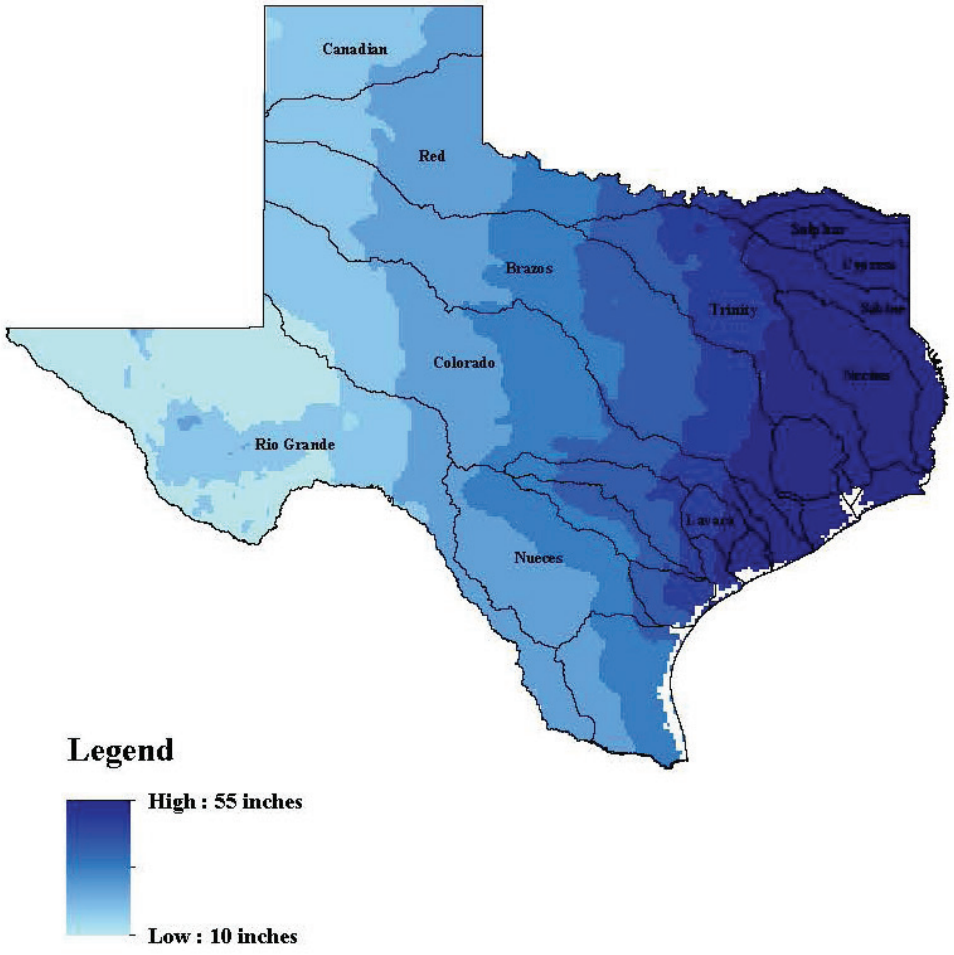
DROUGHT HAZARD AND VULNERABILITY MAPS FOR TEXAS

Texas Water Resources Institute TR-475
May 2015



Texas Water Resources Institute
make every drop count

Deepthi Rajsekhar
Department of Biological and Agricultural Engineering
Vijay P. Singh
Department of Biological and Agricultural Engineering,
Zachry Department of Civil Engineering
Texas A&M University

DROUGHT HAZARD AND VULNERABILITY MAPS FOR TEXAS

by

Deepthi Rajsekhar

Department of Biological and Agricultural Engineering

Vijay P. Singh

Caroline & William N. Lehrer Distinguished Chair in Water Engineering

and

Distinguished Professor, Department of Biological and Agricultural Engineering and Zachry Department of Civil Engineering

Texas A & M University

Submitted to Texas Water Resources Institute

Texas Water Resources Institute
Technical Report-475, May 2015

ABSTRACT

Natural disasters like droughts have a huge socio-economic impact on society. Despite being an important component of mitigation, the concept of vulnerability in association with extreme events has not been explored much. This report presents a systematic approach for the assessment of drought hazard and identification of drought vulnerability indicators pertinent to the state of Texas. A novel drought index known as Multivariate Drought Index (MDI) was used to simultaneously quantify multiple physical forms of drought. A composite risk assessment was then carried out by considering both hazard and vulnerability components. The risk, hazard, and vulnerability components were quantified using standardized indices like Drought Hazard Index (DHI), Drought Vulnerability Index (DVI), and Drought Risk Index (DRI). A suitable classification scheme was adopted for these indices to group regions into classes ranging from low to high. Mapping of DHI, DVI, and DRI classes led to the generation of risk, hazard and vulnerability maps for Texas. The report emphasizes the importance of including vulnerability of the study area in the event of drought while drafting planning measures. Ultimately, the study aims at bridging the gaps existing in the current drought research, which even though substantial, still fails to address some of the issues, and for developing a comprehensive framework for better understanding of droughts in Texas which will help decision makers to formulate a more effective adaptation and mitigation strategy in future.

Keywords: Hazard and Vulnerability maps, Multivariate Drought Index.

ACKNOWLEDGEMENTS

This work is a continuation of the Texas Water Resources Institute (TWRI) report TR-474 titled: "Hydrological Drought Atlas for the State of Texas". We wish to thank the United States Geological Survey (USGS) for providing us with financial support to carry out this research (Grant Number: 2009TX334G).

