1-1-2010

Drought Analysis under Climate Change by Application of Drought Indices and Copulas

Wen Yang
Portland State University

Drought Analysis under Climate Change

by Application of Drought Indices and Copulas

by

Wen Yang

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Civil and Environmental Engineering

Thesis Committee:
Hamid Moradkhani, Chair
Scott Wells
Yangdong Pan

Portland State University
2010

Abstract

Drought is a recurrent extreme climate event with tremendous hazard for every specter of natural environment and human lives. Drought analysis usually involves characterizing drought severity, duration and intensity. Similar to most of the hydrological problems, such characteristic variables are usually not independent. Copula, as a model of multivariate distribution, widely used in finance, actuarial analysis, has won increasingly popularity in hydrological study. Here, the study has two major focuses: (1) fit drought characteristics from Streamflow Drought Index (SDI) or Standardized Runoff Index (SRI) to appropriate copulas, then using fitted copulas to estimate conditional drought severity distribution and joint return periods for both historical time period 1920-2009 and future time period 2020-2090. SDI is calculated based on long term observed streamflow while SRI is based on simulated future runoff. Parameters estimation of marginal distribution and copulas are provided, with goodness fit measures as well; (2) investigate the effects of climate change on the frequency and severity of droughts. In order to quantify the impact, three drought indices have been proposed for this study to characterize the drought duration, severity and intensity changes under the climate change in Upper Klamath River Basin. Since drought can be defined as different types, such as meteorological drought, agricultural drought, hydrological drought and social economical drought, this study chooses Standardized Precipitation Index (SPI), Palmer Drought Severity Index (PDSI) and Surface Water Supply Index (SWSI) to estimate the meteorological, agricultural and hydrological

drought, respectively. Climate change effects come from three sources: the inherent reason, the human activity and the GCMs uncertainties. Therefore, the results show the long term drought condition by calculating yearly drought indices, and compared in three ways: First, compare drought characteristics of future time periods with base period; second, show the uncertainties of three greenhouse gas emission scenarios; third, present the uncertainties of six General Circulation Models (GCMs).

Acknowledgement

I would like to thank all the people who helped me, encouraged me and inspired me. This thesis would not have been possible without all your help and support.

Foremost, I owe my deepest gratitude to my dedicated advisor, Dr. Hamid Moradkhani for his support of my study and research, for his able guidance, wisdom and patience. He always puts high standards to our work, and dedicated his enthusiasm and profound knowledge to kindly guide us to achieve this goal during my research and thesis writing. I feel sincerely thankful to him being my advisor, who always believe in my potential and encourage me to do what I want to do. He is not only my academic advisor, but also a mentor for life.

Besides, I am extremely grateful to the rest of my thesis committee: Dr. Scott Wells and Dr. Yangdong Pan. Dr. Wells always present his class in a delicate balance between rigor and humor, tough but simply stated. I was delighted to interact with Dr. Yangdong Pan, because by attending his class, I get to learn a wonderful programming language and statistical tool, which plays an important part to fulfill my research and thesis work.

Also, I thank all my lab buddies at the Water Resources and Remote Sensing Lab. Thanks to Sue Wherry, Shahrbanou Madadgar, Reza Najafi, and Caleb Dechant, for that you all made our lab a nice and sweet place to work and study, for so kindly helping me editing my thesis, and for always pump me up.

I especially want to show my appreciation to my family for their care and unconditional love. I am indebted to my Mother, Hui Long, who always back up for me. I share all my joy and happiness moment with my Mom, and she knows how to cheer me up and helped me rebuild confidence when I get depressed. Lots of gratitude also goes to my Grandma, who always keeps a sharp mind, read a lot of books, and she is so thoughtful to us, most importantly, she respects to my decisions and helped me to pursue my dream.

Last but not least, thanks to my American host family, The Neufelds', for giving me great memories of my studying abroad life. Special thanks should go to my dear friends: Wen Fang, Jiali Ju, Ziyang Liu, Chengyu Dai and Chengyu Xin, I never feel bored because of your companionship.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1. Drought Types

Drought is a recurrent extreme climate event with tremendous hazard for every aspecter of natural environment and human lives. Comparing to other natural disasters, the consequence of drought is even more severe and costly. The U.S. Federal Emergency Management Agency (FEMA 1995) has estimated that drought costs the U.S. $ 6–8 billion annually. As for the worst recent event, Riebsame estimate the total cost of the 1988 drought at $39.2 billion (Riebsame et al. 1991).

Although there is not a universal definition of drought, in the most general sense, drought can be defined with different disciplinary perspectives, namely, meteorological drought, agricultural drought, hydrological drought and socioeconomic drought (National Drought Mitigation Center). While instead of independent from each other, different types of drought are closely related and interacted with each other. As Dingman (Dingman 1994) indicated, drought originates from a deficiency of precipitation over an prolonged time period, often but not always, accompanied by unusually high temperature, high winds, low humidity, and high solar radiation which result in increased evapotranspiration, known as meteorological drought. These situations would produce extended periods of abnormal low soil moisture, and then affect agriculture and natural plant growth, known as agricultural drought. Deficit of

precipitation might result in stream flow, lake, wetland, reservoir levels and water-table elevations declining to unusually low levels, which then is defined as hydrological drought. Low streamflow could cause reduced hydroelectric power generation, which would lead to the socioeconomic drought.

1.2. Drought Mitigation

Unlike tornado, earthquake or other natural disasters, drought usually evolve slowly in time, which makes it possible to build an effective drought mitigation plan depending on appropriate drought forecasting and monitoring systems (Cancelliere et al. 2007).

1.2.1 Drought Forecasting

According to National Integrated Drought Information System (NIDIS) Implementation Plan, drought forecasting mainly refers to two basic approaches: First is the prediction of hydrological conditions and second is the prediction of drought indices.

Prediction of hydrological conditions usually involves climate prediction and streamflow forecasting. Climate prediction methods include statistical methods which train a statistical model using available data, and dynamical methods which numerically solve the physical equations governing the climate system. Current seasonal streamflow forecasting system is based on Ensemble Streamflow Prediction (ESP)(Day 1985) by using the Variable Infiltration Capacity (VIC) hydrologic model. According to the sources

of the surface forcing data, the streamflow forecasting approach is sub-divided to three branches: ESP-based method, Climate Forecasting System (CFS)-based method and Climate Prediction Center (CPC) based method. The ESP-based method uses meteorological forcing taken from previous years, beginning on the same day that the forecast is initialized. The CFS-based method uses meteorological forecasts from the NCEP Climate Forecast System (CFS) ensemble dynamical model prediction. The CPC-based method uses meteorological forecasts from the NCEP Climate Prediction Center (CPC) Official Seasonal Outlook to drive VIC land model. The CFS is a fully coupled model representing the interaction between the Earth's oceans, land and atmosphere, and Saha et al. (2006) stated that this ocean-land-atmosphere dynamical seasonal prediction system, spatially and temporally complements the skill of the statistical methods used by the NCEP Climate Prediction Center (CPC). Other studies about prediction of hydrological conditions include that Gobena and Gan(2010) incorporate the seasonal climate forecasts in the ensemble streamflow prediction system by downscaling monthly numerical weather prediction (NWP) forecast outputs to station location and then disaggregating the monthly forecasts into daily input weather data to drive a hydrologic model. In addition, the North American Land Data Assimilation System (NLDAS) is also providing seasonal prediction products to support drought monitoring and seasonal drought prediction at CPC and NIDIS.

Prediction of drought indices analyzes duration features of dry and wet periods by statistical approaches to forecast future drought indices' series. Dupuis(2010) built a

statistical model of Palmer Drought Severity index series to get the dry period interarrival times, then estimated drought return periods. Cancelliere et al.(2007) proposed an analytical approach to find the transition probabilities of Standardized Precipitation Index (SPI), under the hypothesis that monthly precipitation aggregated at various time scales is uncorrelated and normally distributed. ARIMA, a linear stochastic model, and multiplicative Seasonal Autoregressive Integrated Moving Average (SARIMA) models were applied to Kansabati river basin to forecast droughts up to 2 months of lead-time(Mishra and Desai 2005). Other studies have been done by applying nonlinear artificial neural network, for example, Karamouz et al.(2009) developed an hybrid index (HDI) by combing SPI, Surface Water Supply Index (SWSI) and Palmer Drought Severity Index (PDSI) and then used artificial neural networks to predict HDI values.

1.2.2 Drought Characteristics Analysis

It is hard to identify the onset and the end of a drought event, but with tools like drought indices one can monitor the drought monitoring and analyze the drought characteristics. Drought characteristics mainly refer to duration, severity and intensity. Duration is the length of period which the index values are less than truncation level, and is selected by the analyst; severity is the cumulative index value based on the duration time; and the Intensity, sometimes known as magnitude, is defined as severity divided by duration(Dingman 1994; Shiau 2006). Another important characteristic of drought is interarrival time, which is defined as the time between the beginning of one

drought and the beginning of the next(Shiau 2006). Figure 1 explains the drought duration, severity and interarrival time.

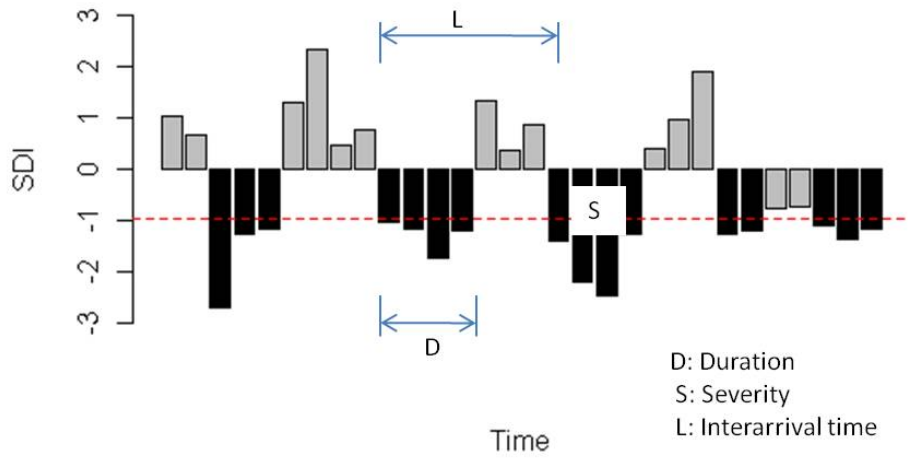


Figure 1: Characterization of drought events: SDI is streamflow drought index.

Analysis of drought characteristics has been studied extensively. Interarrival times of dry periods has been studied by statistical modeling of the monthly PDSI(Dupuis 2010), which models dry periods and wet periods as independent gamma distribution and proposed the exceedances over threshold approach to model the large interarrival times. In meteorological drought assessment, Pashiards and Michaelides (2008)employed the Standardized Precipitation Index (SPI) and the Reconnaissance Drought Index (RDI) by using historical monthly climate data for the identification of drought intensity and areal extent in Cyprus, and a similar study was conducted by Borg (2009) in Malta. In hydrological drought assessment, Dracup et al. (1980) used

truncation level of long-lead average to diagnose the dry and wet periods in historical time series of river flow and Smakhtin (2001)studied the frequency and severity of low flows. Based on the SPI developing concepts, Streamflow Drought Index (SDI) and Standardized Runoff Index (SRI) were applied for characterizing hydrological drought. SDI is defined based on cumulative streamflow volumes introduced by Nalbantis (2008), applied to two river basins in Greece. Instead of using observed streamflow, Shukla and Wood (2008) employed simulated runoff from semi-distributed Variable Infiltration Capacity (VIC) hydrological model to obtain the standardized runoff index. In addition, the links between severe hydrological droughts and weather types were also explored, and a new index-Regional Drought Area Index (RDSI), based on daily streamflow, was proposed to represent the drought affected area in north-western Europe (Fleig et al. 2010).

Considering the recent and potential future increases in global temperatures and changes to precipitation, it has been indicated that climate change may come with changes in the frequency and severity of extreme events such as droughts (IPCC 2007). Therefore drought analysis with respect to the effect of climate change is becoming needed and critical. Blenkinsop and Fowler (2007)utilize six regional climate models (RCMs) for assessing changes of drought frequency, severity and duration for British Isles. Their results show short-term summer drought is projected to increase and long-term drought would become less severe, although with high uncertainty. Loukas et al. (2008)evaluated climate change effects on drought severity in the region of Thessaly,

Greece by using SPI with one Circulation Model and two socioeconomic scenarios (A2 and B2). By combining multi-GCM model and multi-scenario, Sheffield and Wood (2008)analyzed changes in drought occurrence using soil moisture data for Scenarios A2, A1B and B1, while Ghosh and Mujumdar (2007)addressed the GCM and scenario uncertainty by nonparametric methods for determine the probability distribution function (PDF) of SPI. Climate change may also have favorable impacts on some areas. Cunderlik and Simonovic (2005) assesses the potential impact of climate change on the timing and magnitude of hydrological extremes in a southwestern Ontario river basin, concluded that future maximum river flows and low flows will be less extreme, more variable in terms of magnitude and more irregular in terms of seasonal occurrence.

1.3. Objective of This Study

Most of the studies mentioned above estimate drought return periods based on solely drought duration or severity, so the results could be misleading since those variables are typically correlated. Therefore, the first objective of this study is to employ copulas for analyzing drought characteristics. Copulas first introduced by Sklar (1959), can provide multivariate distribution with respect to margins' dependence structure. With this inherent advantage, copulas were broadly used in finance and actuarial science to model dependent mortality and losses, derivative pricing and risk management (Embrechts et al. 2003; Frees and Wang 2005; Yan 2006). Recently, copulas have garnered the attention of hydrologists in various applications such as flood and drought

analysis. (Salvadori and De Michele (2007) reviewed some advances of copulas in hydrological application. The first objective of this study is using copulas for drought analysis, density contours and conditional probability calculated from bivariate copulas; return periods calculated from single variable, bivariate and trivariate cases are also provided. The second objective is to use three widely-used indices of meteorological, agricultural and hydrological droughts, namely, Standardized Precipitation Index (SPI), Palmer Drought Severity Index (PDSI) and surface water supply index (SWSI), combining six General Circulation Models (GCM) and three scenarios (A2, A1B and B1, which represent high, medium and low future greenhouse emission level) for a comprehensive drought analysis at upper Klamath basin, OR, in the United States, specifically focus on the identification of drought duration, severity, intensity with respect to climate change.

Chapter 2: Copula

2.1. Copula Definition

A copula is a joint multivariate distribution in which the marginal distribution is uniform over (0, 1). Sklar's Theorem (Sklar 1959) showed that an n-dimensional Cumulative Distribution Function (CDF) H with univariate marginals F1, … , Fn, there exists an n-copula C such that

$$H(x_1, x_2, \dots, x_n) = C[F_1(x_1), F_2(x_2), \dots, F_n(x_n)] = C(u_1, \dots, u_n) \quad (2.1)$$

Where $u_1, \dots, u_n$ refer to the CDF of $x_1, \dots, x_n$. The most appealing advantage of copulas is that they provide a method to model multivariate distributions by separately dealing with copula marginals and the respective dependence structure. Copulas can be a good supplement of outcomes from regression analysis, which have been broadly used in hydrologic studies, but limited to independent predictors.

2.2. Copula Family

There are many families of copulas including elliptical (normal and t), Archimedean(Clayton, Gumbel, Frank, and Ali-Mikhail-Haq), extreme value(Gumbel, Husler-Reiss, Galambos, Tawn, and t-EV), and other families (Plackett and Farlie-

Gumbel-Morgenstern)(Yan et al. 2009). Archimedean and elliptical copulas are two most widely used copula families.

2.2.1 Archimedean copulas

Archimedean copulas have either symmetric or asymmetric forms. Symmetric copula refers to one-parameter copulas, expressed as:

$$C(u_1, \dots, u_n) = \varphi^{-1}[\varphi(u_1) + \cdots + \varphi(u_n)] \quad (2.2)$$

Where $\varphi(t)$ is the generator (see the definition for generator in Table 2.2.1)

$\varphi^{-1}(s)$ is the inverse of the generator φ.

Clayton, Frank and Gumbel copulas are commonly used one-parameter copulas, where their generator and inverse of generator functions are described in Table 2.2.1(Yan 2006).

Table 2.2.1: Constructors of three one-parameter (symmetric) Archimedean Copulas of three one-parameter (symmetric) Archimedean Copulas

Family	Generator $\varphi(t)$	Generator Inverse
Clayton	$t^{-\alpha} - 1$	$(1 + s)^{-1/\alpha}$
Frank	$-ln\frac{e^{-\alpha t} - 1}{e^{-\alpha} - 1}$	$-\alpha^{-1}\ln(1 + e^{-s}(e^{-\alpha} - 1))$
Gumbel	$(-\ln t)^{\alpha}$	$\exp(-s^{\frac{1}{\alpha}})$

Asymmetric copula refers to copulas having more than two parameters, expressed as:

$$C(u_1, \dots, u_n) = C_1\{u_n, C_2[u_{n-1}, \dots, C_{n-1}(u_2, u_1) \dots]\}$$

$$= \varphi_1^{-1}\left[\varphi_1(u_n) + \varphi_1\left(\varphi_2^{-1}(\varphi_2\{u_{n-1} + \cdots + \varphi_{n-1}^{-1}[\varphi_{n-1}(u_2) + \varphi_{n-1}(u_1)] \dots\})\right)\right] \quad (2.3)$$

By taking trivariate case as an example, we can write

$$C(u_1, u_2, u_3) = C_1[C_2(u_1, u_2), u_3]$$

$$= \varphi_1^{-1}(\varphi_1\{\varphi_2^{-1}[\varphi_2(u_1) + \varphi_2(u_2)] + \varphi_1(u_3)\}) \qquad (2.4)$$

Again, $\varphi(t)$ is copula generator and $\varphi^{-1}(s)$ is the inverse of generator φ.

Due to the positive results found in previous studies in applying copulas to drought analysis(Shiau 2006; Wong et al. 2010), the Gumbel copula for the Archimedean family is used in this study. Considering that the generator function and its inverse for Gumbel copula are known, the trivariate Gumbel copula can be written as

$$C_1[C_2(u_1, u_2), u_3] = \exp\{-[(-\ln\ u_1)^{\theta_2} + (-\ln\ u_2)^{\theta_2}]^{\frac{\theta_1}{\theta_2}} + (-\ln\ u_3)^{\theta_1}\}^{\frac{1}{\theta_1}} \quad (2.5)$$

It is worth noting that θ_2 represents dependence of the highest correlated pair (u_1, u_2).

2.2.2 Elliptical copulas

Elliptical copulas are copulas that correspond to elliptical distributions. Though this family does not have closed form expressions, the benefit of elliptical copulas, over the Archimedean family is they can specify correlation between each pair of marginals.

Normal copula and t-copula are the most commonly used elliptical copulas. While normal copula often used for modeling in finance, t-copulas have recently been used for modeling hydrological extremes, like flood and drought events, because of their ability to characterize tails of a distribution.

