left\{-\left[(-lnu)^{\theta}+(-lnv)^{\theta}\right]^{\frac{1}{\theta}}\right\}$	$\theta \geq 1$
Galambos	$uv \exp\left\{\left[(-lnu)^{-\theta}+(-lnv)^{-\theta}\right]^{-\frac{1}{\theta}}\right\}$	$\theta \geq 0$
Plackett	$\frac{1}{2(\theta-1)}(s-q)$	$\theta \geq 0$

Note: u and v represents two dependent CDFs, θ is the copula parameter,

$s = 1+(\theta-1)(u+v)$ and $q = \sqrt{s^2-4uv\theta(\theta-1)}$

Table 2: Drought categories related to dry (D) conditions for SPI (McKee et al., (1993))

Drought Category	SPI
D0: Mild (abnormal) Drought	0 to -0.99
D1: Moderate Drought	-1.00 to -1.49
D2: Severe Drought	-1.50 to -1.99
D3: Extreme Drought	≤ -2

Table 3: Statistics of the validity indices for precipitation series.

Clusters	Si	Fpi	Pe
2	0.67	0.83	0.87
3	0.57	0.77	0.78
4	0.35	0.74	0.73
5	0.39	0.72	0.74
6	0.41	0.76	0.77

Table 4: The severe drought events (top five) based on SPI12 (IMD precipitation data) for each drought homogeneous region (1962-2017).

Region	Starting month	Ending month	Severity	Duration (months)
Region 1	August 1971	August 1973	-26.42	13
	October 1984	April 1988	-39.23	43
	September 2000	May 2002	-18.56	21
	July 2011	September 2013	-35.39	27
	July 2014	July 2016	-28.61	26
Region 2	July 1972	May 1975	-32.21	35
	July 1982	August 1983	-17.78	14
	June 1987	August 1988	-20.77	15
	January 1996	April 1998	-29.45	28
	July 2004	August 2005	-26.93	14
Region 3	August 1971	July 1973	-32.34	24
	September 1991	September 1993	-29.43	25
	August 2001	June 2005	-40.29	47
	June 2008	August 2010	-33.95	27
	July 2014	June 2016	-37.96	24
Region 4	August 1965	July 1967	-24.81	24
	July 1974	August 1975	-25.29	14
	June 1981	June 1983	-32.17	25
	August 1997	July 1999	-38.63	24
	June 2008	August 2010	-26.32	27

Table 5: Drought properties for each homogeneous region based on IMD and GCM dataset.
[a]Interarrival time and duration of drought were measured in months.

Region	Drought characteristic	IMD	GCM			
		1962-2017	1962-2005	2006-2039	2040-2069	2070-2099
Region 1	No. of droughts	17	21	16	19	17
	Mean interarrival time[a]	27.3	20	25.9	23.9	31.9
	Mean severity	6.7	5.7	6.8	5.6	7.2
	Maximum severity	39.23	43.6	27.6	29.7	35.3
	Mean duration[a]	6.1	3.9	5	4.3	6.3
	Maximum duration[a]	43	25	24	27	23
Region 2	No. of droughts	26	24	19	17	18
	Mean interarrival time[a]	23.6	27	29.4	26.7	23.7
	Mean severity	8.1	7.9	9.34	8.2	7.3
	Maximum severity	32.2	31.9	40.9	30.8	27.9
	Mean duration[a]	5.6	5.5	7.1	5.8	5.4
	Maximum duration[a]	35	27	28	30	23
Region 3	No. of droughts	21	18	16	16	17
	Mean interarrival time[a]	29.3	25.4	27.3	29.3	20.3
	Mean severity	12.5	8.75	8.6	8.1	7.8
	Maximum severity	40.29	42.6	40.9	41.6	36.8
	Mean duration[a]	9.6	5.6	6.1	5.8	4.8
	Maximum duration[a]	47	26	32	38	25
Region 4	No. of droughts	26	29	21	16	19
	Mean interarrival time[a]	21.6	19.3	23	24.3	20.6
	Mean severity	8.4	5.6	7.1	7.9	5.9
	Maximum severity	38.63	29.8	31.6	30.9	34.5
	Mean duration[a]	6.8	3.9	4.7	5.8	4.1
	Maximum duration[a]	27	28	30	35	28

Table 6: Performance of different probability distributions for fitting the drought severity and duration (IMD-SPI12). Kolmogorov-Smirnov (K-S) and Anderson-Darling (A-D) are applied to test the goodness of fit.

Variable	Distribution	K-S	A-D	K-S	A-D	K-S	A-D	K-S	A-D
		Region 1		**Region 2**		**Region 3**		**Region 4**	
Drought severity	Exponential	0.22	1.19	0.25	1.93	0.23	1.57	0.28	2.32
	Normal	0.29	3.01	0.26	1.81	0.29	1.46	0.27	2.27
	Log normal	**0.12**	**0.48**	0.21	1.38	0.16	0.58	**0.17**	**1.15**
	Gamma	0.21	1.01	**0.19**	**1.33**	0.16	0.72	0.22	1.39
	Weibul	0.17	0.82	0.21	1.34	**0.16**	**0.66**	0.21	1.23
	Gumbel	0.25	2.06	0.25	1.96	0.24	1.43	0.28	2.31
Drought duration	Exponential	0.21	0.87	**0.24**	**1.93**	0.18	0.9	0.27	1.52
	Normal	0.29	2.06	0.24	1.98	0.25	1.28	0.29	2.02
	Log normal	0.22	0.91	0.28	2.26	0.18	0.72	0.22	1.38
	Gamma	0.23	1.01	0.26	2.07	0.17	0.73	0.27	1.46
	Weibul	**0.17**	**0.86**	0.27	2.06	**0.17**	**0.72**	**0.21**	**1.32**
	Gumbel	0.28	1.62	0.25	2.27	0.21	1.14	0.3	2.06

Table 7: The best fitted copula model with corresponding copula parameter (θ), log-likelihood (L-L) and AIC values (IMD-SPI12).

Copula	parameter	L-L	AIC	Parameter	L-L	AIC
	Region 1			**Region 2**		
Clayton	**6.38**	**17.51**	**-55.75**	5.67	17.24	-45.84
Frank	24.54	29.64	-65.27	18.9	20.59	-48.14
Gumbel	5.36	24.94	-59.88	**4.21**	**16.12**	**-42.21**
Galambos	4.85	22.38	-57.25	3.55	17.68	-43.65
Plackett	28.71	30.12	-64.01	18.45	20.15	-44.09
	Region 3			**Region 4**		
Clayton	5.63	15.07	-42.87	5.31	16.05	-37.77
Frank	17.15	17.58	-43.16	**13.79**	**15.83**	**-36.92**
Gumbel	**3.93**	**14.87**	**-37.25**	5.12	19.89	-39.52
Galambos	3.78	18.28	-39.32	4.77	20.22	-40.74
Plackett	20.35	22.18	-40.96	24.36	25.35	-42.15