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

TABLE Page

1. INTRODUCTION

Drought is the leading cause of losses due to natural disasters in the state of Texas. Due to the relative size and population, Texas is likely to suffer more from drought related losses. There has been at least one serious drought in some part of the state during every decade of the twentieth century. Recent climate change impact studies also suggest an increase in the likelihood of occurrence of extreme events like droughts due to a projected rise in temperature and reduction in precipitation events as a result of anthropogenic emissions. A significant part of the economy of Texas is agriculture and livestock, which increases the need for careful water management and planning. The motivation for this study arises from the need for drought research considering the wide varieties of sectors affected by drought, and the fact that there is a higher probability of occurrence of longer, more severe droughts in future. Thus, the need to clearly define and quantify multiple physical forms of drought taking into consideration the socio-economic factors relevant to the study area, and implementing subsequent changes in water resources planning and management is of prime importance.

This report aims to address the following research problems:

1 Because of its widespread economical, societal, and ecological effects, multiple drought types can simultaneously affect a region. Existing univariate drought indices can only consider one drought type at a time. This necessitates the need for a non–linear, multivariate measure that can quantify all the physical forms of drought together.

2 There is a strong correlation between the effect of drought events and the coping ability of the location affected by the event. The risk associated with drought

events, thus, depends on both its magnitude as well as the vulnerability of the society towards droughts. The range of vulnerability towards droughts is very wide since it depends on highly variable factors like social and economic resilience towards disasters. In particular, rural areas lack the resources to prepare for and respond to disasters due to its social and economic setup. A combined hazard and vulnerability assessment which takes into account relevant socio–economic factors is required for an accurate drought risk analysis instead of the usual frequency analysis of drought properties alone.

The report is organized into following sections: section 2 discusses briefly the study area, and section 3 explains data requirements for conducting this study. Section 4 gives a systematic approach to be followed for multivariate drought hazard and vulnerability assessment. This section also includes relevant literature review and explains how the methodology followed in this report tries to fill the gaps in the existing research. Section 5 shows the results of this study, and section 6 discusses the obtained results and the conclusions drawn from the results.

2. STUDY AREA

The study area considered is the state of Texas. Due to its size and geographic location, it is affected by a wide variety of local and regional climatic influences. Texas experiences five distinct climate types ranging from arid to sub–tropic humid zones. The basic climate patterns in Texas are fairly simple: the annual mean temperature increases from north to south, and annual mean precipitation increases from west to east. In Sabine River basin in east Texas, the mean annual rainfall is nearly 60 inches and annual evaporation is less than 70 inches, whereas in Rio–Grande basin in west Texas, mean annual rainfall ranges from 8 to 20 inches and annual evaporation is as much as 105 inches. These climate patterns strongly control the flows of rivers and streams in Texas. Out of the 13 major river basins in Texas that vary greatly in size, shape and stream patterns, east Texas rivers flow year around and most of the west Texas streams flow only part of the year (Bureau of Economic Geology, 1996). The vegetation and land use patterns also vary greatly, with forests in the east, coastal plains in the south to the elevated plateaus and basins in the north and west (Benke and Cushing, 2005). The land surface elevation follows a decreasing trend from west to east, with arid climate zone covering higher elevation areas, whereas most of the sub tropic humid zone and parts of subtropic semi humid zone covers the low lying regions in Texas. Figure 3.1 shows the river basin map of Texas and the precipitation (annual average in inches) gradient within the state.

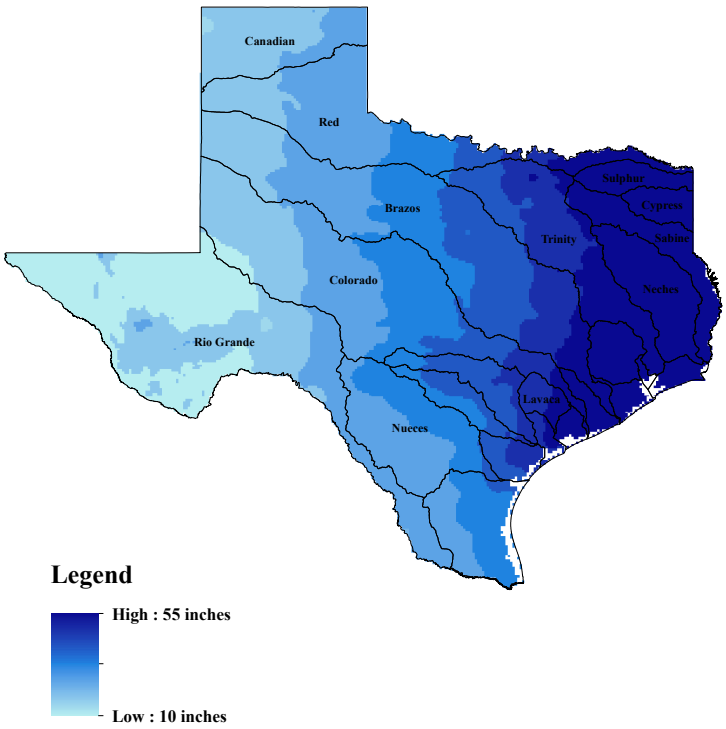


Figure 2.1: River Basin Maps of Texas With Precipitation Pattern

3. DATA

In this study, the authors employed a multivariate approach for quantifying drought. A new drought index known as Multivariate Drought Index (MDI) was used for deriving the drought characteristics. The hydroclimatic variables considered for calculating the Multivariate Drought Index (MDI) include: precipitation (P), runoff (R), evapotranspiration (ET) and soil moisture (SM) for a time period of 1950–2012 on a monthly time scale. Since long term observed records for these variables is sparse, a macro-scale land surface model known as Variable Infiltration Capacity (VIC) model was employed to simulate spatially and temporally continuous R, ET, and SM monthly data for the state of Texas. A brief description of the model processes and the concepts behind it is discussed in section 4. The model generates the aforementioned variables for grids of 1/8^o resolution. All the input files required for running the model also had the specified resolution of 1/8^o. The VIC model overcomes limitations that exist due to the lack of long–term observed databases in the case of soil moisture and evapotranspiration, and non–uniform distribution of flow gaging stations which are mostly concentrated in eastern Texas and poorly distributed in western Texas. A brief discussion on hydroclimatic variables used in this study is now given.