Because of the benefits of using the t-copula for hydrologic extremes, this study also utilizes t-copula for drought characteristics analysis. Multivariate t distribution is usually defined by its Probability Density Function (PDF), since it does not have an analytical expression for CDF. The procedure of constructing t-copula was illustrated by Wong (2010). Suppose we have d-dimensional random vector X, with mean vector μ , positive-definite correlation matrix R and degrees of freedom ν, then the PDF is

$$f(x) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\pi\nu)^d|R|}}\left(1 + \frac{(x-\mu)'R^{-1}(x-\mu)}{\nu}\right)^{-(\nu+d)/2} \quad (2.6)$$

Then according to Sklar theorem [Eq. (1)], the copula of X can be written as

$$C^t_{\nu,R}(u) = t^t_{\nu,R}[t_\nu^{-1}(u_1), \dots, t_\nu^{-1}(u_n)] \qquad (2.7)$$

Where, the parameters for trivariate t-copula include degree of freedom and a 3×3 correlation matrix, which determines the dependence structure.

2.3. Parameters' estimation

A two-stage approach, Inference Functions for Margins (IFM) method (Joe and Xu 1996), is used in the estimation of copula parameters for this study. Suppose we have p-

dimensional multivariate distribution with n observations for each margin, $\{(X_{i1}, \dots, X_{ip})^T : i = 1, \dots, n\}$.The first step of IFM is to use Maximum Likelihood Estimation (MLE) method to find the vector of marginal parameters β, which maximizes the likelihood function(Ricci 2005; Yan 2006):

$$logL(X_{ij}; \beta) = \sum_{i=1}^{n} \sum_{j=1}^{p} \log f_i(X_{ij}; \beta_j) \quad (2.8)$$

Where $f(\cdot)$ is the marginal PDF.

If the likelihood function is simple, we can easily set its partial derivates equal to zero, however the optimization procedure is usually conducted by iterative methods. Then the estimated $\widehat{\beta}_{IFM} = (\hat{\beta}_1^T, \dots, \hat{\beta}_p^T)^T$ (where, $\hat{\beta}_1^T, \dots, \hat{\beta}_p^T$ are transpose of $\hat{\beta}_1, \dots, \hat{\beta}_p$) from step 1 along with sample data, are used to estimate the copula parameters α, which maximizes the likelihood function:

$$logL(X_{ij}; \alpha, \beta) = \sum_{i=1}^{n} \log c\ (F_1(X_{i1};\ \hat{\beta}_1), \dots, F_p(X_{ip}; \hat{\beta}_p)) \quad (2.9)$$

Where $F(\cdot)$ is the marginal CDF

Again, iterative methods are applied to optimize the likelihood function to obtain the copula parameters $\hat{\alpha}_{IFM}$.

2.4. Goodness of fit test

The Kolmogorov-Smirnov test (ks test) is employed for this study to measure the goodness of fit for marginal distributions based on the null and alternative hypotheses(Ricci 2005):

H_0: Sample data come from the stated distribution

H_a: Sample data do not come from the stated distribution

The test compares the empirical distribution function (ECDF) with the theoretical CDF from the stated distribution. If the maximum absolute value difference between the ECDF and the theoretical CDF is greater than a given the critical value, which corresponds to a given significance level, we reject the null hypotheses. Conversely, if the p-value from ks test is greater than significance level, we accept our null hypothesis, meaning that the sample is of the stated distribution.

Chapter 3: Drought indices

3.1. What are drought indices?

Drought indices assimilate thousands of bits of data on rainfall, snowpack, streamflow and other water supply indicators into a comprehensive big picture(Hayes 2003). Drought indices are useful for monitoring drought conditions, because they provide a quantitative method for determining the onset and end of a drought event, and because the index value indicates the level of drought severity. It is difficult to say which drought index is the best, since some drought indices like SPI, which is based on precipitation provide good measurement for meteorological drought while other indices, like PDSI, that incorporate temperature in computations, emphasize soil moisture condition which better assesses agricultural drought. Still others like SWSI, take streamflow and reservoir data into account and are more appropriate for hydrological drought studies.

3.2. Drought indices used for this work

Four drought indices have been used for this work, which are Streamflow Drought Index (SDI), Standardized Precipitation Index (SPI), Palmer Drought Severity Index (PDSI), and Surface Water Supply Index (SWSI). Drought characteristics calculated from SDI are fitted to copulas to obtain conditional probability and return periods (Chapter 4), while the latter three drought indices are used for climate change effects study which

compares two future time periods 2020-2050 and 2060-2090 with reference time period 1975-2005 (Chapter 5).

3.2.1 Standardized Precipitation Index (SPI)

The SPI was developed by McKee et al. (1993) to quantify the precipitation deficit for multiple time scales. These time scales reflect the impact of drought on the availability of the different water resources. The SPI has also been used by the National Drought Mitigation Center (NDMC) to monitor moisture supply conditions.

The key advantage of SPI is that it can be calculated for different time scales. Suppose we have a time series of monthly total precipitation P_t , the equations calculating the SPI for a chosen time scale m are as follows:

$$R_t = P_t + P_{t-1} + \cdots + P_{t-m+1} \qquad (3.1)$$

Where R_t is the aggregated precipitation for t-month time scale

By aggregating the precipitation for a desired period, each month has a new value determined from the previous m months. Next, a two-parameter Gamma distribution is fitted to the $\{R_t\}$ and the parameters are estimated by the MLE method discussed in Chapter 2.

$$g(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}, 0 \le x \le \infty, \alpha > 0, \beta > 0 \quad (3.2)$$

After obtaining the estimated shape parameter α and scale parameter β, we then perform a "Z-score" to get standardized $\{R_t\}$ data.

$$SPI_t = \frac{R_t - \overline{R_t}}{S_{R_t}} \qquad (3.3)$$

Where $\overline{R_t} = \alpha\beta$ is the mean estimated from the fitted gamma distribution;

$S_{R_t} = \sqrt{\alpha\beta^2}$ is the standard deviation estimated from the fitted gamma distribution.

Therefore, SPI is normally distributed and can also be applied to wet periods in order to quantify precipitation surplus. As indicated in Table 3.2.1, moderate drought or worse, which is the focus of this study, corresponds to the SPI values of less than -0.99, associating with a cumulative probability of 16%. The SPI computation procedure is straightforward and a program is available online, by which the SPI for 3 month, 6 month, 12 month and 24 month rainfall totals can be obtained with only monthly precipitation input—relatively long time records of 30 years data are recommended.

For this study, SPI is calculated for the 12-month time scale, over the hydrological year: October –September. For example, the reference period is from Sep. 1975 to Oct. 2005, and the estimated drought type is meteorological drought.

Table 3.2.1: SPI Classification

SPI Values	Category	Time in Category	Cumulative Probability
-2.0 and less	extreme drought	2.3%	2.3%
-1.99 to -1.5	severe drought	4.4%	6.7%
-1.49 to -1.0	moderate drought	9.2%	15.9%
-0.99 to 0.99	near normal	68%	83.9%
1.0 to 1.49	moderate wet	9.2%	93.1%
1.5 to 1.99	severe wet	4.4%	97.5%
2.0 and more	extreme wet	2.3%	99.8%

3.2.2 Streamflow Drought Index (SDI)

As suggested by Mckee et al. (1993), the SPI procedure can also be applied to other water variables, such as soil moisture, snowpack, streamflow, reservoir and groundwater. The SDI developed by Nalbantis and Tsakiris (2009) and the Standardized Runoff Index (SRI) developed by (Shukla and Wood 2008), have computation procedures very similar to that of SPI. The difference between SDI and SRI is that the SDI uses observed streamflow data, while the SRI uses simulated runoff data from hydrological models.

According to the original SDI developers, (Nalbantis and Tsakiris 2009), the SDI calculation based on monthly observed streamflow volumes $V_{i,k}$ for a chosen reference period k of the i th year is as follows:

$$SDI_{i,k} = \frac{y_{i,k} - \overline{y_k}}{s_{y,k}} \quad i = 1,2,\dots, \quad k = 1,2,3,4 \quad (3.4)$$

$$y_{i,k} = \ln\left(V_{i,k}\right), i = 1,2,\dots, \quad k = 1,2,3,4 \quad (3.5)$$

Where $\overline{y_k}$ is the average of all $y_{i,k}$' values

$s_{y,k}$ is the standard deviation of all $y_{i,k}$' values

And k=1 for October-December, k=2 for October-March, k=3 for October-June, and k=4 for October-September. Noticing that every hydrological year has only one SDI value for a chosen k, those equations present above are suitable for annual SDI calculation when k=4, as for the monthly SDI, the procedure should follow the SPI computation method discussed above. Conversely, annual SPI calculation could also use the SDI developing equations (3.4) and (3.5) by replacing streamflow data with precipitation data.

While Mckee suggested that gamma distribution could be applied to streamflow, (Nalbantis 2008; Shukla and Wood 2008) state that log-normal distribution is a better choice for streamflow or runoff. In this work, we fit streamflow to both gamma and log-normal distribution and find that the goodness of fit is dependent on the choice of time scale. So for copula application, which will be discussed in detail in Chapter 4, cumulative streamflow data is fitted to both distributions for different time scales to select the best candidate distribution.

3.2.3 Palmer Drought Severity Index (PDSI)

The Palmer Drought Severity Index (PDSI) developed by Palmer (1965), which is based on the water balance concept and takes into account temperature, precipitation, soil moisture, runoff and other climate and hydrological properties, is a widely used measure of agricultural and general drought. The U.S. Department of Agriculture uses PDSI to determine when to grant emergency drought assistance(Hayes 2003).

Simply stated, PDSI uses temperature data to calculate potential evapotranspiration (PE) from the Thornthwaite method, and then uses precipitation and PE as inputs to compute basic hydrological cycle components, such as evapotranspiration, soil moisture, and runoff. This is, in essence, a soil moisture accounting algorithm. The computation involves two soil layers with the local Available Water Capacity (AWC) of each layer, however these AWC values are usually decided subjectively, which should be kept in mind when using them. According to the Soil Survey of Klamath County, the AWC can range from 3 inches to more than 13 inches. For example, the minimum AWC for the flood plain of Klamath River is 11 inches, the estimated AWC in Sprague sub-basin is 7 inches, and most of the irrigation land in Klamath Basin has AWC ranging from 9 inches to 13 inches. For this study, AWC of 10 inches was assumed for study area, with 1 inch assigned to the surface layer and 9 inches to the lower layers. The most important part for PDSI is the water balance routing procedure, and it can be illustrated as follows (Palmer 1965):

If PE is greater than precipitation (P) for the month, soil moisture loss (L) is assumed to occur:

$$L_s = \min\left(S_s', (PE - P)\right) \qquad (3.6)$$

The loss for surface layer is the minimum between available water capacity in this layer and the deficit of PE and P, and

$$L_u = (PE - P - L_s)\frac{S_u'}{AWC}, \quad L_u \leq S_u' \qquad (3.7)$$

Where L_s =moisture loss from surface layer,

S_s' = Available moisture stored in the surface layer at the beginning of the of month,

L_u =Loss from underlying levels,

S_u' =Available moisture stored in underlying levels at the beginning of the month, and

AWC =Combined available capacity of both levels

Evapotranspiration (ET) losses are the sum of P and L. Potential loss (PL) for the surface layer is defined as the minimum value between PE and AWC for this layer:

$$PL_s = \min(S_s', PE) \qquad (3.8)$$

For the underlying layers, the PL computation is similar to the L computation for this layer.

$$PL_u = (PE - PL_s)\frac{S'_u}{AWC} \qquad (3.9)$$

1) When P is greater than PE, there is no soil moisture loss; instead, recharge (R) is assumed to occur.

Potential recharge (PR) is defined as the amount of moisture required for the soil to reach its AWC:

$$PR = AWC - S' \qquad (3.10)$$

Where S'= the amount of available moisture in both layers of soil at the beginning of the month.

Runoff (RO) is assumed to occur when P is greater than AWC, and potential runoff (PRO) is defined as AWC minus PR:

$$PRO = AWC - PR = S' \qquad (3.11)$$

Therefore, the difference (d) between actual precipitation and Climatically Appropriate For Existing Conditions (CAFEC) is determined as follows:

$$d = P - (\alpha PE + \beta PR + \gamma PRO - \delta PL) \qquad (3.12)$$

Where α=ET_{avg}/PE_{avg} ,

β=R_{avg}/PR_{avg} ,

γ=RO_{avg}/PRO_{avg} ,

δ=L_{avg}/PL_{avg} .

The average values are the mean values of the month (from 1 to 12) for all calculated years.

A moisture anomaly index (Z) can be calculated as:

$$Z = Kd \qquad (3.13)$$

Where the climatic characteristic (K) is determined by:

$$K = \frac{17.67\ K_{avg}}{\sum_{i=1}^{12} D_i \times\ K_{avg}} \qquad (3.14)$$

$$K_{avg} = 1.5log_{10}\left(\frac{\left[\frac{PE_{avg}+R_{avg}+RO_{avg}}{P_{avg}+L_{avg}}\right]+2.8}{D_i}\right) + 0.5 \qquad (3.15)$$

Where D_i =the average of the absolute values of d for the month i

After all the Z values for each month have been calculated, we can compute the PDSI:

For the first month: $PDSI_{(1)} = \frac{1}{3}Z_{(1)}$ (3.16)

$$PDSI_{(j)} = 0.897 \times PDSI_{(j-1)} + \frac{1}{3}Z_{(j)} \qquad (3.17)$$

It is important to notice that until now all of the PDSI values have been monthly based. In order to get the yearly index, this study took the mean values of all 12 months for each year.

The Palmer drought classifications can be found in Table 3.2.2. The trigger level of drought events in this study is -1.99, corresponding to the moderate or worse drought situations.

Table 3.2.2: Drought Classification defined by PDSI

PDSI Values	Drought Class
4.0 or more	extremely wet
3.0 to 3.99	very wet
2.0 to2.99	moderately wet
1.0 to 1.99	slightly wet
0.5 to 0.99	incipient wet spell
0.49 to -0.49	near normal
-0.5 to -0.99	incipient dry spell
-1.0 to -1.99	mild drought
-2.0 to -2.99	moderate drought
-3.0 to -3.99	severe drought
-4.0 or less	extreme drought

The Thornthwaite Monthly Water Balance Model was developed by the U.S. Geological Survey (USGS), which can be used to compute the monthly water balance components of the hydrologic cycle by using the Thornthwaite method(McCabe and Markstrom 2007). This model provides Potential Evapotranspiration (PE) data to start the PDSI calculation, and serves as a future streamflow simulation tool. This model has seven input parameters: runoff factor, direct runoff factor, soil-moisture storage capacity,

latitude of location, rain temperature threshold (T_{rain}), snow temperature threshold (T_{snow}), and maximum snow-melt rate of snow storage. When the mean monthly temperature is above T_{rain}, all the precipitation is considered to be rain, and when the temperature is below T_{snow}, all the precipitation is considered to be snow. Direct runoff factor is the fraction of liquid precipitation, which directly becomes infiltration-excess overflow, while the maximum snow-melt rate is the fraction of snow storage that melts in a month (McCabe and Markstrom 2007). Soil-moisture storage capacity is set to 254 mm, in order to correspond to the total Available Water Capacity (10 inches) mentioned in the PDSI calculation. As recommended, T_{rain} =3.3°C, Runoff Factor=0.5, and Maximum Snow Melt Rate =0.5 work for most sites(McCabe and Wolock 1999; Wolock and McCabe 1999), whereas T_{snow} , and the Direct runoff factor need to be calibrated due to their sensitivity. Details for parameter selection are as follows (Table 3.2.3):

Table 3.2.3: Parameters of Thornthwaite Monthly Water Balance Model chosen for this study

	Calibrated at Reference Period	
Model Parameters		Note
Runoff Factor	0.5	recommended
Direct Runoff Factor	0.23	sensitive, calibrated
Soil-Moisture-Storage Capacity	254 mm	based on location
Maximum Snow Melt Rate	0.5	recommended
Latitude of Study Area	42 °	based on location
Rain Temperature	3.3 °C	recommended
Snow Temperature	-10 °C	sensitive, calibrated

Figure 2 shows that the accumulated 6-month scale simulated runoff from the Thornthwaite model shows a good agreement with the observed streamflow when the proposed parameters in Table 3.2-3 are applied to the validation period of 1975-2005. We can also see that the model underestimated the low flow during the earlier years, and overestimated some peak flows during the later years.

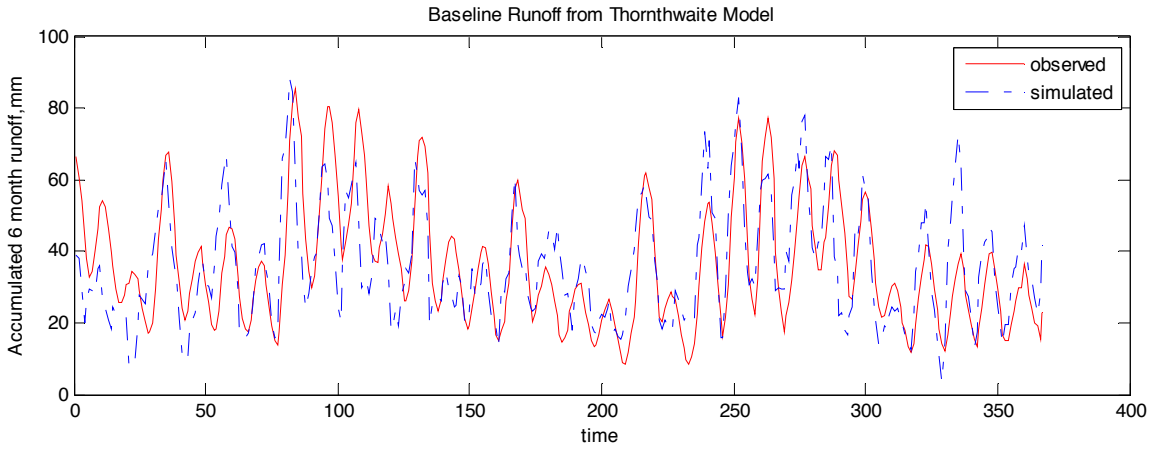


Figure 2: Accumulated 6-month runoff from observed streamflow and simulated runoff from Thornthwaite Model for validation period: 1975-2005

Statistical measurements, which include the Root Mean Square Error (RMSE) and Percent Bias for both monthly and 6-month scale, are summarized in Table 3.2.4. Since the both the monthly and 6-month scale percent bias are less than 2%, the model simulation is considered satisfactory.