3.1 Precipitation

Precipitation data was obtained from National Oceanic and Atmospheric Administration (NOAA) Cooperative Observer (Co–op) stations. The point data was then gridded to the 1/8^o resolution using the Synergraphic Mapping System (SYMAP) algorithm (Shepard, 1984). This gridded precipitation data was then scaled to match the long–term average of the parameter–elevation regressions on independent slopes

model (PRISM) precipitation climatology. This was done to ensure that the precipitation data used in the study is statistically adjusted to capture the influence of local topography. The scale factor would be the ratio of mean monthly PRISM precipitation for the period 1961–1990 to the unscaled mean monthly observed precipitation for the grid during 1961–1990. For each grid, there would be a different scaling factor for each month.

3.2 Runoff

Monthly runoff values for grids of 1/8° resolution in Texas was obtained using the VIC model. Using a stand-alone routing model (Lohmann et al., 1996, 1998) as a post processing tool, stream flow values were generated at points of interest by performing within-grid cell routing and then channel routing. After routing, the simulated stream flow values were validated against the USGS hydro climatic data network (HCDN) naturalized stream flow data.

3.3 Evapotranspiration

The total evapotranspiration over a grid cell computed by VIC model will have three components: evaporation from the canopy layer (E_c, mm) of each vegetation tile, transpiration (E_t, mm) from each of the vegetation tiles, and evaporation from the bare soil (E_1, mm) (Liang et al. 1994). These individual factors will then be weighted by the respective surface cover area fractions to give the total evapotranspiration for the respective grid cell. The model simulated evapotranspiration values were validated against the actual evapotranspiration data obtained from the Texas ET network and Texas Water Development Board (TWDB).

3.4 Soil Moisture

Since the VIC model does not consider lateral flow between grids in the top two soil layers, the movement of moisture can be characterized by one–dimensional Richard's equation. The soil moisture percentiles were simulated for the top 40 cm of soil. The simulated soil moisture percentiles were validated using soil moisture data obtained from soil climate analysis network (SCAN) stations maintained by Natural Resources Conservation Service (NRCS), and climate reference network stations.

3.5 Vulnerability Indicators

A number of socio-economic factors relevant to the economy of Texas were considered to conduct the drought vulnerability assessment. The factors include population density, municipal water demand, and non–municipal water demand (irrigation, livestock, industrial, mining and steam electric plant). Texas county level data for these indicators were obtained from Texas Water Development Board (TWDB).

The methodology followed for calculation of MDI, drought hazard and vulnerability assessment is explained in the following section.

4. METHODOLOGY

The methodology followed for multivariate drought hazard and vulnerability assessment for Texas is explained as simple steps in the following sub-sections.

(1) Simulation of input hydroclimatic variables required for drought quantification,

(2) Calculation of Multivariate Drought Index (MDI),

(3) Drought hazard assessment and calculation of Drought Hazard Index (DHI),

(4) Drought vulnerability assessment and calculation of Drought Vulnerability Index (DVI), and

(5) Integrated risk assessment which includes hazard and vulnerability aspects.

4.1 Simulation of Hydroclimatic Variables

Since the availability of long-term, spatially continuous data of precipitation, runoff, evapotranspiration, and soil moisture is a pre-requisite for drought analysis, Variable Infiltration Capacity (VIC) model was used to simulate these variables.

4.1.1 Rationale for using VIC Model

This particular model was chosen, since it focuses on simulating hydrological processes relevant to the water and energy balance over the land surface for studying the effects of climate change on natural processes like droughts. Additionally, while assessing the climate change impact on droughts and to understand the behavioral properties of future droughts, there arises the need to simulate the drought indicating variables by coupling VIC with General Circulation Model (GCM) data.

Distinguishing characteristics of VIC model include the sub–grid variability in land surface vegetation classes, sub–grid variability in the soil moisture storage capacity, and drainage from the lower soil moisture zone (base flow) as a nonlinear recession. The VIC model has been well calibrated and applied in a number of large river basins over the continental United States and the globe, and has participated in the Intercomparison of Land Surface Parameterization Schemes (PILPS) project and the North American Land Data Assimilation System (NLDAS), where it has performed well relative to other schemes and to available observations (Bowling et al. 2003a, 2003b, Lohmann et al., 1998). The VIC model has been widely used, particularly for runoff and soil moisture simulation. Abdulla et al. (1996), Nijssen et al. (1997), Lohmann et al. (1998), and Nijssen et al. (2001) used VIC primarily for runoff simulation. Sheffield et al. (2004), Andreadis and Lettenmaeir (2006), Sheffield and Wood (2008), and Shukla and Wood (2008) demonstrated the use of VIC simulated soil moisture and runoff in the context of droughts.

4.1.2 Model Description

The VIC–3L is a large scale land surface model and is used for simulating land–atmosphere fluxes by solving water and energy balance at a daily or sub–daily temporal scale (Liang et al., 1994). The land surface is essentially divided into grids of specified resolution. Each of these cells is simulated independent of each other. Land surface is divided into different vegetation covers in such a way that multiple vegetation classes can exist within a cell. The soil moisture distribution, infiltration, drainage between soil layers, surface runoff, and subsurface runoff are all calculated for each land cover tile at each time step. Then for each grid cell, the total heat fluxes (latent heat, sensible heat, and ground heat), effective surface temperature, and the total surface and subsurface runoff are obtained by summing over all the land cover

tiles weighted by fractional coverage. It should thus be noted that the VIC model does not account for the interflow between the grids. Typically in VIC–3L model, the soil is partitioned into three layers vertically, with variable soil depths and the main soil parameters include hydraulic conductivity, thickness of each soil layer, soil moisture diffusion parameters, initial soil moisture, bulk density and particle density. The vegetation parameters considered by the model include root depth, root fraction, Leaf Area Index (LAI), stomatal resistance, albedo, etc.

4.1.3 Model Processes

The water balance in the VIC model follows the continuous equation for each time-step:

$$\frac{\partial S}{\partial t} = P - E - R \tag{4.1}$$

where $\frac{\partial S}{\partial t}$, P, E, and R, are the changes in water storage, precipitation, evapotranspiration, and runoff, respectively. The major processes simulated by VIC and the concepts behind them are briefly discussed in the sections below.

4.1.3.1 Evapotranspiration

The VIC model considers three types of evaporation: evaporation from the canopy layer of each vegetation tile (E_c), transpiration from each vegetation tile (E_t), and evaporation from the bare soil (E_1) (Liang et al. 1994). Total evapotranspiration over a grid cell is computed as the sum of the above components, weighted by the respective surface cover area fractions.The formulation of the total evapotranspiration is:

$$E = \sum_{n=1}^{N} C_n.(E_{c,n} + E_{t,n}) + C_{N+1}.E_1 \tag{4.2}$$

where C_n is the vegetation fractional coverage for the n^{th} vegetation tile, C_{N+1}

is the bare soil fraction, and $\sum_{n=1}^{N} C_n = 1$.

4.1.3.2 Runoff

The VIC model uses the variable infiltration curve (Zhao et al., 1980) to account for the spatial heterogeneity of runoff generation. It assumes that surface runoff from the upper two soil layers is generated by those areas for which precipitation, when added to soil moisture storage at the end of the previous time step, exceeds the storage capacity of the soil. The formulation of subsurface runoff follows the Arno model conceptualization (Franchini and Pacciani, 1991). The soil moisture and runoff algorithms for the VIC-3L is explained with details in Liang et al. (1994). Total runoff Q is given by:

$$Q = \sum_{n=1}^{N+1} C_n.(Q_{d,n} + Q_{b,n}) \tag{4.3}$$

where $Q_{d,n}$, $Q_{b,n}$, and C_n, are the direct runoff (surface runoff), base flow (subsurface runoff), and the vegetation fractional coverage for the n^{th} vegetation tile, respectively.

4.1.3.3 Soil Moisture Content

The VIC model assumes that there is no lateral flow in the top two soil layers; therefore, the movement of moisture can be characterized by the one–dimensional Richards equation:

$$\frac{\partial \theta}{\partial t} = \frac{\partial D(\theta).\frac{\partial \theta}{\partial z}}{\partial z} + \frac{\partial K(\theta)}{\partial z} \tag{4.4}$$

where θ is the soil moisture content, D(θ) is the soil water diffusivity, K(θ) is the hydraulic conductivity, and z is the soil depth.

4.1.4 Routing Model

Since the grid–based VIC model simulates the time series of runoff only for each grid cell, a stand–alone routing model (Lohmann. et al., 1996, 1998a) is employed to transport grid cell surface runoff and base flow to the outlet of that grid cell and then into the river system. In this routing scheme, the surface runoff simulated by VIC in each grid cell is transported to the outlet of the grid cell using a unit hydrograph approach. Then, by assuming that all runoff exits a cell in a single flow direction, it is routed through the channel using a linearized Saint–Venant equation. In the routing model, water is never allowed to flow from the channel back into the grid cell. Once it reaches the channel, it is no longer part of the water budget scheme.

Because of the absence of long–term observed data for evaporation, soil moisture and runoff for each grid, to evaluate the model simulation results, the routing model will be used as a post–processing tool to produce stream flow at locations for which observed records are available for comparison.

4.1.5 Data Requirements for the Model

VIC requires high quality daily gridded meteorological forcing data as input. Precipitation (mm), maximum and minimum temperature (^oC), and wind speed (m/s) constitute the major input forcing data. In addition to this, the model also requires soil and vegetation data as input. The model was run at its default resolution of $1/8^o$ (Salathe, 2003). Hence, all input files, including forcing files, soil and vegetation parameter files, also have this resolution. This resolution was chosen by also taking into consideration the availability of gridded daily forcing data which was needed to drive the model at $1/8^o$ (Maurer et al., 2002).

The time period of data used was 1949–2012. The year 1949–1950 was considered as the spin up year for the model. The daily precipitation and temperature data were

obtained from National Oceanic and Atmospheric Administration (NOAA) Cooperative Observer (Co–op) stations and National Climate Data Center (NCDC), respectively. Synergraphic Mapping System (SYMAP) algorithm introduced by Shepard (1984) was used to grid the forcing data to match model resolution. The gridded forcing data were subsequently rescaled to match long term averages of Parameter Elevation Regressions on Independent Slopes Model (PRISM) climate data.

The soil characteristics which will not be considered for calibration were taken from gridded $1/8^o$ datasets developed as part of the Land Data Assimilation System (LDAS) project (Mitchell et al. 1999). Within the conterminous United States, these datasets are based on the 1–km resolution dataset produced by the Pennsylvania State University (Miller and White 1998). Soil texture in the LDAS dataset is divided into 16 classes for each of the layers, inferring specific soil characteristics (e.g., field capacity, wilting point, saturated hydraulic conductivity) based on the work of Cosby et al. (1984), Rawls et al. (1993), and Reynolds et al. (2000). These LDAS datasets were used to specify the relevant soil parameters required by the VIC model directly. For the remaining soil characteristics (e.g., soil quartz content), values were specified using the soil textures from the 1–km database, which were then indexed to published parameter values [the primary source was Rawls et al. (1993)], and aggregated to the $1/8^o$ model resolution.