Table 3.2.4: Thornthwaite Model calibrated for 1975-2005

	Model Performance	
Runoff	RMSE	Percent Bias

Monthly	5.23	0.68%
6 month	13.9	1%

3.2.4 Surface Water Supply Index (SWSI)

The SWSI originally designed by Shafer(Shafer and Dezman 1982), incorporated snowpack, streamflow, precipitation, and reservoir storage to complement the Palmer Index for moisture conditions across Colorado. The basic concept of SWSI is a rescaled weighted sum of the nonexceedance probabilities of snowpack, streamflow and reservoir storage and precipitation, expressed as:

$$SWSI = \frac{aP_{snow} + bP_{prec} + cP_{strm} + dP_{resv} - 50}{12} \qquad (3.18)$$

Where a, b, c and d are weights for each of the hydrological components, specifically, in winter a+b+d=1, and in summer b+c+d=1. Although it seems quite appealing to incorporate all the major water supply components into a single index, David C. Garen pointed out the original SWSI has two conceptual weaknesses: first, all the weights are usually obtained by subjective judgments, which makes the inter-basin less comparable and make it hard to optimize the weights; second, statistically, the weighted sum of the nonexceedance probabilities of each hydrological component is not equal to the joint nonexceedance probability of these variables.

In consequence, a revised SWSI formulation was proposed for better application to other states such as Oregon, Montana, Idaho and Utah(Garen 1993). Since the main

sources of water in the West are reservoir storage and streamflow, the modified SWSI is the sum of the nonexceedance probabilities of reservoir storage (Mar 31st) and streamflow (April-September). Therefore SWSI can be calculated with the following formula:

$$SWSI = \frac{P - 50}{12} \qquad (3.19)$$

The key step in determining the SWSI is a frequency analysis of the sum of streamflow and reservoir components. In this study, the streamflow and reservoir data for the reference period is observed data from NWS and USGS, which are volumetric data, with units in kaf. For future time periods, the simulation of runoff from Thornthwaite Water Balance model was treated as future streamflow. As for the future reservoir data, it was replaced by the mean reservoir storage over the reference period. The truncation value of drought events for SWSI is -1.99, corresponding to moderate or more severe droughts.

The ranges of SWSI values and their particular descriptors are similar to the Palmer index, as classified in Table 3.2.5(Garen 1993).

Table 3.2.5: SWSI classifications

SWSI Values	Classification
+2 or above	abundant supply
-2 to +2	near normal
-3 to -2	moderate drought

-4 to -3	severe drought
-4 or below	extreme drought

SWSI is designed to serve areas where are mountain water dependent and areas that snowmelt runoff is the main source of water supply(Shafer and Dezman 1982). Upper Klamath River Basin locates at Pacific Northwest, so it is appropriate to use SWSI for this region.

Some limitations of SWSI should be acknowledged. Hayes (2003) discussed that the discontinuance of any station means that new stations need to be added to the system and new frequency distributions need to be determined for that component. Although the revised SWSI overcome the subjective judgments of component weights, to some extent, it also masks important information about the behavior of each component(Garen 1993).

Chapter 4: Drought analysis by application of copula

In this chapter, copula techniques are applied to characterize drought duration, severity and intensity. First, Streamflow Drought Index (SDI) calculated for reference period (1906-2009) is used to derive drought duration, severity and intensity. Second, bivariate and trivariate Gumbel and/or t-copula were constructed to estimate severity conditional probability and drought return periods. Third, for future period (2020-2090), similar procedure is applied to drought characteristics derived from Standardized Runoff Index (SRI) to estimate the climate change effects on severity conditional probability and drought return periods.

4.1. Study Area and Data

The Upper Klamath Basin drains the south-central Oregon and the north-central California with the drainage area of 5,158,340 acres (20,875.0614 km^2). The basin consists of various land types like-wetlands and the rivers, desert, and semiarid areas with under average snowpack and precipitation, and hence the cycles of flood and drought events makes it a big challenge to manage the water resources in the basin. For example, the competition between irrigation water for agricultural commodities and the water needs of Coho salmon and other endangered fish species has become a difficult problem for water suppliers. The Upper Klamath Basin comprises 6 sub-basins:

Williamson River, Sprague River, Upper Klamath Lake, Lost River, Butte Creek, Upper Klamath East basins (Figure 3).

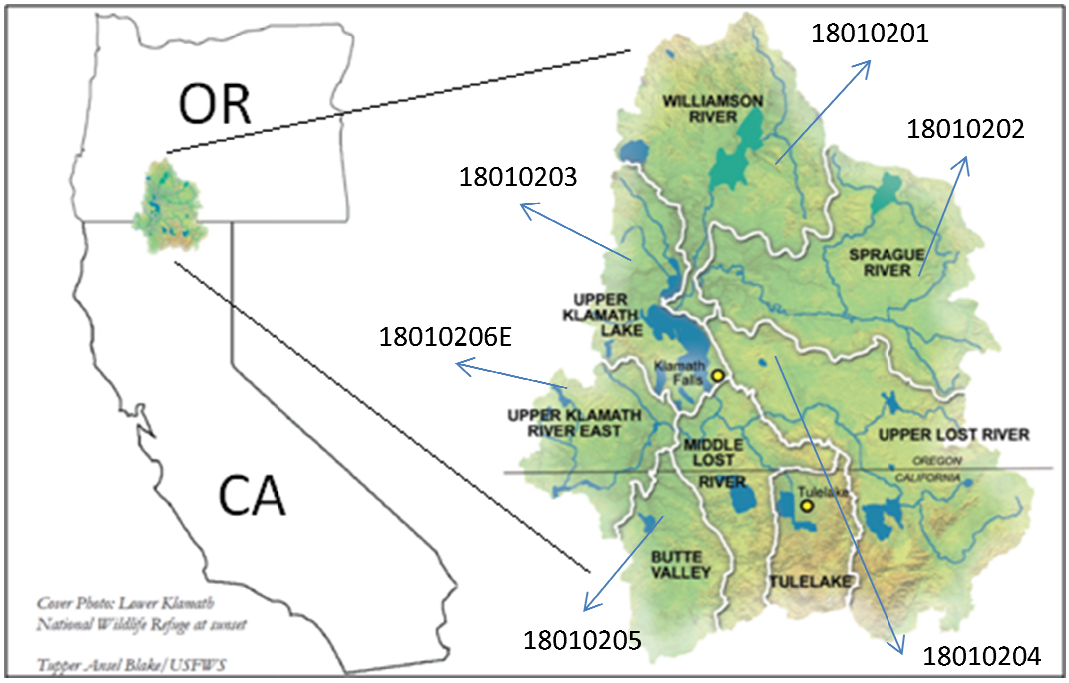


Figure 3: Study area—Upper Klamath River Basin with sub-basins and the numbers are indicating their corresponded Hydrologic Units

Before discussing the data used in this study, it is necessary to provide a brief description on the available stations and the flow direction. Figure 4 is a simplified schematic of the major river flow direction and available gage stations in the study area. There might be some other reservoirs in Upper Klamath River basin, but Crater Lake and Upper Klamath Lake are the two sites with available reservoir elevation data from USGS. Besides, the drainage area of Crater Lake is far smaller than that of Upper Klamath Lake, the reservoir storage data is only retrieved for Upper Klamath Lake. Similarly, since the

gage station of Upper Klamath River, Klamath Falls, locates at the outlet of Upper Klamath Lake, where the Williamson and Sprague Rivers meet, the streamflow data is only collected from this site. As mentioned in chapter 3, SDI and SRI are very similar and the difference is distinguished based on whether the data is from observed or simulated streamflow. The monthly observed streamflow data (units in kaf) from National Weather Service (NWS) Water Resource Forecasts, located in Upper Klamath Falls for time period (1906-2009) are used to calculate SDI. Then bivariate and trivariate copulas are constructed to investigate the historical drought situation.

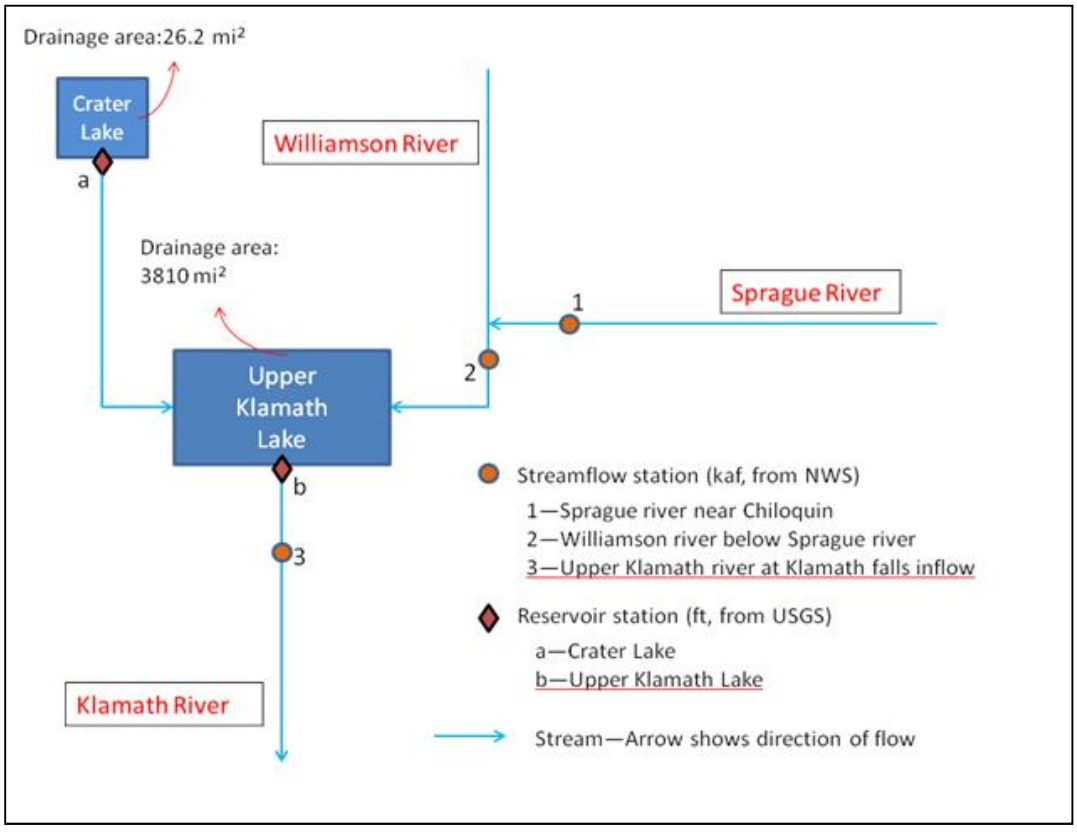


Figure 4: River flow direction and gage stations in Upper Klamath River Basin

Climate data for future from 2020-2090 using Statistically Downscaled WCRP CMIP3 Climate Projections are considered as input of Thornthwaite water balance model to simulate future runoff. Similarly, bivariate and trivariate copulas can be constructed to investigate the future drought situation. It's generally acknowledged that the projections from General Circulation Models (GCMs) have inherent uncertainty due to model structures or other factors; and so, an ensemble of six GCMs is used in this study to quantify the range of future joint drought return periods. The detail information about the selected GCMs can be found in Table 4.1.1.

Table 4.1.1: General Circulation Models used for this study

Modeling Group, Country	**Model Abbreviations**
Bjerknes Centre for Climate Research	bccr_bcm2_0
Meteo-France/Centre National de Recherches Meteorologiques, France	cnrm_cm3
CSIRO Atmospheric Research, Australia	csiro-mk3_0
US Dept. of Commerce/NOAA/Geophysical Fluid Dynamics Laboratory, USA	gfdl_cm2_1
Institut Pierre Simon Laplace, France	ipsl_cm4
Hadley Centre for Climate Prediction and Research / Met Office, UK	ukmo_hadcm3

There are two major benefits to choose WCRP CMIP3 data instead of the original GCM simulations. First, the CMIP3 data is already bias-corrected over 1950-1990 by using quantile mapping technique to remove GCMs warm/cool/dry/wet tendencies. Second, the typical GCM resolution is 2 degree grid, nearly 200km square, while the resolution

for CMIP3 data is 1/8 degree grid, approximately 12km*12km. So comparing to rather coarse GCM resolution, CMIP3 climate simulations is more preferable for a basin-relevant analysis. The downscaling method can be stated as following: Factor values for precipitation and temperature were computed by comparing the bias corrected GCM data to observation data at a 2 degree resolution. Then inverse-distance-squared interpolation was employed to get the Factor values based on 1/8 degree, lastly apply the 1/8 degree grid based Factor values to the observation data to get the downscaled bias-corrected GCM data, which becomes CMIP3 data.

4.2. SDI and margins estimation

Theoretically, SDI can be computed for any accumulation period, nevertheless, for the following reasons, 6-month time scale was chosen for this study: first, we can avoid zero values, which would be present for 1-month streamflow data; second, 6-month time period encompasses a reasonable duration of hydrological droughts.

Table 4.2.1: Statistical results to fit distributions for different time-scales of accumulated streamflow

	3-month streamflow	6-month streamflow	12-month streamflow
Ks-test	p-value	p-value	p-value
gamma	0.676	<2.2e-16	<2.2e-16
log-normal	0.008	0.7432	0.0095

To fit appropriate distributions to accumulated streamflow data, Gamma distribution(McKee et al. 1993) and Log Normal (LN) distribution(Nalbantis 2008; Shukla and Wood 2008) are tested in this study. Table 4.2.1 compares the results of applying different time-scales to accumulate streamflow data. The Ks test is chosen to test the goodness of fit for assumed distributions. The null hypothesis of Ks test is that the observed data is come from the proposed distribution, and hence, higher p-values are interpreted as tendency toward accepting the null hypothesis. According to the p-values, fitting 6-month streamflow data to log-normal distribution has the best performance among the others, and then the 6-month time scale is used for future analysis. Table 4.2.2 shows the estimated parameters and p-values of fitted gamma and log-normal distribution to 6-month streamflow data. The p-value for fitting gamma distribution is less than 0.05, while the p-value of fitting log-normal distribution is very high; greater than 0.7, and so the log-normal distribution seems a reasonable choice for streamflow.

Table 4.2.2: Marginal distribution choices for 6-month streamflow data

Marginal Distribution	Estimated parameters	p-value
6-month streamflow		
gamma	shape=5.7153,rate=0.008	< 2.2e-16
log-normal	meanlog=6.3832,sdlog=0.4389	0.7432

By comparing both graphically (Figure 5) and statistically (Table 4.2-2), the 6-month time scale and log-normal distribution are applied to fit the long-term streamflow data.

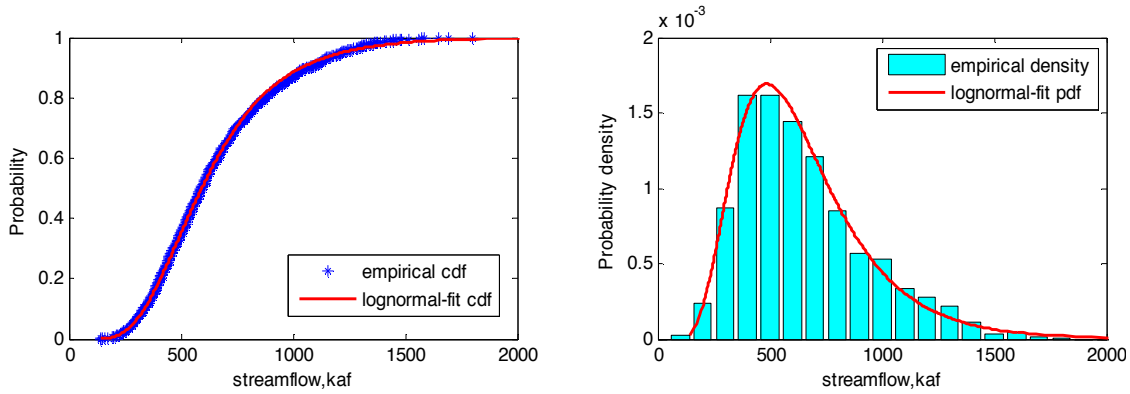


Figure 5: Comparison of empirical CDF and PDF with those of Log-normal distribution

Monthly SDI values for the time period of 1907-2009 are depicted in Figure 6. As shown, the period from June to October is detected as the driest period in a water year, and the period from January to April are relatively wet, with rather high variation. In addition, according to the reports by Balance Hydrologics, Inc., 1905-1912 would be considered as pre-Klamath Project, which, comparing to the long term records, experienced the precipitation amounts much larger than the normal values. To avoid any bias made by such precipitation trend, we decide to pick 1920-2009 period to show the application of copula technique.

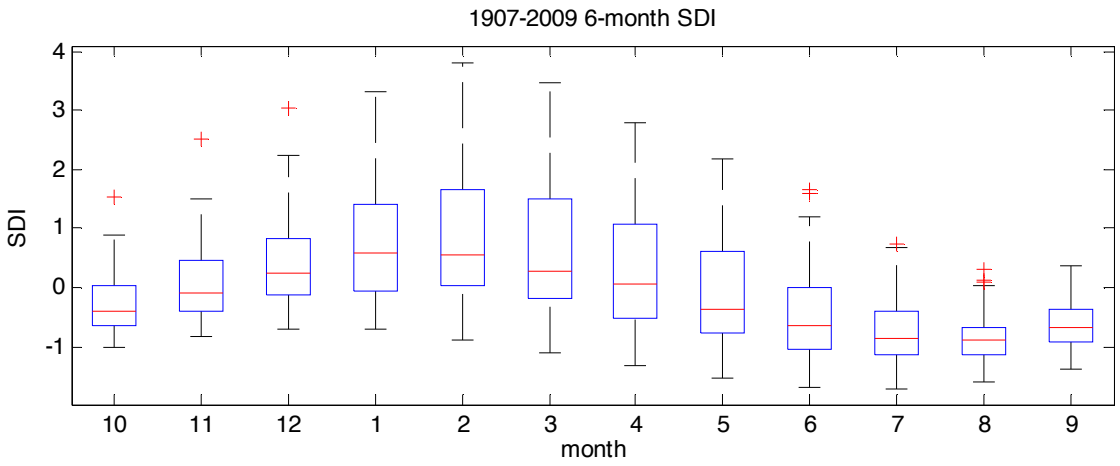


Figure 6: Monthly SDI box plot for the time period 1907-2009

To find the marginals of variables-duration, severity and intensity of hydrological droughts, these variables should be first calculated from SDI values and then the appropriate distribution for each variant should be found and fitted. The candidate distributions for duration include Exponential distribution (Shiau 2006) , Lognormal and Weibull distribution (Wong et al. 2010). For severity, most suggest gamma distributions(Shiau 2006). Therefore, those four distributions are all fitted to duration, severity and intensity, and the best choice for each variable is determined based on the p-value from Ks test.

Since the null hypothesis of Ks test is that the data are from the proposed distribution, so the higher p-value, the more reliable to accept the null hypothesis; i.e. to choose the proposed distribution. Hence, according to the p-values in Table 4.2.3, the marginal distribution choices for duration, severity and intensity are weibull, gamma and gamma

distributions, respectively. The p-values associated with chosen distributions are all greater than 0.05, which is the significance level. Besides, the closeness between the empirical and theoretical CDFs shown in Figure 7 visually supports the selected combination of marginal distributions.

Table 4.2.3: Results of fitting candidate marginal distribution to duration, severity and intensity variables

Marginal Distribution	Estimated parameters	p-value
Duration		
exponential	rate=0.3129	0.0025
log-normal	meanlog=0.9954, sdlog=0.6059	0.1145
gamma	shape=3.161,rate=0.9892	0.3281
weibull	shape=1.9456, scale=3.616	0.4979
Severity		
exponential	rate=0.2573	0.01757
log-normal	meanlog=1.1463, sdlog=0.6799	0.5575
gamma	shape=2.5196, rate=0.6482	0.7948
weibull	shape=1.6864,scale=4.3741	0.7383
Intensity		
exponential	rate=0.8553	1.25E-13
log-normal	meanlog=0.1519, sdlog=0.0932	0.758
gamma	shape=113.6474, rate=97.206	0.804
weibull	shape=10.22, scale=1.221	0.5017

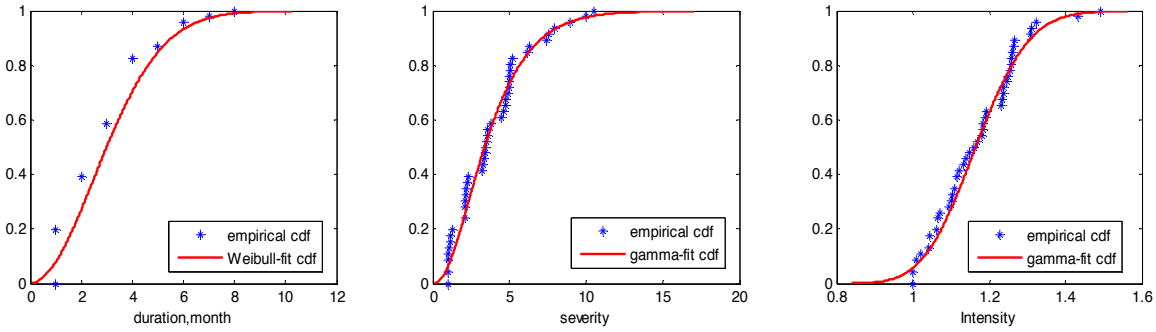


Figure 7: CDFs of observed drought duration, severity and intensity comparing with the chosen theoretical distributions.

4.3. Trivariate Copulas Application

As mentioned in Chapter 2, the trivariate Gumbel copula can be expressed as:

$$C_1[C_2(u_1,u_2),u_3] = \exp\{-[(-\ln\ u_1)^{\theta_2} + (-\ln\ u_2)^{\theta_2}]^{\frac{\theta_1}{\theta_2}} + (-\ln\ u_3)^{\theta_1}\}^{\frac{1}{\theta_1}} \quad (4.1)$$

Since θ_2 represents the existing dependency between the highest correlated pair (u_1, u_2), the correlations between three drought variable pairs should be calculated first. Pearson correlation coefficient and Kendall's tau correlation were applied to both drought variables (i.e. duration, severity and intensity) and their associated CDFs. Pearson's correlation coefficient is defined as the covariance of the two variables X and Y divided by the product of their standard deviationsσ_X, σ_Y.

$$\rho_{X,Y} = corr(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \quad (4.2)$$

The Kendall's tau coefficient is defined as (Djehiche B. 2004):

$$\tau = \frac{(number\ of\ concordant\ pairs) - (number\ of\ discordant\ pairs)}{\frac{1}{2}n(n-1)} \quad (4.3)$$

Where two pairs (x_1, y_1) and (x_2, y_2) are said to be concordant if

$$(x_1 - x_2)(y_1 - y_2) > 0 \quad (4.4)$$

And discordant if

$$(x_1 - x_2)(y_1 - y_2) < 0 \quad (4.5)$$

Pearson correlation coefficient is probably the most frequently used method to measure the linear relationship between variables.The superiority of Kendall's tau is that it is not sensitive to the data outliers and it is invariant under strictly increasing transformations. Table 4.3.1 summarizes the Pearson correlation coefficient as well as the Kendall's tau values between all possible pairs of drought variables. Results show that Duration and Severity has the highest correlation among the other pairs, suggested by both methods, and it is also concluded that the transformation has little effect on correlation, especially when the Kendall's tau is employed to assess the relationships.

Table 4.3.1: Correlations of drought variables and transformed uniform drought variables

Correlations of drought variables		
	Pearson	Kendall's tau
Duration - Severity	0.992	0.916
Duration - Intensity	0.801	0.679
Severity - Intensity	0.847	0.78
Correlations of transformed uniform drought variables		
	Pearson	Kendall's tau
Duration - Severity	0.997	0.916
Duration - Intensity	0.798	0.679
Severity - Intensity	0.83	0.78

When it is determined that the highest correlation value exists between duration and severity, the IFM method, as illustrated in Chapter 2, is utilized to estimate the parameters θ_1 and θ_2 in trivariate Gumbel copula, as well as to estimate the correlation matrix and degree of freedom in trivariate t-copula. The estimated parameters of trivariate Gumbel copula and t-copula are summarized in the Table 4.3.2.

Table 4.3.2: Estimated Parameters of Trivariate Gumbel copula and t-copula

	Estimated Parameters
Gumbel copula	$\hat{\theta}_1$=13.72, $\hat{\theta}_2$=2.44
t-copula	$R = \begin{bmatrix} 1 & 0.99 & 0.76 \\ 0.99 & 1 & 0.81 \\ 0.76 & 0.81 & 1 \end{bmatrix}$, γ= 7.25

In order to compare the shapes of t-copula and Gumbel copula density functions, their marginal bivariate PDF contours are depicted in Figure 8. The first row of figure 8 shows the relationship between duration and severity; the contours are narrow stretching for both t-copula and Gumbel copula, and the reason refers to the high linear relationship between duration and density variables.

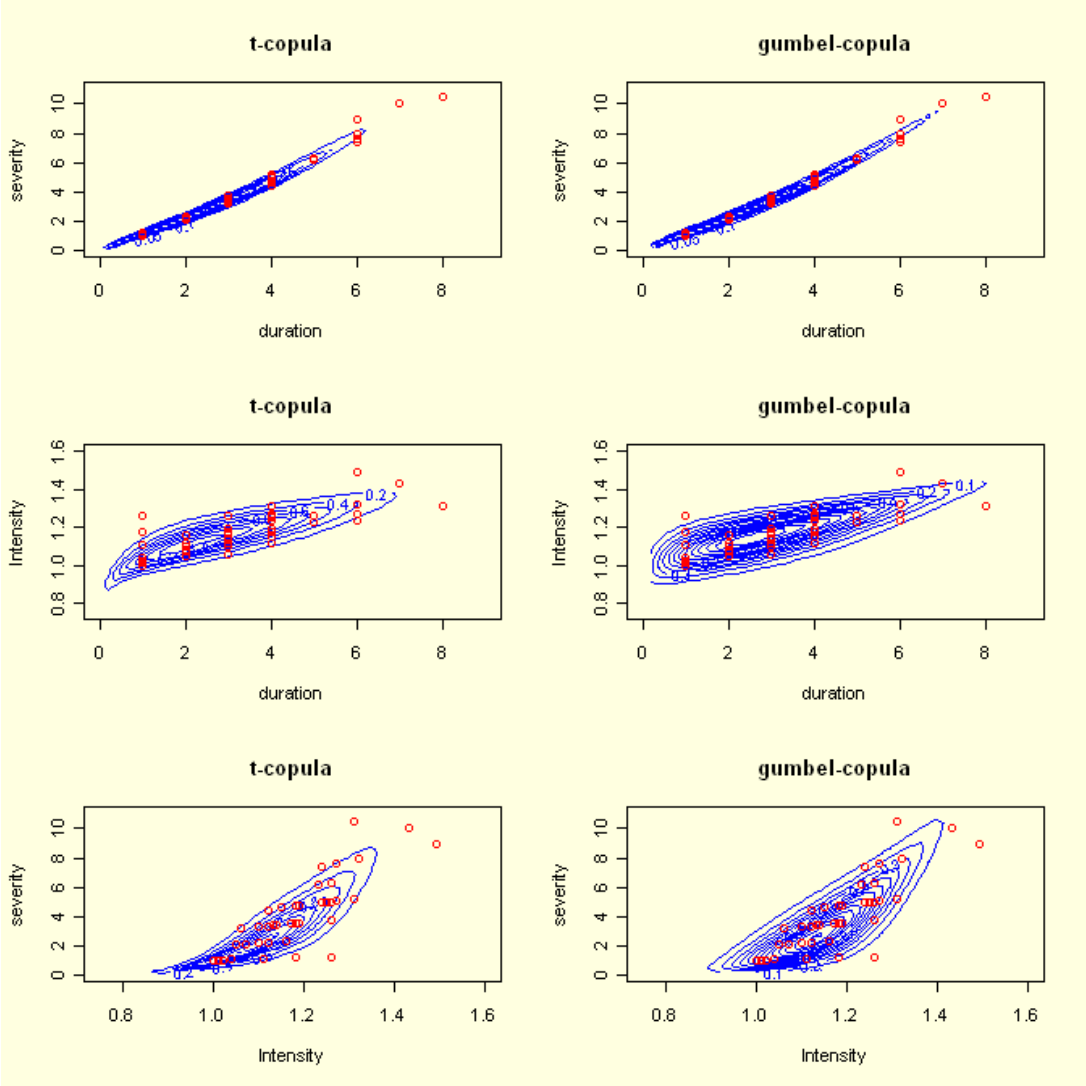


Figure 8: Marginal bivariate PDF contours of t-copula and Gumbel copula

The observed measurements are superimposed to contours as showing in red dots. It seems that bivariate Gumbel copula performs better than t-copula since it captures more observed data points. Besides, confirmed with the 3-D surface plots in Figure 9, the contours of Gumbel copula in the lower tail are more spread out than that of t copula. However the dependence in the upper tail of t-copula is as strong as that of GH copula, since the contours of both types are spread out.

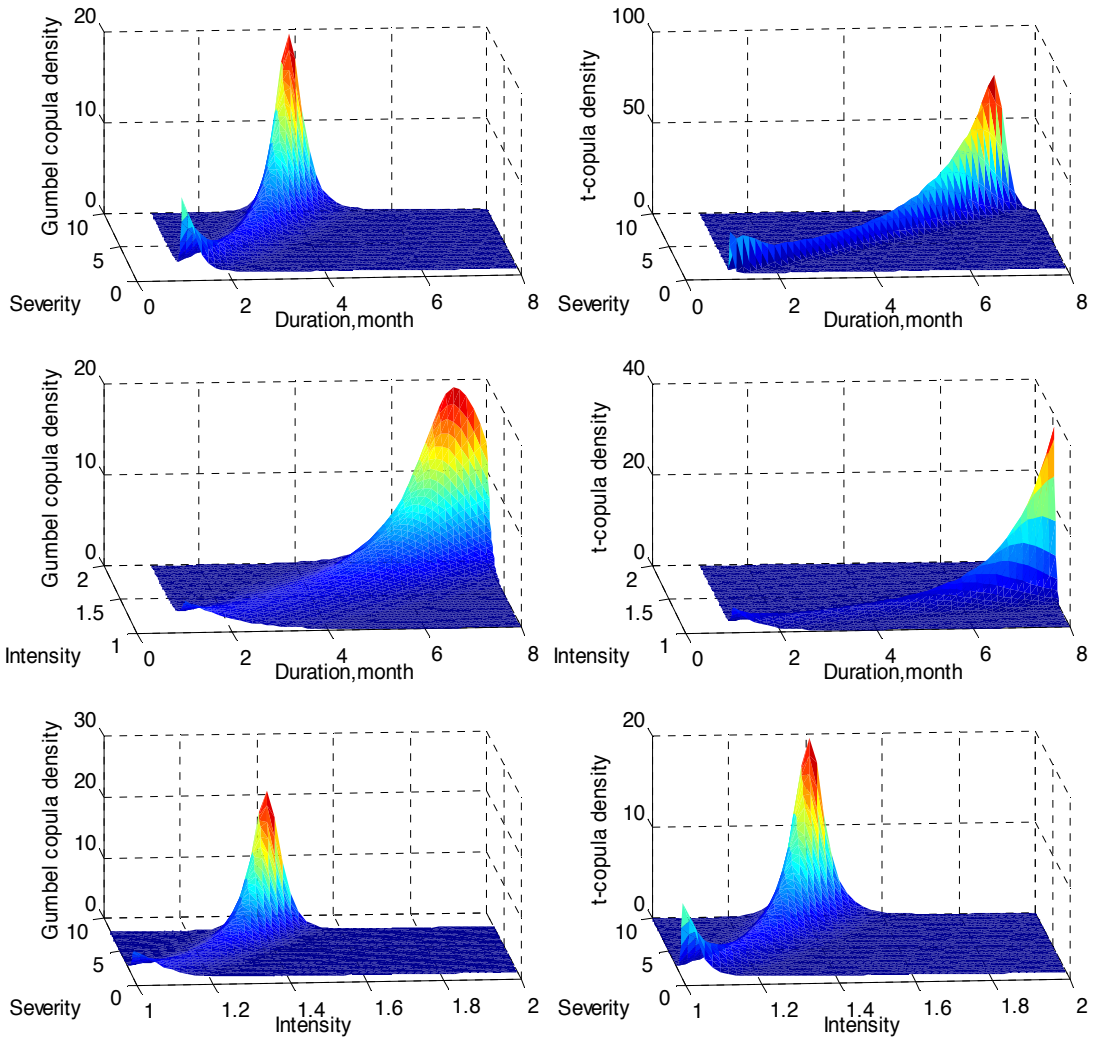


Figure 9: 3-D surface plots of marginal bivariate PDFs of t-copula and Gumbel copula

4.3.1 Conditional Probability

In water-supply management systems, the concerns are not only limited to discern if the drought duration and severity simultaneously exceed certain thresholds to trigger a drought contingency plan. To evaluate the drought severity distribution given drought duration values exceeded a certain threshold d', the conditional probability is also important for sake of managing water-supply systems.

$$P(S \leq s | D \geq d') = \frac{P(D \geq d', S \leq s)}{P(D \geq d')} = \frac{F_S(s) - F_{DS}(d', s)}{1 - F_D(d')}$$

$$= \frac{F_S(s) - C(F_D(d'), F_S(s))}{1 - F_D(d')} \quad (4.6)$$

Thus, if the bivariate copula of duration and severity is known, it is easy to derive conditional severity distribution given specific duration criteria according to the Equation (4.3.1)(Shiau 2006). Figure 10 demonstrates the severity distribution given duration value exceeding specified number of months when the bivariate t-copula is in use. Note that if given duration exceeds 8 months, the severity of any levels is almost zero, which implies that drought duration may hardly prolong for 8 months. The severity distribution in Figure 10 enables the water planners to estimate the probability of

drought events with a certain range of severity, like between 4 and 6, given an assumed exceeding drought duration.

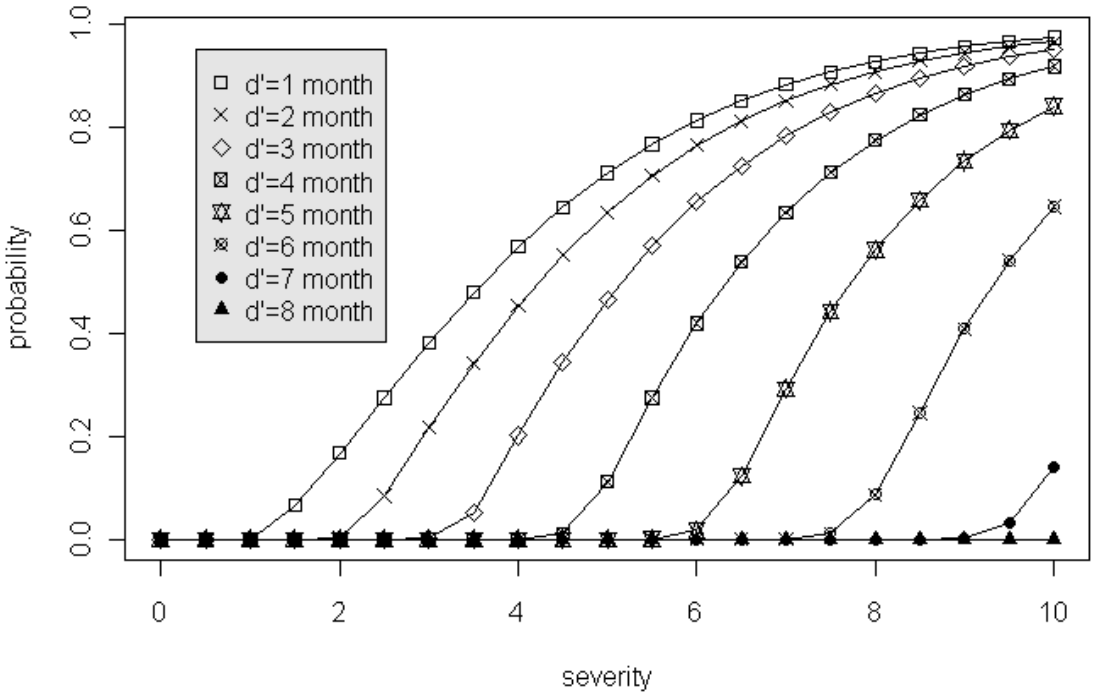


Figure 10: Conditional distributions of drought severity given drought duration exceeding d' by applying t-copula

4.3.2 Return Periods

Return period of certain drought event usually associates with a specified exceedence probability. Unlike flood frequency analysis, a specific drought event could happen multiple times in one year, and might also prolong more than one year, then, an important drought characteristic--expected drought interarrival time is needed to

include for estimating drought return periods. The return periods may be separately defined for the variables of duration, severity and (Shiau and Shen 2001) :

$$T_S = \frac{E(L)}{1 - F_S(s)} \qquad (4.7)$$

$$T_D = \frac{E(L)}{1 - F_D(d)} \qquad (4.8)$$

$$T_I = \frac{E(L)}{1 - F_I(i)} \qquad (4.9)$$

Where, T_S is the return period for droughts with severity greater than or equal to certain values; T_D is the return period for drought s with duration longer than or equal to certain months; T_I is the return period for droughts with intensity greater than or equal to certain values; and $E(L)$ is the expected drought interarrival time. In this study, the average drought interarrival time is estimated as 23.733 months for the observed period of 1920 to 2009. Table 4.3.3 summarizes drought severity, duration and intensity of single-variant return periods lasting 5, 10, 20, 50, 100, and 1000 years.