Vegetation parameters needed were also obtained from LDAS. Land cover characterization was based on the University of Maryland global vegetation classification described by Hansen et al. (2000), which has a spatial resolution of 1 km, and a total of 14 different land cover classes. From these global data, the land cover types present in each $1/8^o$ grid cell in the model domain and the proportion of the grid cell occupied by each type were identified (Maurer et al., 2001). The leaf area index (LAI) needed was derived from the gridded ($1/4^o$) monthly global LAI database of

Myneni et al. (1997), which is inverted using the Hansen et al. (2000) land cover classification to derive monthly mean LAIs for each vegetation class for each grid cell.

The model results need to be validated and for this purpose, routing model was used as a post processing tool to produce stream flow at the points of interest. Several United States Geological Survey (USGS) stream gages which come under the Hydro–Climatic Data Network (HCDN) were considered for model validation.

The data needed for the routing scheme include a fraction file, flow direction file, Xmask file, flow velocity and diffusion files, and unit hydrograph file. ArcMap was used for the preparation of these files. The required Digital Elevation Model (DEM) files were obtained from the USGS hydro 1–k datasets.

Figure 4.1 shows a schematic diagram depicting the working of the VIC model.

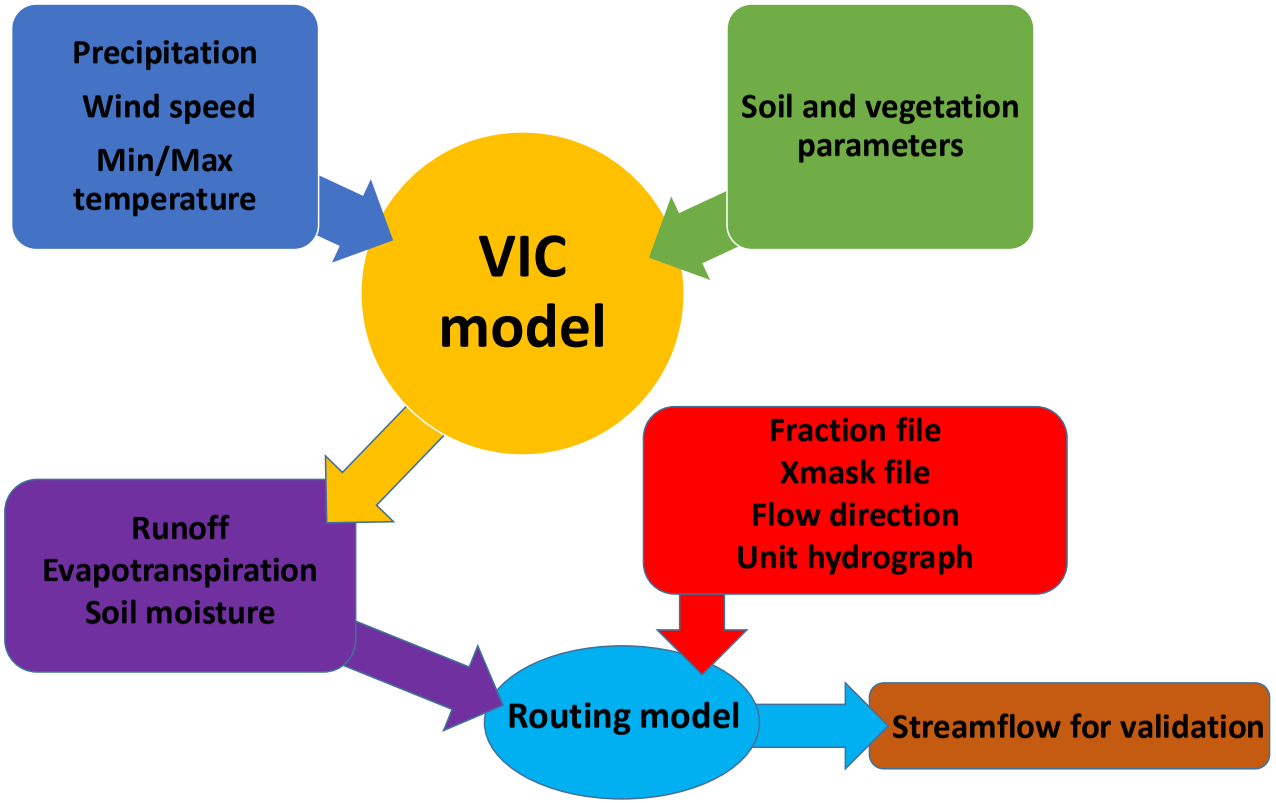


Figure 4.1: Schematic Diagram Showing Working of the VIC Model

The next step is calculation of MDI which is a combination of P, R, ET, and SM.

4.2 Integrated drought quantification using Multivariate Drought Index (MDI)

4.2.1 Rationale for development of MDI

Droughts are classified into four categories: meteorological or climatological, agricultural, hydrological, and socioeconomic (The American Meteorological Society, 1997; Mishra and Singh, 2010). The relationship between hydroclimatic variables and different types of droughts is complex and hence it is difficult to develop an accurate index to quantify and compare droughts.

Currently, there exist a number of drought indices that are used to represent different types of droughts. Many of the commonly used drought indices like Palmer Drought Severity Index (PDSI; Palmer, 1956), Standardised Precipitation Index (SPI; McKee, 1993), Standardized Runoff Index (SRI; Shukla and Wood, 2008), etc. has many limitations. PDSI, in particular, has several limitations like lack of physical meaning, slowness in detecting the onset of drought events, unclear temporal scale and problems with Thornthwaite's method used for calculation of PDSI. SPI and other standardized indices similar to it like SRI, Standardized Stream Flow Index (SSFI; Modarres, 2007), Crop Moisture Index (CMI; Palmer, 1968) for agricultural drought, Vegetation Condition Index (VCI; Kogan, 1995), Climate prediction center (CPC) Soil Moisture Index (SMI; Huang et al., 1996), and Standardized Precipitation Evapotranspiration Index (SPEI; Serrano et al., 2010), etc. considers one specific physical form of drought: hydrological, meteorological, or agricultural. This might not be adequate to get a comprehensive idea of the drought condition, since it is dependent on multiple variables. Hence, in general it can be concluded that the drought status indicated by one drought index might not be consistent with the

findings obtained while using a different drought index.

To overcome these limitations, a group of indices that consider multiple variables to represent drought were developed. The drought monitor developed by Svoboda et al. (2002) considers an Objective Blend of Drought Indicators (OBDI) which is the linear weighted average of several drought indices. Aggregated Drought Index (ADI; Keyantash and Dracup, 2004) comprehensively considers all physical forms of drought through variables like precipitation, stream flow, evapotranspiration, reservoir storage, soil moisture content and snow water content. ADI aggregates all these variables into a single time series through principal component analysis (PCA). However, the use of PCA has several limitations like linearity assumption in data transformation, and the assumption that most information is contained in those directions where input data variance is maximum. These assumptions however need not be always met in reality. Recently, bivariate drought indices have been derived using copulas to quantify the joint behavior of drought types. Kao and Govindaraju (2010) introduced a Joint Drought Index (JDI) using copula for obtaining the joint probabilities while considering precipitation and stream flow. Hao and Agakouchak (2013) introduced Multivariate Standardized Drought Index (MSDI) which uses copula to form joint probabilities of precipitation and soil moisture content. The use of copula for multivariate analysis is, no doubt, highly effective. However, for higher dimensional cases (i.e. more than three variables), this method will not be a feasible choice due to the lack of flexibility in modeling the dependence structure.

Feature extraction technique is an effective approach to aggregate the various drought types into a single index. In particular, the kernel based methods, like the kernel principal component analysis (KPCA) and kernel partial least squares (KPLS), have attracted a lot of attention, particularly in the last decade as an effective non–linear approach for dimensionality reduction. These methods target at

finding projections that maximize the variance of input data in the feature space. However, the method assumes that the maximum information that can be obtained from the input data is oriented along the direction of maximum variance. In this study, Multivariate Drought Index (MDI), a new index formulated by Rajsekhar et al. (2014) is used for comprehensive drought assessment. MDI uses a kernel entropy component analysis (KECA) for extracting a drought index named as multivariate drought index (MDI) from the set of input variables representing the various physical forms of drought. We consider the variables: precipitation (P), runoff (R), evapotranspiration (ET), and soil moisture (SM), thus accounting for all the major elements in the water balance model. The method is essentially a novel feature extraction technique that combines the concept of entropy and KPCA. KECA is superior to KPCA since it has been proved that entropy is a much better measure of information than variance (Dionisio et al., 2007). Entropy is related to the higher order moments of a distribution, and thus, unlike the variance, it can offer a better characterization of the input data, since it uses more information from the probability distribution (Ebrahimi et al.,1999).

The advantages of this approach can be summarized as: (1) It does not make the linearity assumption; (2) final multivariate index is obtained in such a way that it preserves the entropy of the input data, which means it tries to preserve the maximum amount of information of the input data; and (3) unlike KPCA, it does not make the assumption that the maximum information from the input data is oriented along the direction of maximum variance. KPCA essentially preserves only the second order statistics of data set, whereas KECA preserves the higher order statistics also through the use of entropy.

The following subsection explains in detail the formulation of MDI.

4.2.2 Mathematical Formulation of MDI

The mathematical formulation of MDI follows two steps:

(1) The input hydroclimatic variables used for the calculation of MDI were transformed into standard normal variates.

(2) The information theory based feature extraction technique called KECA was then utilized to extract the MDI time series that maximally preserved the entropy of the standardized input dataset.

The following sections discuss in detail the steps involved in the calculation of MDI.

4.2.2.1 Standardization of Input Variables

The primary step involved in the mathematical formulation of MDI is the transformation of each input variable into an index that is a standard normal variate. The procedure followed for the calculation of these standardized indices consists of the identification of a suitable probability distribution fitted to the monthly time series of the variable under consideration, followed by the construction of cumulative density function which is then transformed to standard normal distribution function.

Having approximated to the normal CDF, the respective standardized index for the time series of the given variable was obtained as the standard normal variate with zero mean and unit standard deviation.

Since both precipitation and evapotranspiration were considered as input variables, a combined standardized drought index popularly known as standardized precipitation evapotranspiration index (SPEI) developed by Vincente–Serrano et al. (2010) was used, instead of calculating separate standardized indices for precipitation and evapotranspiration. A differential timeseries $D = P - ET$ formed the basis of SPEI. The D timeseries was fitted to a three parameter log–logistic distribution

to get the cumulative probabilities (Vincente–Serrano et al., 2010). These cumulative probabilities were converted to standard normal variates by following the steps outlined above in order to obtain SPEI. Likewise, the log–normal distribution was used to fit the runoff time series and obtain the CDFs which were subsequently converted to standard normal CDFs and a Standard Runoff Index (SRI) was obtained (Shukla and Wood, 2008; Rajsekhar et al., 2013). A non–parametric approach was used to obtain the empirical probabilities of soil moisture data using the Gringorten plotting position. These empirical probabilities were then transformed to standard normal CDF and a standardized Soil Moisture Index (SMI) was obtained (Hao and Aghakouchak, 2013).

Thus, the input dataset used for formulating MDI consisted of SPEI, SRI and SMI. Note that all of these indices have multiscalar property like SPI. This property was acquired by MDI as well. Thus, an n–month MDI was calculated by considering the n–month totals of each input drought variable.