Table 4.3.3: Drought characteristics for single-variant return periods

Return Period(years)	quantile	severity	duration	intensity
5	0.604	3.99	3.46	1.2
10	0.802	5.66	4.62	1.26
20	0.901	7.17	5.55	1.31
50	0.96	9.03	6.59	1.37
100	0.98	10.38	7.29	1.41
1000	0.998	14.64	9.25	1.51

A methodology proposed by (Shiau 2003) defined the bivariate joint return periods considering both duration and severity variables. The joint drought duration and severity return periods can be characterized in two cases: return period for $D \geq d$ and $S \geq s$; return period for $D \geq d$ or $S \geq s$. These joint return periods for copula-based drought events are described as:

$$T_{DS} = \frac{E(L)}{P(D \geq d, S \geq s)} = \frac{E(L)}{1 - F_D(d) - F_S(s) + F_{DS}(d,s)}$$

$$= \frac{E(L)}{1 - F_D(d) - F_S(s) + C\big(F_D(d), F_S(s)\big)} \quad (4.10)$$

$$T'_{DS} = \frac{E(L)}{P(D \geq d \; or \; S \geq s)} = \frac{E(L)}{1 - F_{DS}(d,s)} = \frac{E(L)}{1 - C(F_D(d), F_S(s)} \quad (4.11)$$

Where T_{DS} denotes the joint return period for $D \geq d$ and $S \geq s$; T'_{DS} denotes the joint return period for $D \geq d$ or $S \geq s$.

Because there are various combinations of drought duration and severity which result in the same return period, the joint return periods are demonstrated in contour lines as shown in Figure 11. To use the plot, for d=3.5 months and s=4, for instance, the following return periods are obtained: T_{DS}= 5.229 years and T'_{DS}=4.8044 years.

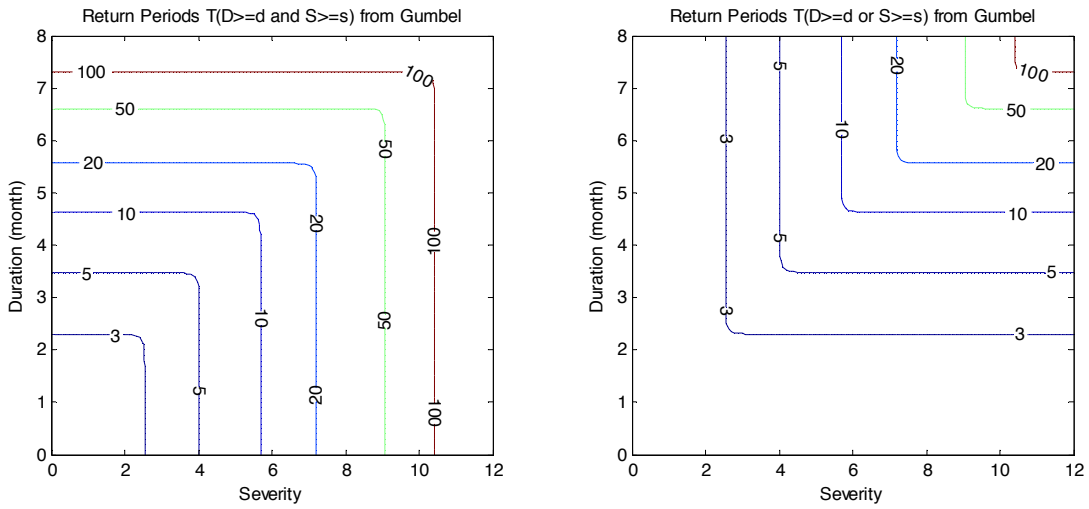


Figure 11: Joint drought duration and severity return period T_{DS} (left figure) and T'_{DS} (right figure) from Gumbel copula

Similarly, we can move a step forward to derive return periods defined for trivariate case, which should be presented as the following:

$$T_{DSI} = \frac{E(L)}{P(D \geq d, S \geq s, I \geq i)}$$

$$= \frac{E(L)}{1 - F_D(d) - F_S(s) - F_I(i) + F_{DS}(d,s) + F_{DI}(d,i) + F_{IS}(i,s) - F_{DSI}(d,s,i)}$$

$$= \frac{E(L)}{1 - F_D(d) - F_S(s) - F_I(i) + F_{DS}(d,s) + F_{DI}(d,i) + F_{IS}(i,s) - C\big(F_D(d), F_S(s), F_I(i)\big)}$$

(4.12)

Where T_{DSI} denotes the joint return period for $D \geq d$ and $S \geq s$ and $I \geq i$.

The joint return period for $D \geq d$ or $S \geq s$ $or I \geq i$, it can be expressed as:

$$T'_{DSI} = \frac{E(L)}{1 - F_{DSI}(d,s,i)} = \frac{E(L)}{1 - C(F_D(d), F_S(s), F_I(i))} \quad (4.13)$$

Table 4.3.4 compares the single-variant with trivariate return periods using t-copula. For instance, if we only consider single variable, then return period is equal to 10 years meaning that severity is greater than 5.66, duration is longer than 4.62 months, and intensity is greater than 1.26, respectively. But if we consider the joint behavior of those three variables, we can find that return period for case $D \geq 4.62 \; and \; S \geq 5.66 \; and \; I \geq 1.26$ is equal to 13.79 years, while return period for case $D \geq 4.62 \; or \; S \geq 5.66 \; or \; I \geq 1.26$ is equal to 7.17 years. In practical , single return period of 10 years means the probability of this kind of drought is 20%, while considering the joint behavior, the probability of all drought variables exceeding is 7.25%.

Table 4.3.4: Trivariate joint return period VS single-variant return period using t-copula

	Trivariate t Copula	
Return period for single (years)	Return period for Tdsi (years)	Return period for T'dsi (years)
5	6.32	3.9
10	13.79	7.17
20	28.84	13.57
50	73.09	32.01
100	146.97	62.43
1000	1464.4	594.91

4.3.3 Measure of Goodness of fit

In order to apply the copula method to drought analysis under climate change, it is required to test the goodness of fit measurement for both t-copula and Gumbel copula. It is hard to employ any statistical test to assess copula's goodness of fit, but we can graphically compare the closeness between empirical copula and theoretical copula. To construct the empirical copula, first the variables of duration, severity and intensity should be ranked based on their observed values, and then, their empirical CDFs are made based on the ranks. Thereafter, the empirical CDFs are used to evaluate the empirical trivariate copula. Briefly, in empirical copula, the theoretical CDFs of drought duration, severity and intensity are replaced by their empirical CDFs. Figure 12 shows the quantile-quantile plot for empirical copula and theoretical copula, and both t-copula and Gumbel copula show good performance since the points are close to the red diagonal line.

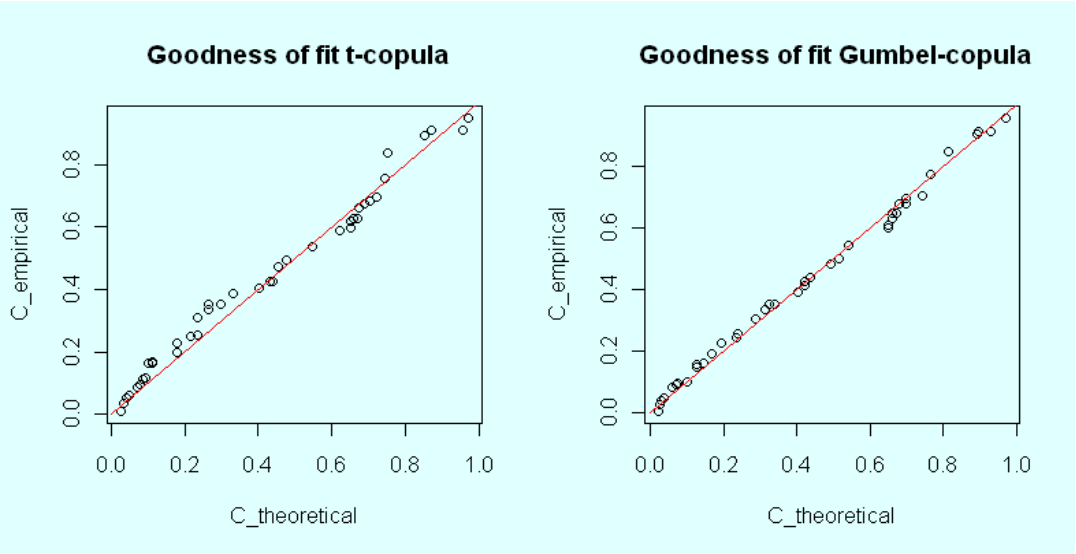


Figure 12: Empirical copula against theoretical t-copula (left) and Gumbel copula (right)

4.4. Copula analysis on climate change

Copula application to hydrology has become popular not long ago, while most studies employ the merit of joint distribution to analyze the historical and present drought conditions, rare of them has used GCM data to assess the potential drought events under climate change effects. As illustrated in before, copula construction is based on identifying drought duration, severity and intensity derived from SDI values calculated from the observed streamflow data, so before using this technique to study climate change impacts, the corresponding elements in copula equations should be found using the future time period data. First, SRI is used to replace SDI, since their definitions are very similar, which are the standardized streamflow or runoff index. Second, multi-GCM climate data can be used to simulate future runoff, which would be an ensemble of runoff time series. Then we can apply the similar copula procedure to estimate the potential changes of drought characteristics with respect to conditional probability and joint return periods.

4.4.1 Conditional probability

As suggested in figure 13, for the same drought duration 1 month, future droughts predicted by all six GCMs are less severe than historical droughts. Specifically, GCM-ipsl_cm4 is evidently wetter than other GCMs, while bccr_bcm2_0 and cnrm_cn3 cause

very close conditions. All in all, except ipsl_cm4, other five GCMs have very similar forecast results. The reason to choose the duration of 1-month is that most of the drought events during 2020-2090 persist for only one month, and rarely reaches to five months. It is worth noting that the largest duration value for period of 1920-2009 is 8 months. So we can inferred that, in comparison with historical droughts, the future drought events would have less severity in most of the time.

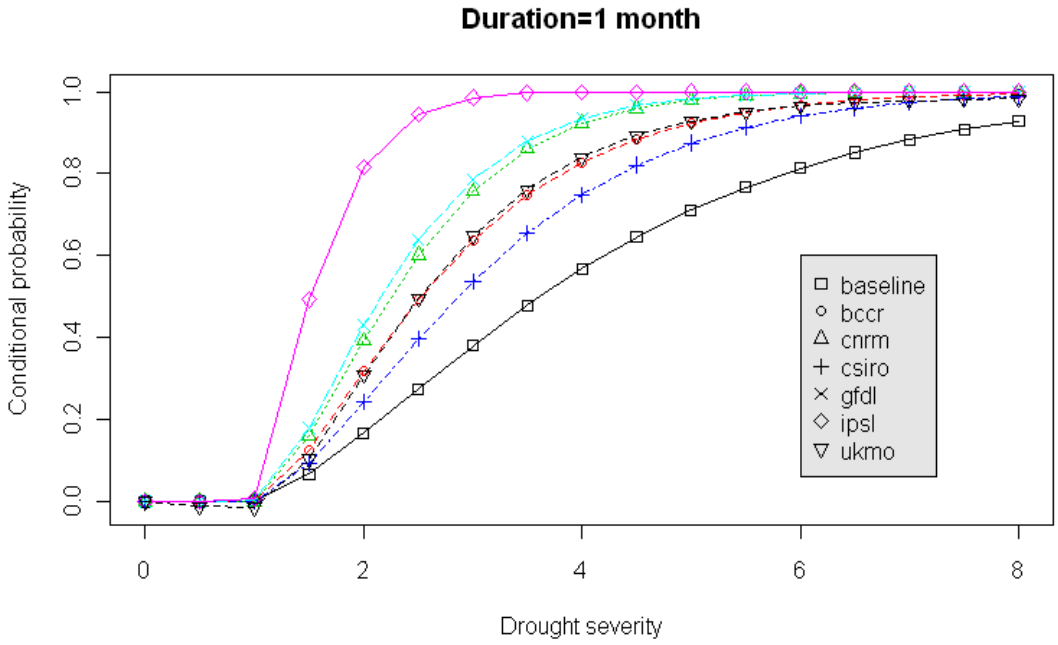


Figure 13: Conditional drought severity distribution given duration=1 month (baseline period corresponds to 1920-2009, and the future period is taken as 2020-2009): bccr, cnrm, csiro, gfdl, ipsl and ukmo are six General Circulation Models.

4.4.2 Return periods

Since the results of conditional probability analysis show that bccr_bcm2_0 performs very close to the average level of employed GCMs', and the scenario a1b is known as the medium greenhouse gas emission scenario, the bivariate return period contours are made based on bccr_bcm2_0 with a1b emission scenario (Figure 14). Comparing the contours associated with the future climate conditions with the historical conditions (Figure 11) reveals some rather remarkable differences. For example, the maximum duration shrinks from 8 months to 6 months, and the maximum severity is also reduced from 12 to 8. Moreover, d=3.5 months and s=4 result in T_{DS} of 5.229 years and 16.4429 years for historical and future periods, respectively. As a result, for the same short return periods, the duration and severity of future events are both smaller than those of historical events. AS a brief summary, the potential drought events in future would become more frequent yet less severe than the historical drought events, and the possibility of duration longer than 5.5 months is less than 0.01.

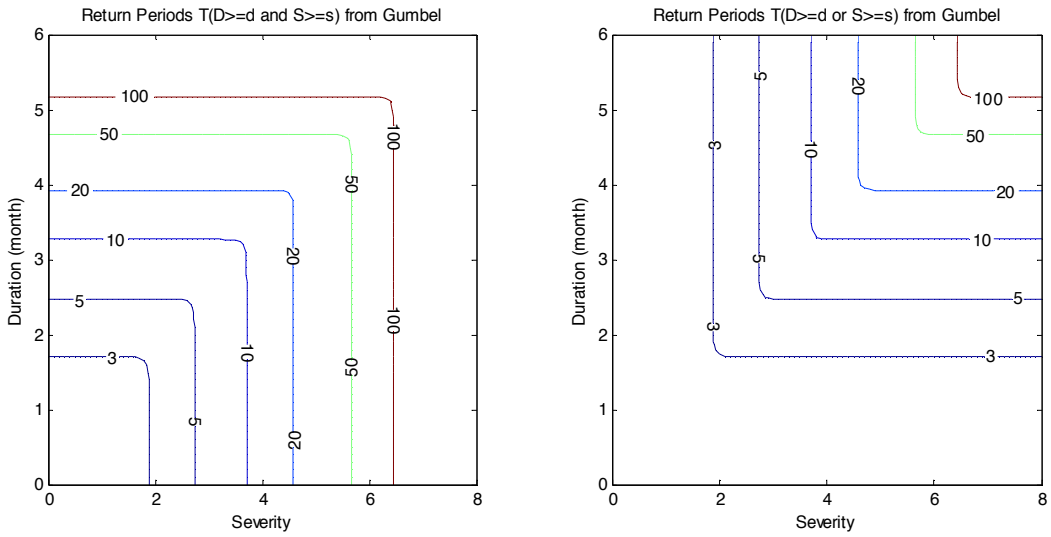


Figure 14: Bivariate return period from GCM bccr_bcm2_0 with scenario a1b for 2020-2090: the left panel is return periods estimated by duration and severity both exceeding certain levels, and the right panel is return periods estimated by either duration or severity exceeding certain levels.

Further application to trivariate return periods for 2020-2090 is accomplished in this study (Figure 15), and the major outcomes from comparing the results with those of historical data (1920-2009) are enlisted as the following:

- For the joint return period with D≥ d and S≥ s and I≥ i:

  (1) For 20-year drought or less severe drought conditions (like 5 or 10 year drought), most GCMs predict less frequent drought in future, while few GCMs predict more frequent droughts in future.

  (2) For 20-year drought or more severe drought conditions (like 50 or 100 year drought), all the GCMs predict less frequent drought in future.

- For the joint return period with D≥ d or S≥ s or I≥ i:

  (1) All the GCMs predict more frequent drought in future except GCM ipsl_cm4.

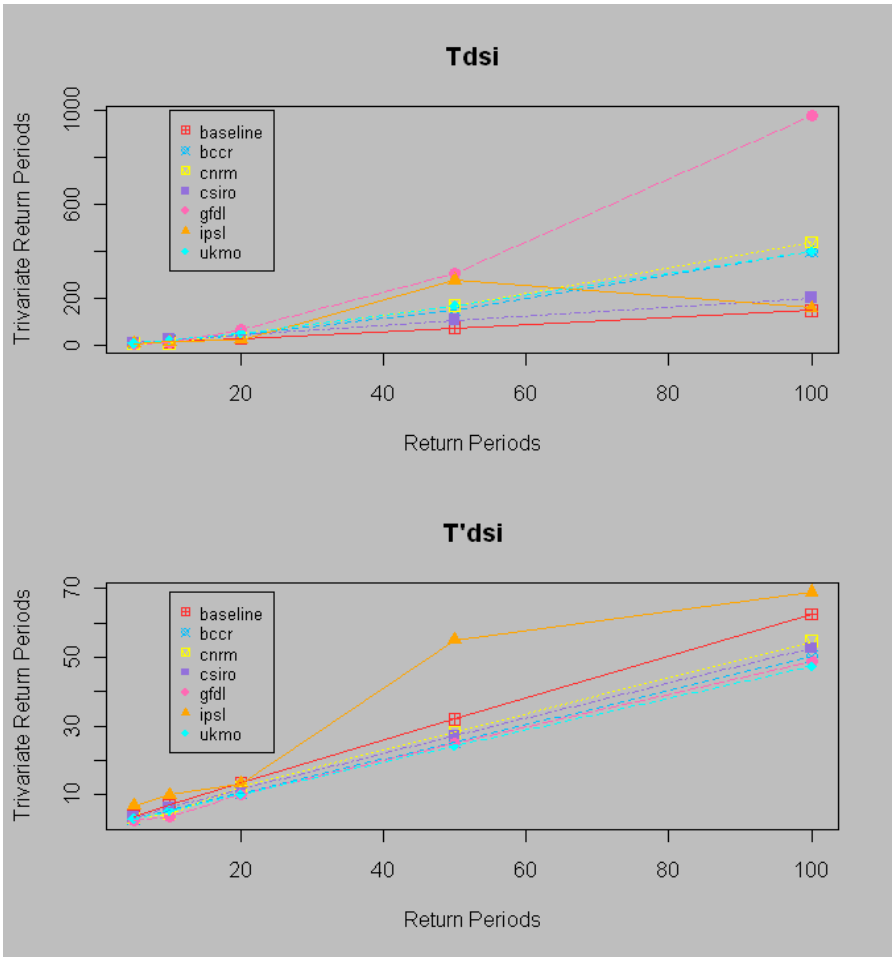


Figure 15: Trivariate return period comparison from t-copula: baseline period corresponds to 1920-2009; bccr, cnrm, csiro, gfdl, ipsl and ukmo are six General Circulation Models.

Chapter 5: Climate change effects predicted by SPI, PDSI and SWSI

In chapter 1 it was discussed that drought types include meteorological drought, agricultural drought, hydrological drought and socio-economic drought. Except for the last type, which is usually considered as the secondary drought impact, the first three types of drought can be considered as the direct drought impact outcomes. Chapter 4 did a comprehensive demonstration of climate change effects on hydrological drought by using copulas; this chapter focuses on the trend of different types of drought and different greenhouse gas emission scenarios.

5.1. Data

This study calculate yearly SPI, PDSI and SWSI for three time periods: reference time period which is from 1975-2005, and two future time periods which are 2020-2050, 2060-2090 respectively. Data sources are coming from:

1) Historical observed climate data (monthly precipitation and mean temperature) from the United States Historical Climatology Network, which was used for the reference time period: 1975-2005; the selected station (Klamath Falls 2 SSW,

OR:354506) is near to the outlet of Klamath Lake with the elevation of 1249.1 meters.

2) Six GCMs as explained in Chapter 4, with three greenhouse gas emission scenarios (A2, A1b and B1) are downloaded from Statistically Downscaled WCRP CMIP3 Climate Projections. SRES A2, A1b and B1 represent the higher, medium and lower future emissions, respectively, with respect to the balance of economic structures, technological change, and energy utilization as well as the population growth.

3) Observed reservoir elevation at USGS station NO.11507001, in Upper Klamath Lake near Klamath Falls (42.25N, 121,815W) for reference time period (1975-2005). The reservoir storage can be obtained for Upper Klamath Lake by the Elevation-Capacity relationship. As the revised Surface Water Supply Index (SWSI) formulation indicates, the reservoir storage should be estimated at the beginning of the snowmelt season. Specifically, We first downloaded the real time reservoir elevation data for March 31st every year during reference period, then estimated the related reservoir storage (units in kaf, meaning 1000*acre*feet) by the Elevation-Capacity relationship curve. Notice that reservoir storage presents the current water resources condition, so there is no real time data available for future time periods, thus, the mean reservoir storage over reference period was used to calculate future SWSI, which can be illustrated

as: future water resources condition in the Upper Klamath Lake is as same as the average reservoir storage of reference period.

5.2. Methodology

The aim of this study is to evaluate the impact of climate change on meteorological, agricultural and hydrological drought by adopting three drought indices: SPI, PDSI and SWSI. Applying Drought analysis on these three indices can characterize the duration, severity and intensity of drought for both historical and future time periods. The procedure includes calculating SPI, PDSI and SWSI and identifying drought characteristics. The SPI calculation can be conducted by SPI program, PDSI and SWSI calculation details can be found in Chapter 3.

5.3. Results

The purpose of this study is to analyze three important drought characteristics under climate change impact, which are severity, intensity and duration, respectively. The moderate drought situation has been chosen as the truncation level, which means the situations when SPI values are less than -0.99, PDSI and SWSI values are less than -1.99. Notice that the duration is a positive value, severity and intensity are negative values. The bigger the duration number, and the smaller the severity and intensity values, indicates a high severe drought.

5.3.1 Trend of drought characteristics over time

Figure 16 shows the drought events under scenario a1b, GCM bccr_bcm2_0, estimated by different drought type over three time periods. As indicated by the SPI values, the meteorological drought becomes less severe in the future time periods, but the occurrence of extreme low SPI values in 2020-2050 suggests that this period may experience extreme low precipitation in 2021 and 2036. Compared to the PDSI values of the reference period, the first future time period (2020-2050) see less drought events, and no drought events lasting more than 1 year; while the second future time period turns to the opposite: showing there would be two drought events lasting for 3 years. The changes in SWSI values are not significant, with close duration, severity and intensity in all three periods. Notice that the agricultural drought and hydrological drought estimated by PDSI and SWSI respectively, are usually 1 to 2 years lag than meteorological drought, which is estimated by SPI, and this phenomenon exists in both

reference period and future periods.

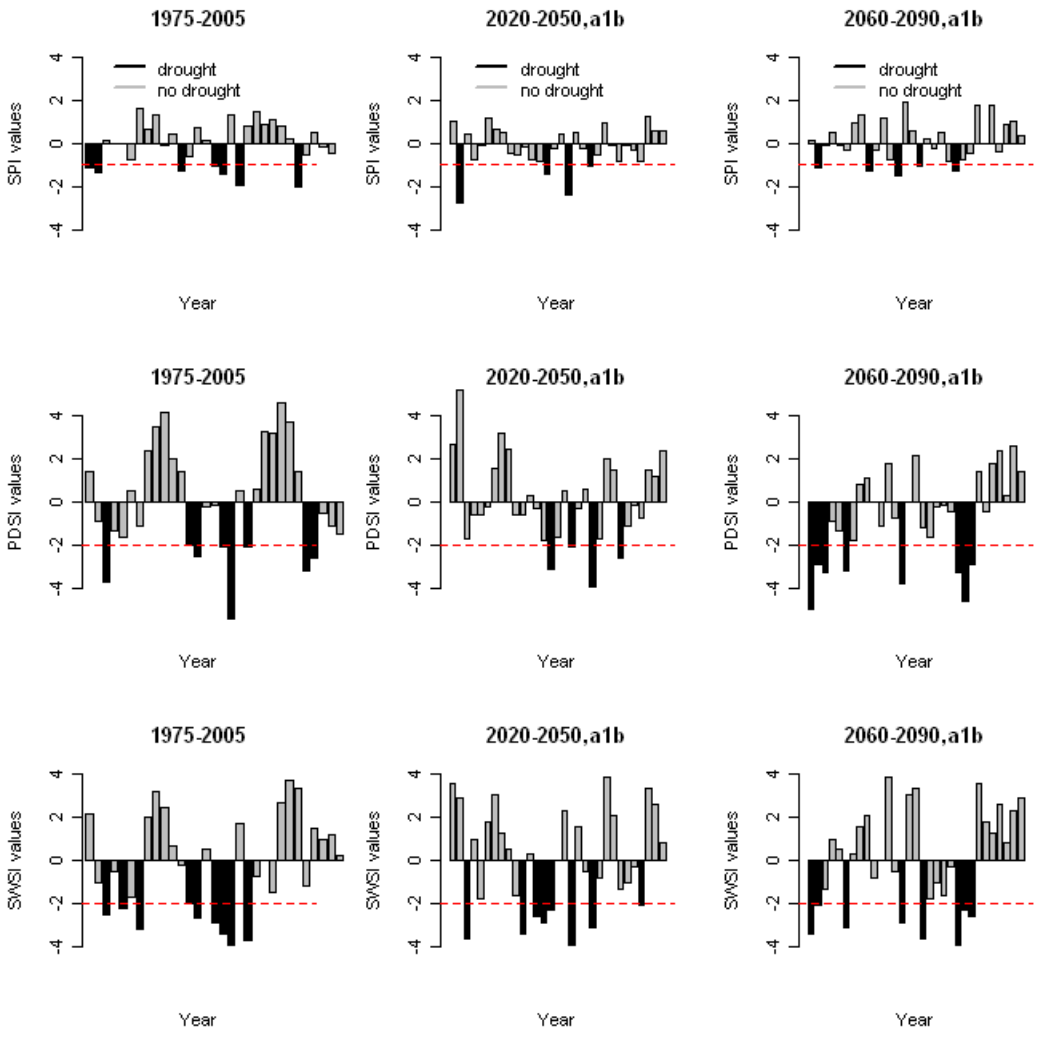


Figure 16: Drought events of different drought type under SRES a1b, GCM bccr_bcm2_0

A close look at the relative change of drought characteristics for each drought type, showing in figure 17, would reveal that the total drought years decreased for the first future time period (2020-2050) for all the drought type. Although the total drought

years of the second future time period (2060-2090) see a decrease or an even level, comparing to reference period (1975-2005), all three indices show that the average duration would increase in 2060-2090, which means the persistency is enhanced in this period. In addition, the average severity improved for all three indices for 2020-2050, but the severity of agricultural drought (PDSI) for 2060-2090 experienced an evident decline even though with slightly increased SPI.

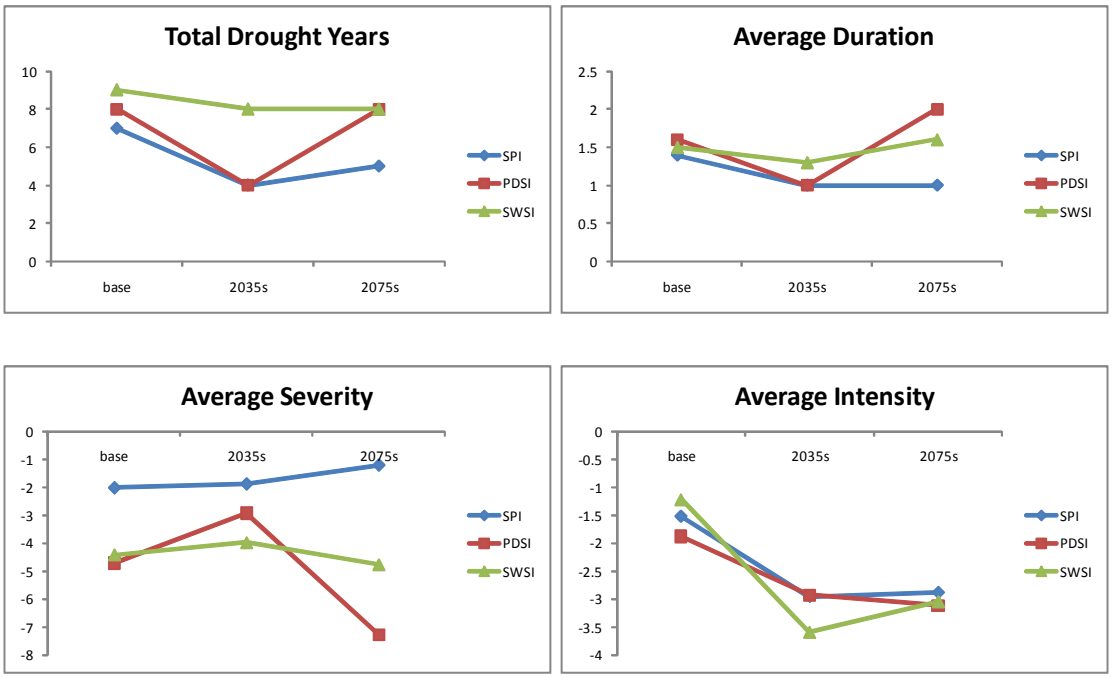


Figure 17**:** The trend of Drought characteristics of SPI, PDSI and SWSI over time

This could be explained by the influence of the global warming, so the enhanced evapotranspiration offset the slightly increased precipitation effects. That's why the

severity of PDSI and SWSI decreased whilst the severity of SPI improved. The intensity has the similar trend as severity, which also experiences a worse situation for agricultural and hydrological drought for 2060-2090 (figure 18). From the time point of view, the simulation from single GCM model, middle emission scenario a1b, suggest that the first future period 2020-2050 has less drought risk while the second future period of 2060-2090 would experience more severe agricultural drought comparing to reference period of 1975-2005. The primary reason is that although both future time periods have increased precipitation, specifically, increasing 28.8% and 33% for the 2020-2050 and 2060-2090 respectively, the temperature for both future time periods also see a 4% and 18% increases. From the results we can see that PDSI values representing the agricultural drought are very sensitive to temperature, and the enhanced evapotranspiration resulted from increased temperature compensates the slightly increased precipitation in this period.

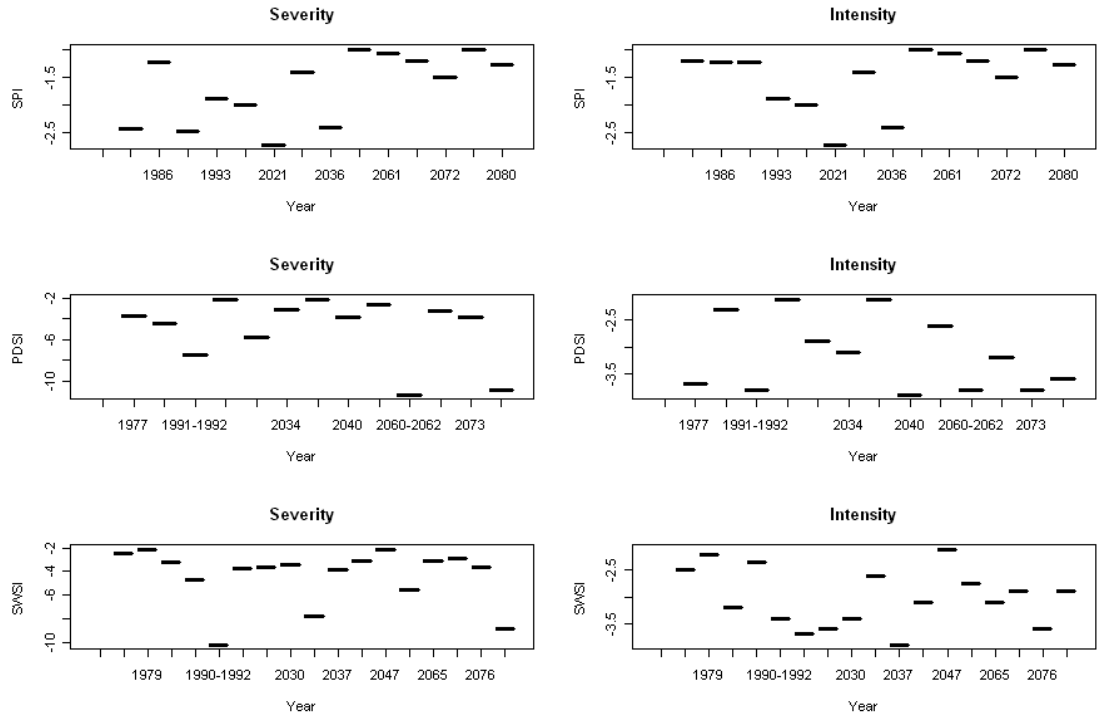


Figure 18: Severity and Intensity of SPI, PDSI and SWSI over three time periods.

5.3.2 Comparison of different greenhouse gas emission scenarios

Future period 2020-2050 suggests a relative wet period comparing to other two periods, as shown in figure 16 and figure 17 indicated for scenario a1b. In order to see how much difference the human activity could affect the climate change, based on the concentration of greenhouse gas, three emission scenarios have been chosen, which are a1b, a2 and b1. According to IPCC(Nakicenovic et al. 2000), SRES a2, known as “higher” emissions path, represents technological change and economic growth become more fragmented and slower, and population growth is higher. SRES a1b, known as “middle”

emissions path, represents technological change in the energy system is balanced across all fossil and non-fossil energy sources, where balanced energy system is defined as not relying too heavily on one particular energy source. SRES b1, known as “lower” emissions path, represents a rapid change in economic structures toward service and information, emphasis on clean, sustainable technology, reduced material intensity as well as improved social equity.

Figure 19 shows the drought events under different scenarios for 2020-2050, simulated by GCM bccr_bcm_2_0. High emission scenario a2 predicted the largest number of total drought years, longest duration and the highest severity for all three indices, especially for PDSI, which also has the largest intensity. At the same time, the largest intensity for SPI and SWSI happened under the middle greenhouse gas emission scenario a1b.

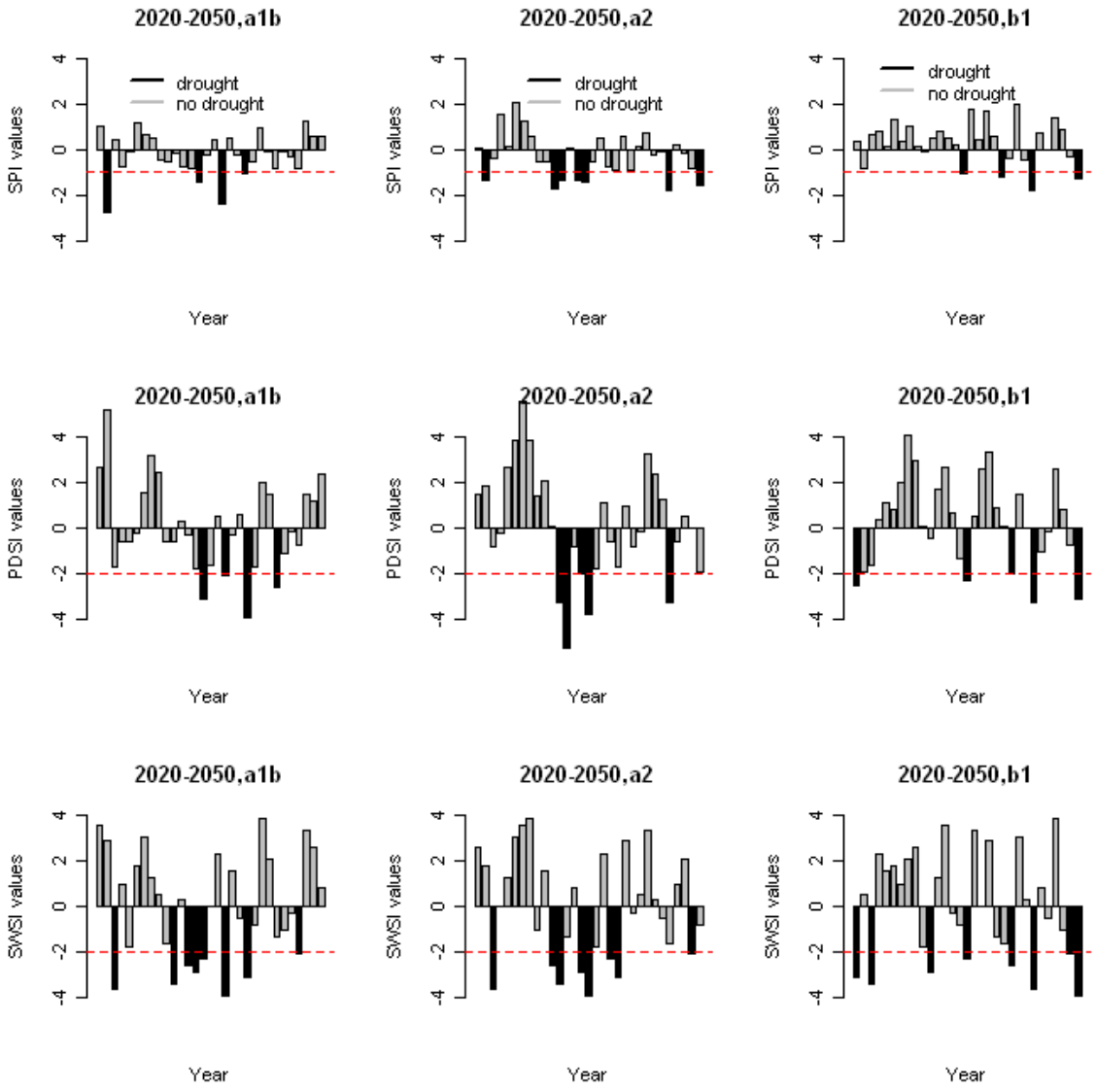


Figure 19**:** Projected drought events from single GCM, multi-scenario for multi-drought type for future period 2020-2050: a2, a1b and b1 represent the higher, medium and lower future greenhouse gas emission scenarios

A close look at the comparison of the three different scenarios is shown in Figure 20, which depicts the severity, intensity and the total amount of years for three indices, respectively. As shown by the first column of figure 6, the fifty percentile for three drought indices under scenario a2 indicates the most dangerous condition for agricultural and hydrological drought. For meteorological drought index SPI, it takes place under SRES a1b. Notice that there exists a difference between fifty percentile and the average mean values for SPI, which points to different scenarios with respect to the severity. The average SPI values measured by mean values indicate that the largest severity happens under SRES a2, while the fifty percentile of SPI values shows the largest severity under SRES a1b. The remaining mean values of drought indices and the fifty percentile values correspond to each other very well for all three indices, considering the drought duration, amount of years, severity and intensity.