4.2.3 Spectral Methods for Data Transformation

The approach used for aggregating the input data set into MDI was through a data transformation technique that combined the concept of entropy with kernel principal component analysis. Data transformation techniques are basically used to convert high dimension data into an alternative lower dimensional representation that preserves the structure of the original data. Several data transformation methodologies have been reported in the literature. Spectral methods are the most popular technique used for this purpose, and it is based on the eigen values and eigen vectors of spatially constructed data matrices. Saul et al. (2005) give a detailed review of the spectral methods for data transformation. This subsection discusses Principal Component Analysis (PCA) and its extensions.

Principal Component Analysis (PCA) is a linear dimensionality reduction technique. PCA aims at developing a lower dimensional data representation of the original data in such a way that the transformed data preserves the covariance structure. The input patterns, $X = x_1, ...x_n; x_i \epsilon R^d$ are projected onto an m-dimensional subspace that has maximum variance. The output obtained through PCA are the coordinates of the input data in this subspace, using the directions specified by the top m eigen vectors as the principal axes. The procedure for computation of eigen vectors is explained below.

For matrix X that contains an n–dimensional input data, there exists an eigen vector Y corresponding to each eigen value λ such that:

$$(X - \lambda I_n)Y = 0 \tag{4.5}$$

The eigen values of the covariance matrix represent the variance in the eigen–directions of data space. Hence, the eigen vector corresponding to the largest eigen value is the direction in which the data is most stretched out. The second direction is orthogonal to it and picks the direction of the largest variance in that orthogonal subspace, and so on and so forth. Thus, the number of significant eigen values determines the dimensionality of the subspace that explains most of the original data's variance.

As an advancement from linear methods, a number of nonlinear spectral data transformation methods like kernel PCA have been proposed (Scholkopf et al., 1999). Kernel PCA (KPCA) performs like traditional PCA in a so called kernel feature space which is nonlinearly related to the input space. Suppose we are given a real–valued function $K : R^d X R^d \rightarrow R$ with the property that there exists a map $\phi : R^d \rightarrow H$ into a dot product feature space H such that for all x, $x' \epsilon \mathrm{R^d}$, we have $\phi(x).\phi(x') = K(x, x')$. The kernel function $K(x, x')$ can be viewed as a nonlinear

similarity measure (Scholkopf and Smola, 2002). The covariance matrix in this case can be given as:

$$C = \frac{1}{n}\sum_{i=1}^{n}\phi(x_i)\phi(x_i^T) \tag{4.6}$$

The top m eigenvectors of C are denoted as $(v_\alpha)_{\alpha=1}^{m}$ and their respective eigenvalues as $(\lambda_\alpha)_{\alpha=1}^{m}$. The kernel matrix K may be eigendecomposed as $K = EDE^T$, where D is the diagonal matrix storing all the eigenvalues $\lambda_1, ..., \lambda_n$ and E is a matrix with the corresponding eigenvectors $v_1, ..., v_n$ as columns (Williams, 2002). The lower dimensional outputs of KPCA are thus given by $\phi_{pca} = \sqrt{D_m}{E_m}^T = \sqrt{\lambda_\alpha}.{v_\alpha}^T$. D_m stores the top m eigenvalues of K, and E stores the corresponding eigenvectors as columns. Using the fact that the equivalence between PCA and another linear data transformation method called Metric multidimensional scaling (MDS; Borg and Groenen, 2005) holds for KPCA as well (Williams, 2002), the KPCA outputs can be seen as solution to a minimization problem which is analogous to the mathematical formulation for MDS. The minimization problem for KPCA outputs can be formulated as:

$$\phi_{pca} = D_m^{\frac{1}{2}}E_m^T : \min_{\lambda_1, v_1 ...., \lambda_n, v_n} I^T(K - K_{pca})^2 I \tag{4.7}$$

where $K_{pca} = E_m D_m {E_m}^T$ and I is an $(n \times 1)$ matrix of ones. KPCA shares all the statistical and mathematical properties of PCA with the modification that they become valid over the feature space H rather than R^d.

To better understand the procedure involved in dimensionality reduction using KPCA, a simple step–step scheme is given below:

1 Consider an n–dimensional data matrix $X = x_1, ...x_n$.

2 Subtract the mean from all data points.

3 Choose an appropriate kernel $k(.,.)$ from the various available kernels like polyno-

mial, Gaussian, tanh kernel, etc.

4 Form an $n \times n$ Gram matrix (inner–product matrix), K, which is given by the dot product: $[K(x, x^{'})]$.

5 Form the modified Gram matrix:

$$\bar{K} = (I - \frac{1_{n\times n}}{n})^T K (I - \frac{1_{n\times n}}{n}) \tag{4.8}$$

where $1_{n\times n}$ is an $n \times n$ matrix with all entries equal to 1.

6 Diagonalize $\bar{K}$ to get its eigen values λ_n and eigen vectors v_n.

7 Normalize the eigen vectors as $\frac{v_n}{\sqrt{\lambda_n}}$

8 Retain the top m eigen vectors corresponding to the largest eigen values so that the desired variance is captured.

9 Project the data points on the eigen vectors:

$$\phi = v^T (I - \frac{1_{n\times n}}{n})(\begin{pmatrix} k(x_1, x) \\ . \\ k(x_n, x) \end{pmatrix} - K\frac{1_{n\times 1}}{n}) \tag{4.9}$$

where $1_{n\times 1}$ is an $n\times 1$ matrix with all entries equal to one. Now, use the projections instead of data points.

Note, however, that the KPCA transformation is based on the selection of eigenvectors solely on the basis of size of eigenvalues, and hence it might end up choosing uninformative eigenvectors from an entropy perspective. To overcome this issue, a new data transformation method called kernel entropy component analysis has been employed in this study, which is explained below.