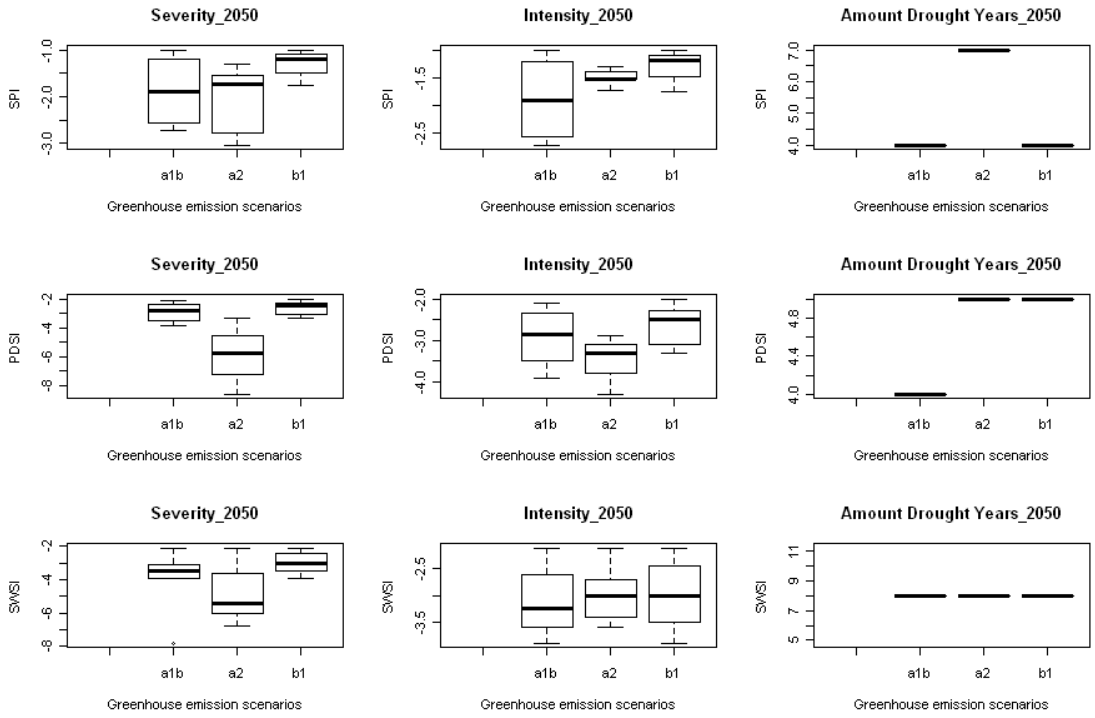


Figure 20**:** Boxplot show projected drought severity, intensity and total amount of drought years from single GCM, 2020-2050 for multi-scenario

The reason of SRES a2 to be the most dangerous drought condition seems understandable since it is referred as the “higher” emissions path. Why the “middle” emissions path SRES a1b also predicted some dangerous phenomena? To find the implied reason, one needs to compare the average precipitation and temperature simulated by the three scenarios. For SRES a1b, the mean precipitations is 37.82mm per month and mean temperature is 9.13 °C; SRES a2, the mean precipitation is 36.20mm per month and mean temperature is 8.7°C; SRES b1, the mean precipitation is 41.98mm

and the mean temperature is 8.5°C. So we can see that SRES b1 simulates the most amount precipitation with lowest temperature, which means more input of water and less output because of the lower evapotranspiration rate, so SRES b1 has less drought risk. As for SRES a1b, the amount of precipitation and temperature are both slightly more than the conditions predicted by SRES a2. Besides, according to IPCC, the CO_2 concentration level for SRES a1b and a2 are very close during 2020-2050 period, but after 2060, the CO_2 concentration level would increase significantly under SRES a2, reaching to 29 (Gt C, where, Gigatone (Gt)=10^{12}kg; C=0.273×CO_2) at 2100, while the SRES a1b see an increase of CO_2 concentration level until 2050, after that, it would slowly decrease, and reduce to 13.5 (Gt C) at 2100.

In addition, the largest intensity for the SPI and SWSI occurs under SRES a1b. The reason is the relationship between the nature of these indices and climate facts. Since precipitation is the only required input of SPI computation, the SPI values are very sensitive to precipitation, similarly, SWSI is calculated by frequency analysis for the total of streamflow and reservoir storage, so streamflow is highly correlated with SWSI. This can be found at figure 21: SPI has a very similar trend as precipitation, while streamflow shares its pattern with SWSI.

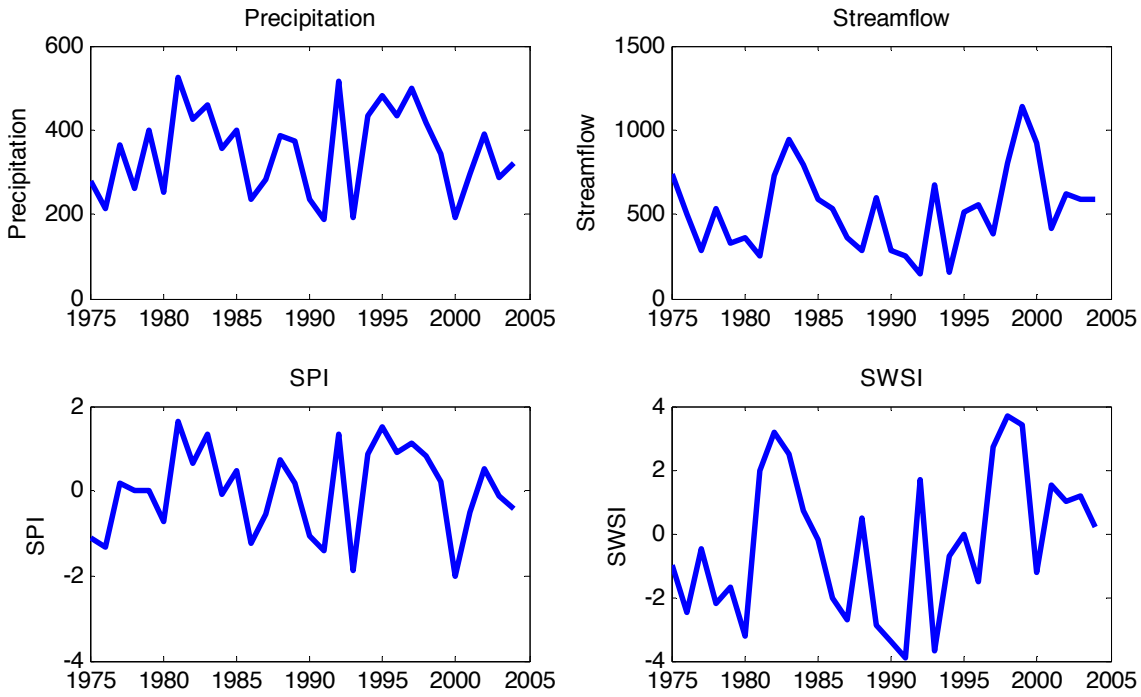


Figure 21: The relationship between Precipitation and SPI and the relationship between streamflow and SWSI from observed period 1975-2005.

So even though SRES a1b experienced both increase in precipitation (4%) and temperature (5%) comparing to SRES a2, but the slightly higher increase rate in temperature than precipitation reveals the sensitivity of SPI and SWSI, which are highly related to precipitation. On the other hand, the most severe agricultural drought estimated by PDSI series seems strongly clings to scenario a2. This can be explained that scenarios a2 has the comparatively "highest" temperature, which stands for the highest evapotranspiration from soil, so the deficit of soil moisture condition would suffer most.

5.3.3 Comparison of different GCMs

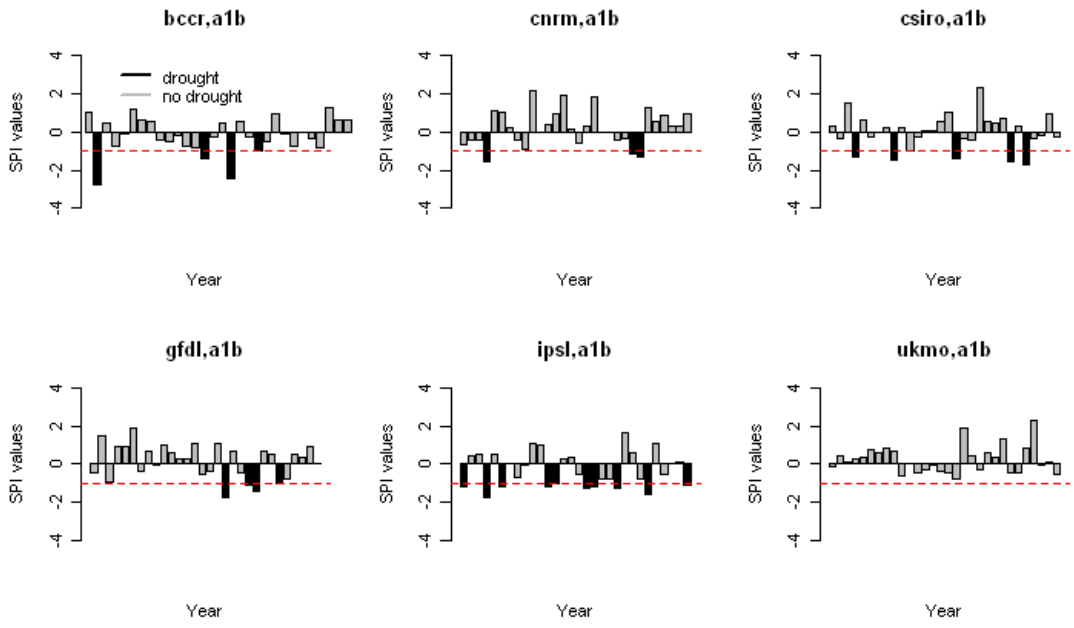
Global circulation models have uncertainties, and some model results may incline to warmer projections, while others may cooler ones. Here 6 GCMs were chosen for different GCM simulations comparison, which are bccr_bcm2_0, cnrm_cm3, csiro_mk3_0, gfdl_cm2_1, ipsl_cm4 and ukmo_hadcm3. Figure 22 shows the drought events projected by each GCM and figure 23 provides a detail comparison between multi-GCM on the aspect of severity, intensity and total amount of drought years.

Notice that figure 22 projectes no drought events for SPI values, but projectes 5 drought years for PDSI and 8 drought years for SWSI values from ukmo_hadcm3 GCMsimulations. This is because for SPI value calculation, only precipitation has been considered, while for PDSI and SWSI calculation, temperature is also involved. In fact, this model simulates the highest mean temperature (equal to 9.8°C)among six GCMs, so the high temperature causing high evapotranspiration rate, results in soil moisture deficit and low streamflow, therefore agricultural drought index and surface water supply index show there will be drought events according to the climate simulation by ukmo_hadcm3 GCM.

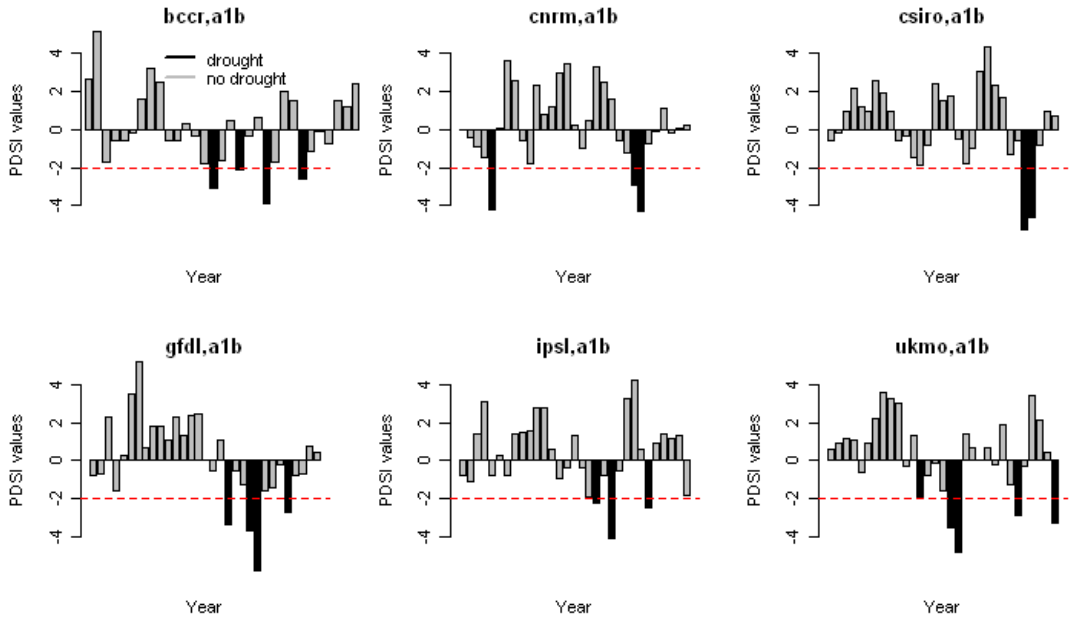
From the third row of figure 22, we can see that all six GCMs predicted the same amount of drought years, which equal to 8 years for Surface Water Supply Index. This is because we choose the same amount of reservoir storage for each GCM. Actually, we don't have any information about the reservoir storage for future periods, and it is not

easy to predict or estimate the water use condition, so the average level of reservoir storage of reference period has been chosen to apply for both two future periods. The reservoirs storage in Upper Klamath River Basin plays an important role in the total water resources, as observed streamflow and reservoir storage data indicated. Some years the reservoir storage may be even larger than the streamflow. For 2020-2050, we take the similar reservoir storage effects for all GCMs, so the total drought years show the same 8 years for SWSI simulations.

(a)



(b)


bccr,a1b
drought
no drought
cnrm,a1b
csiro,a1b
gfdl,a1b
ipsl,a1b
ukmo,a1b
PDSI values
Year

(c)

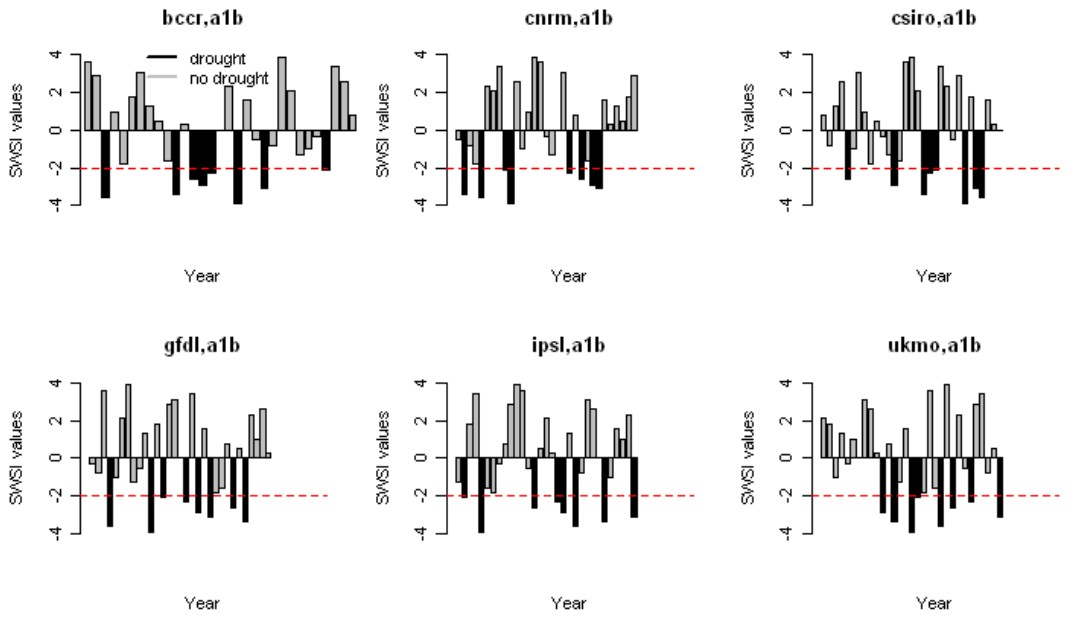

bccr,a1b
drought
no drought
cnrm,a1b
csiro,a1b
gfdl,a1b
ipsl,a1b
ukmo,a1b
SWSI values
Year

Figure 22**:** Drought events projected by multi-GCM for time period 2020-2050 under scenario a1b.

(a) Shows the SPI values, (b) shows the PDSI values and (c) shows the SWSI values.

A close look at the uncertainties from multiple-GCMs could be found at figure 23. The most dangerous agricultural drought happens at the simulation of csiro_mk3_0 (marked as m3 in figure 23). Unlike the climate data analysis, which show m3 has the relatively abundant precipitation and the lowest mean or 0.5 quantile temperatures, and this is supposed to have less drought risk, however the results indicate opposite result. To resolve this paradox phenomenon, we compared the precipitation and temperature from several typical months. We found that m3 predicted the lowest 0.5 quantile temperature at January and April for 2020-2050, so that means snow is the major part of the winter precipitation and at the spring season, the temperature is still low so there is still a large amount of no-melt snow. That's why m3 actually has the most severe agricultural drought due to the lowest liquid precipitation.

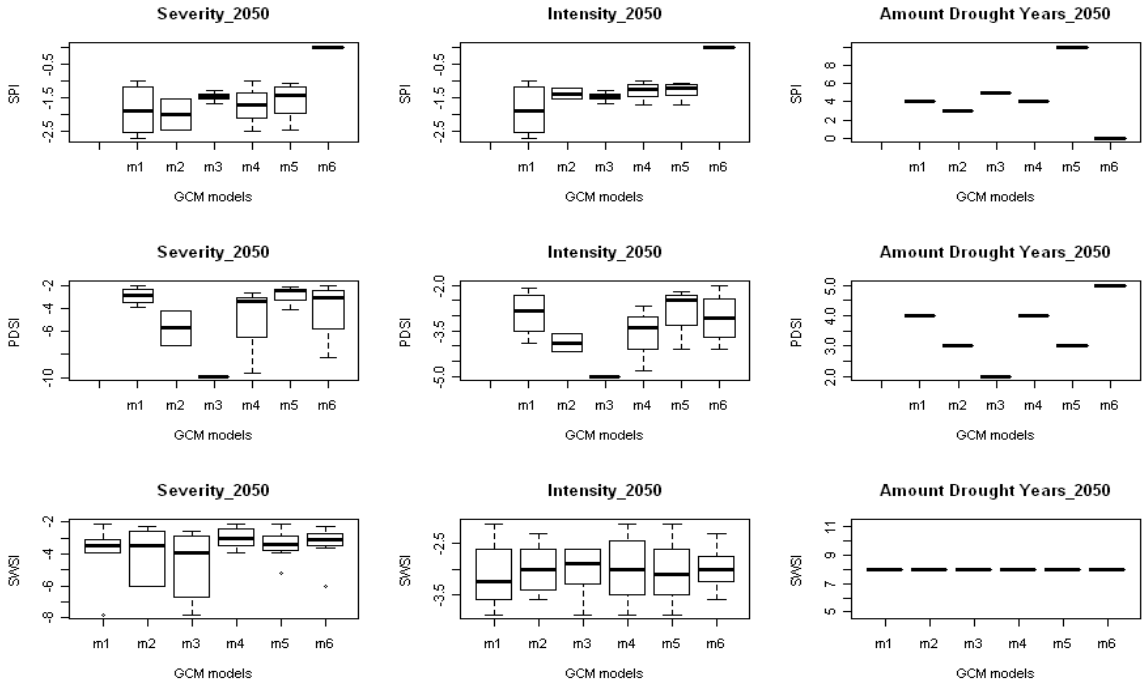


Figure 23**:** Projected drought severity, intensity and total amount of drought years from multi- GCM, 2020-2050 for a1b scenario

For SPI and SWSI simulations, bccr_bcm2_0 (m1) shows the largest intensity. If we consider the total six GCM mean precipitation and temperature as the measure basis, we find that the mean precipitation and temperature simulated by m1 is 3% and 0.9% below the basis. In other words, it offsets the decreased temperature effect, and the lower precipitation still affects the hydrological cycle. Especially for SPI and SWSI, which are highly related to precipitation, the effects seem more obvious. Similarly, as the fifty percentile precipitation and temperature values from cnrm_cm3 (m2) indicated, 3.4%

below the average precipitation and 2% below the average temperature cause the meteorological drought index SPI experience the longest duration and largest severity.

Chapter 6: Discussion and Conclusions

Upper Klamath River Basin is a semiarid area, and has a long history of facing water challenges due to yearly in turn drought and flood events. In order to estimate the drought condition on a long-term view, this study has two objectives: first, apply copula technique to assess the drought duration, severity and intensity joint behavior on hydrological drought defined by SDI or SRI; second, employ three other drought indices, namely SPI, PDSI and SWSI, which are corresponding to meteorological drought, agricultural drought and hydrological drought respectively, to estimate the trend drought characteristics evolving under climate change effects.

The first objective is achieved through conditional probability and return periods estimation for historical and future time periods, addressed in Chapter 4. The main results show that either t-copula or Gumbel copula are suitable for drought analysis. Fundamentally, the reason is because t-copula and Gumbel copula have tail probability, which usually considered as appropriate for estimating extreme events. This is also supported by the closeness of the marginal bivariate copula contours with superimposed observed data. With duration-severity bivariate Gumbel copula, we conducted the conditional severity distribution given drought duration for both historical and future drought events. All six GCMs predict that potential changes in drought characteristics would be decreased drought duration, diminished severity, at the same time; we found that climate data from GCM ipsl_cm4 result in much more wet

tendency than other GCMs. Drought return period is a useful water-supply planning criteria, and trivariate copula based on the joint behavior of duration, severity and intensity is more objective than solely based on single drought variable. In all, the climate change effects on drought estimated by applying copulas is that drought would become more frequent but less severe.

The Chapter 5 majorly discussed climate change effects by assessing other three drought indices, focusing on how different GCMs as well as different greenhouse gas emission scenarios would influence on their corresponding drought types. The results can be viewed as three aspects: First, the change of drought characteristics evolve over time. From simulations of GCM bccr_bcm2_0 under greenhouse gas emission scenario a1b, we can find that future time period 2020-2050 has less drought possibility while 2060-2090 has a high drought potential comparing to the reference period 1975-2005. This can be explained by the global warming at 2060-2090 enhancing evapotranspiration offset the slight increased precipitation effects at this period.

Second is the comparison between three greenhouse gas scenarios. Scenario a2 represent the high level emission condition, while a1b represent the middle level and b1 is the lowest emission scenarios. The results show that scenario a2 is the most drought occasion for PDSI series, in other words, it would cause the most severe agricultural drought under scenario a2. The longest duration and largest intensity for SWSI happened under SRES a1b, which also captured the largest SPI intensity. The reason is

that the huge difference in greenhouse gas concentration level between SRESa2 and SRESa1b is obvious only for the second future time period. For 2020-2050, SRESa1b captured the highest mean temperature, which might be even more danger than SRES a2.

Third, the comparisons between multi-GCM suggest that csiro_mk3_0 (m3) predicted the most severe agricultural drought and the longest duration and severity for hydrological drought. Even though m3 seems have the relatively sufficient precipitation and lower temperature, which will not lose a lot of water by evapotranspiration. But actually the temperature is pretty low especially for winter and spring, so the winter precipitation is consist of large amount of snow and do not get a lot snow-melt in spring, which causing the lowest liquid precipitation comparing to other GCMs for 2020-2050.

Last but not least, since all the indices are yearly based, which indicating long-term drought. Although some results show less drought risk, that may not true when taking into short-term drought account. For example, simulations from GCM ukmo_hadcm3, SRES a1b for 2020-2050 suggest no drought condition predicted by SPI values; still, the results do not mean there will be no short drought during this period.

Other limitations of this study include using average monthly reservoir elevation as a substitute for future SWSI calculation, and due to the data availability only the largest reservoir-Upper Klamath Lake in study area has been used, which might result in inconsistency for some locations.

Reference

Blenkinsop, S., and Fowler, H. J. (2007). "Changes in drought frequency, severity and duration for the British Isles projected by the PRUDENCE regional climate models." Journal of hydrology, 342(1-2), 50-71.

Borg, D. S. (2009). "An Application of Drought Indices in Malta, Case Study."

Cancelliere, A., Mauro, G. D., Bonaccorso, B., and Rossi, G. (2007). "Drought forecasting using the standardized precipitation index." Water Resources Management, 21(5), 801-819.

Cunderlik, J. M., and Simonovic, S. P. (2005). "Hydrological extremes in a southwestern Ontario river basin under future climate conditions/Extrêmes hydrologiques dans un basin versant du sud-ouest de l Ontario sous conditions climatiques futures." Hydrological Sciences Journal, 50(4), 631-654.

Day, G. N. (1985). "Extended streamflow forecasting using NWSRFS." Journal of Water Resources Planning and Management, 111(2), 157-170.

Dingman, S. L. (1994). Physical hydrology, Prentice Hall Upper Saddle River, NJ.

Djehiche B., H. H. (2004). "An Introduction to Copulas with Applications." Svenska Aktuarieföreningen,02 March 2004, Stockholm.

Dracup, J. A., Lee, K. S., and Paulson Jr, E. G. (1980). "On the definition of droughts." Water Resources Research, 16(2), 297-302.

Dupuis, D. J. (2010). "Statistical Modeling of the Monthly Palmer Drought Severity Index." Journal of Hydrologic Engineering, 1, 143.

Embrechts, P., Lindskog, F., and McNeil, A. (2003). "Modelling dependence with copulas and applications to risk management." Handbook of heavy tailed distributions in finance, 8, 329-384.

Fleig, A. K., Tallaksen, L. M., Hisdal, H., and Hannah, D. M. (2010). "Regional hydrological drought in north-western Europe: linking a new Regional Drought Area Index with weather types." Hydrological Processes.

Frees, E. W., and Wang, P. (2005). "Credibility using copulas." North American Actuarial Journal, 9(2), 31-48.

Garen, D. C. (1993). "Revised surface-water supply index for western United States." Journal of Water Resources Planning and Management, 119(4), 437-454.

Ghosh, S., and Mujumdar, P. P. (2007). "Nonparametric methods for modeling GCM and scenario uncertainty in drought assessment." Water Resources Research, 43(7), W07405-W07405.

Gobena, A. K., and Gan, T. Y. (2010). "Incorporation of seasonal climate forecasts in the ensemble streamflow prediction system." Journal of hydrology, 385(1-4), 336-352.

Hayes, M. (2003). "Drought indices." National Drought Mitigation Center, University of Nebraska-Lincoln [http://www. drought. unl. edu/whatis/indices. htm# spi], accessed September, 9, 2007.

IPCC. (2007). "The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change." Cambridge University Press, Cambridge, UK and New York, NY, USA.

Joe, H., and Xu, J. J. (1996). "The estimation method of inference functions for margins for multivariate models." Department of Statistics, University of British Colombia, Technical Report, 166.

Karamouz, M., Rasouli, K., and Nazif, S. (2009). "Development of a Hybrid Index for Drought Prediction: Case Study." Journal of Hydrologic Engineering, 14, 617.

Loukas, A., Vasiliades, L., and Tzabiras, J. (2008). "Climate change effects on drought severity." Advances in Geosciences, 17, 23-29.

McCabe, G. J., and Markstrom, S. L. (2007). "A monthly water-balance model driven by a graphical user interface." US Geological Survey Open-File Report, 1088(6).

McCabe, G. J., and Wolock, D. M. (1999). "Future snowpack conditions in the western United States derived from general circulation model climate simulations." Journal of the American Water Resources Association, 35, 1473-1484.

McKee, T. B., Doesken, N. J., and Kleist, J. "The relationship of drought frequency and duration to time scales." 184.

Mishra, A. K., and Desai, V. R. (2005). "Drought forecasting using stochastic models." Stochastic environmental research and risk assessment, 19(5), 326-339.

Nakicenovic, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grubler, A., Jung, T. Y., and Kram, T. (2000). "Special report on emissions scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change." Pacific Northwest National Laboratory, Richland, WA (US), Environmental Molecular Sciences Laboratory (US).

Nalbantis, I. (2008). "Evaluation of a Hydrological Drought Index."

Nalbantis, I., and Tsakiris, G. (2009). "Assessment of hydrological drought revisited." Water Resources Management, 23(5), 881-897.

Palmer, W. C. (1965). "Meteorological drought." Research paper, 45, 1–58.

Pashiardis, S., and Michaelides, S. (2008). "Implementation of the Standardized Precipitation Index (SPI) and the Reconnaissance Drought Index (RDI) for Regional Drought Assessment: A case study for Cyprus."

Ricci, V. (2005). "Fitting Distributions with R." R project web site http://cran. r-project. org/doc/contrib/Ricci-distributions-en. pdf. Retrieved July, 6, 2007.

Riebsame, W. E., Changnon, S. A., and Karl, T. (1991). "Drought and natural resources management in the United States: impacts and implications of the 1987-89 drought." Westview special studies in natural resources and energy management (USA).

Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., Van den Dool, H. M., Pan, H. L., Moorthi, S., and Behringer, D. (2006). "The NCEP climate forecast system." Journal of Climate, 19, 3483-3517.

Salvadori, G., and De Michele, C. (2007). "On the use of copulas in hydrology: theory and practice." Journal of Hydrologic Engineering, 12, 369.

Shafer, B. A., and Dezman, L. E. "Development of a Surface Water Supply Index (SWSI) to assess the severity of drought conditions in snowpack runoff areas." 164-175.

Sheffield, J., and Wood, E. F. (2008). "Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations." Climate Dynamics, 31(1), 79-105.

Shiau, J. T. (2003). "Return period of bivariate distributed extreme hydrological events." Stochastic environmental research and risk assessment, 17(1), 42-57.

Shiau, J. T. (2006). "Fitting drought duration and severity with two-dimensional copulas." Water Resources Management, 20(5), 795-815.

Shiau, J. T., and Shen, H. W. (2001). "Recurrence analysis of hydrologic droughts of differing severity." Journal of Water Resources Planning and Management, 127, 30.

Shukla, S., and Wood, A. W. (2008). "Use of a standardized runoff index for characterizing hydrologic drought."

Sklar, A. W. (1959). "Fonctions de repartition á n dimensions et leurs marges." Publications de l'Institut de Statistique de l'Universit´e de Paris, 8, 229–231.

Smakhtin, V. U. (2001). "Low flow hydrology: a review." Journal of Hydrology, 240(3-4), 147-186.

Wolock, D. M., and McCabe, G. J. (1999). "Simulated effects of climate change on mean annual runoff in the conterminous United States." Journal of the American Water Resources Association, 35(6), 1341-1350.

Wong, G., Lambert, M. F., Leonard, M., and Metcalfe, A. V. (2010). "Drought Analysis Using Trivariate Copulas Conditional on Climatic States." Journal of Hydrologic Engineering, 15, 129.

Yan, J. (2006). "Enjoy the joy of copulas." Department of Statistics and Actuarial Science, University of Iowa.

Yan, J., Kojadinovic, I., and Yan, M. J. (2009). "Package 'copula'."

SPI program: http://drought.unl.edu/monitor/spi/program/spi_program.htm

Soil Survey of Klamath County: http://soildatamart.nrcs.usda.gov/Manuscripts/OR640/0/or640_text.pdf

Monthly observed streamflow from National Weather Service (NWS) Water Resource Forecasts:http://wateroutlook.nwrfc.noaa.gov/point/data?id=KLAO3&valid=y&ver=table&tablelist=obs&dataobs=1906-01-01&dataobsend=2009-12-01&getdata=Get+Data

Statistically Downscaled WCRP CMIP3 Climate Projections: http://gdo-dcp.ucllnl.org/downscaled_cmip3_projections/

United States Historical Climatology Network: http://cdiac.ornl.gov/epubs/ndp/ushcn/ushcn_map_interface.html

Observed reservoir elevation at USGS: http://waterdata.usgs.gov/or/nwis/uv/?site_no=11507001&PAPAmeter_cd=00062,72020

Appendix: Matlab code for Copula

```
% Matlab code for Copula
% data import
dta=xlsread('drought events.xlsx');
D=dta(2:end,2);
S=dta(2:end,3);
I=dta(2:end,4);
arrT=dta(2:end-1,6); %interarrival time
% fit duration to weibull
Ud=wblcdf(D,3.616,1.9456);
% fit severity to gamma
Us=gamcdf(S,2.5196,1.5428);
% fit intensity to gamma
Ui=gamcdf(I,113.7076,0.0103);

%-----Parameter Estimation for bivariate copulas----------
% fit t-copula
[rho1,nu1]=copulafit('t',[Ud Us]);
[rho2,nu2]=copulafit('t',[Ud Ui]);
[rho3,nu3]=copulafit('t',[Ui Us]);
% fit gumbel
param1=copulafit('Gumbel',[Ud Us]);
param2=copulafit('Gumbel',[Ud Ui]);
param3=copulafit('Gumbel',[Ui Us]);

%---------Parameter Estimation for Trivariate copulas------
% fit t-copula
[rho,nu]=copulafit('t',[Ud Us Ui]);

% fit GH-copula
theta1=copulafit('Gumbel',[Ud Us]);
C2=copulacdf('Gumbel',[Ud Us],theta1);
theta2=copulafit('Gumbel',[C2 Ui]);

%%%------------Gumbel-----Bivariate Return period------
d=[0:0.1:8];
s=[0:0.1:12];
Fd=wblcdf(d,3.616,1.9456);
Fs=gamcdf(s,2.5196,1.5428);
EL=23.7333;
for i=1:length(d)
for j=1:length(s)
   Fds(i,j)=copulacdf('Gumbel',[Fd(i) Fs(j)],13);
   test(i,j)=1-Fd(i)-Fs(j)+Fds(i,j);
   Tds(i,j)=EL/test(i,j)/12;%year
end
end
```

```
v=[3 5 10 20 50 100];
subplot(1,2,1)
[C,h]=contour(s,d,Tds,v);
clabel(C,h);
xlabel('Severity');
ylabel('Duration (month)');
title('Return Periods T(D>=d and S>=s) from Gumbel')

%------------
for i=1:length(d)
for j=1:length(s)
   Fds(i,j)=copulacdf('Gumbel',[Fd(i) Fs(j)],13);
   test(i,j)=1-Fds(i,j);
   Tds(i,j)=EL/test(i,j)/12;%year
end
end
subplot(1,2,2)
[C,h]=contour(s,d,Tds,v);
clabel(C,h);
xlabel('Severity');
ylabel('Duration (month)');
title('Return Periods T(D>=d or S>=s) from Gumbel')

%%%----------t---------Trivariate Return period----
% in order to compare single variable return period, using the same
% quantile
q1=0.604; q2=0.802; q3=0.901; q4=0.96; q5=0.98; q6=0.998;
% compare return period=5 yr all

R5=EL/(1-q1-q1-q1+copulacdf('t',[q1,q1],rho1,nu1)+...
copulacdf('t',[q1,q1],rho2,nu2)+copulacdf('t',[q1,q1],rho3,nu3)...
-copulacdf('t',[q1,q1,q1],rho,nu))/12;
% 10 year
R10=EL/(1-q2-q2-q2+copulacdf('t',[q2,q2],rho1,nu1)+...
copulacdf('t',[q2,q2],rho2,nu2)+copulacdf('t',[q2,q2],rho3,nu3)...
-copulacdf('t',[q2,q2,q2],rho,nu))/12;
% 20 year
R20=EL/(1-q3-q3-q3+copulacdf('t',[q3,q3],rho1,nu1)+...
copulacdf('t',[q3,q3],rho2,nu2)+copulacdf('t',[q3,q3],rho3,nu3)...
-copulacdf('t',[q3,q3,q3],rho,nu))/12;
% 50 year
R50=EL/(1-q4-q4-q4+copulacdf('t',[q4,q4],rho1,nu1)+...
copulacdf('t',[q4,q4],rho2,nu2)+copulacdf('t',[q4,q4],rho3,nu3)...
-copulacdf('t',[q4,q4,q4],rho,nu))/12;
% 100 year
R100=EL/(1-q5-q5-q5+copulacdf('t',[q5,q5],rho1,nu1)+...
copulacdf('t',[q5,q5],rho2,nu2)+copulacdf('t',[q5,q5],rho3,nu3)...
-copulacdf('t',[q5,q5,q5],rho,nu))/12;
% 1000 year
R1000=EL/(1-q6-q6-q6+copulacdf('t',[q6,q6],rho1,nu1)+...
copulacdf('t',[q6,q6],rho2,nu2)+copulacdf('t',[q6,q6],rho3,nu3)...
-copulacdf('t',[q6,q6,q6],rho,nu))/12;
```

```
%%%%--------or
r5=EL/(1-copulacdf('t',[q1,q1,q1],rho,nu))/12;
r10=EL/(1-copulacdf('t',[q2,q2,q2],rho,nu))/12;
r20=EL/(1-copulacdf('t',[q3,q3,q3],rho,nu))/12;
r50=EL/(1-copulacdf('t',[q4,q4,q4],rho,nu))/12;
r100=EL/(1-copulacdf('t',[q5,q5,q5],rho,nu))/12;
r1000=EL/(1-copulacdf('t',[q6,q6,q6],rho,nu))/12;

%-----------------------------------
```