4.2.3.1 Kernel Entropy Component Analysis (KECA)

Recently, it has been shown that there is a close connection between the kernel methods and information theory (Jenssen et al., 2005; 2006). This is a new spectral data transformation method and is fundamentally different from other spectral methods because the data transformation in this method is based on the Renyi entropy of the input space dataset. Jenssen (2010) shows that the Renyi entropy estimator of the input space can be expressed in terms of projections onto the principal axes which are the KPCA axes in kernel feature space. In KECA, the dimensionality reduction is brought about by projecting onto those KPCA axes that contribute to the entropy estimate of the input data set. In general, it need not correspond to the top eigenvalues and eigenvectors of the kernel matrix, as is the case with the KPCA method. Hence, KECA may produce strikingly different results compared to KPCA. The transformed data produced through KECA transformation shows a distinct angular structure, meaning that even nonlinearly related input datasets would be distributed in different angular directions with respect to the origin of the feature space, thus revealing more information about the input dataset.

Entropy, first introduced in the field of information theory by Shannon (1948), is defined for a random variable X as (Lathi,1968):

$$H(X) = -\int_i P(x_i) log_2 P(x_i) dx \tag{4.10}$$

where $P(x_i)$'s are the probabilities associated with the events $X = x_i$'s. $H(X)$ is the marginal entropy of X which measures the information contained in X. Extensions to Shannon's entropy which result in alternate forms of information measures can be found in the literature. The Renyi entropy is a more generalized and flexible form

of Shannon entropy. A general form for the Renyi entropy can be given as:

$$H(X) = \frac{1}{(1-q)}.log \int_i P(x_i)^q dx \tag{4.11}$$

The Renyi entropy becomes Shannon entropy as $q \to 1$. In this study, we focus on Renyi's quadratic entropy wherein $q \to 2$. This is the most heavily used form of Renyi entropy. The Renyi quadratic entropy is given as:

$$H(X) = -log \int_i P^2(x_i) dx \tag{4.12}$$

To estimate the Renyi entropy, we concentrate on the quantity $V(p) = \int_i P^2(x_i) dx$, which can alternately be formulated as expectation w.r.t $P(x)$, and can be calculated using the Parzen window. The Parzen window is a non–parametric density estimation method. Beirlant et al. (1997) introduced this non parametric plug–in entropy estimator. It is known for its consistency and efficiency, and provides a link between information theory and kernel learning. Using the Parzen window, the probability density estimation is given as:

$$\hat{P}(x) = \frac{1}{n} \sum_{x_i \epsilon R_d} K_\sigma(x - x_i) \tag{4.13}$$

where $K_\sigma(x, x_i)$ is the Parzen window or kernel centered at x_i, and σ is the kernel size. K_σ is a Mercer kernel which is continuous, symmetric and positive semi definite. $V(p)$ can then be invoked using the sample mean approximation of expectation operator

as (Jenssen, 2010):

$$\begin{aligned}\hat{V}(p) &= \frac{1}{n}\sum_{x_i \epsilon R^d} \hat{P}(x_i) \\ &= \frac{1}{n}\sum_{x_i \epsilon R^d} \frac{1}{n}\sum_{x_i \epsilon R^d} K_\sigma(x - x_i) \\ &= \frac{1}{n^2} I^T K I\end{aligned} \tag{4.14}$$

where I is an $(n \times 1)$ matrix of ones, and K is the kernel matrix. The kernel matrix K can be eigendecomposed and eq. 4.14 can thus be rewritten as (Jenssen, 2010):

$$\hat{V}(p) = \frac{1}{n^2}\sum_{i=1}^{n}(\sqrt{\lambda_i} v_i^T I)^2 \tag{4.15}$$

The ψ term that denotes $(\sqrt{\lambda_i} v_i^T I)^2$ contributes to the total entropy of input data. Certain eigenvectors contribute more towards the entropy than others. Eq. 4.15 reveals that the Renyi entropy estimator is composed of projections onto all the KPCA axes, wherein the projection onto the i^{th} principal axis is given by $(\sqrt{\lambda_i} v_i^T)$. Only a principal axis with $\lambda_i \neq 0; v_i^T I \neq 0$ contributes to the entropy estimate. Hence, a large eigenvalue λ_i simply does not guarantee that the principal axis contributes to the entropy estimate at all. The KECA transformation for an n dimensional data into a k dimensional subset is done by projecting the feature space ϕ onto a subspace ϕ_k spanned by the k kernel PCA axes that contribute most to the entropy estimate of the input data. Mathematically, this transformation can be denoted by:

$$\phi_{eca} = \phi_k^T \phi = D_k^{\frac{1}{2}} E_k^T \tag{4.16}$$

where D_k is the diagonal matrix containing the eigenvalues $\lambda_1, \lambda_2, ...\lambda_k$ that contribute the most towards the entropy of the input dataset, and E_k contains the corresponding eigenvectors $v_1, v_2, ...v_k$ as columns. Hence, analogous to eq. 4.7, the KECA outputs could be formulated as the solution to a minimization problem:

$$\phi_{eca} = D_k^{\frac{1}{2}} E_k^T : \min_{\lambda_1, v_1 ...., \lambda_n, v_n} \frac{1}{n^2} I^T (K - K_{eca}) I \tag{4.17}$$

where $K_{eca} = E_k D_k E_k^T$. The entropy estimate of the subspace ϕ_{eca} is given as:

$$\hat{V_k(p)} = \frac{1}{n^2} I^T K_{eca} I \tag{4.18}$$

The MDI time series obtained for each grid cell within Texas was then used for calculating the drought properties.

4.2.4 Derivation of Drought Properties Using Theory of Runs

A drought event is characterized by severity, duration and areal extent (Mishra and Singh, 2010). For any drought event, the cumulative deficit of the variable of interest during the drought event is defined as drought severity. Drought duration is the time between the onset and the end of a drought event. Drought magnitude is the average deficit per unit duration. In this study, drought duration and severity were considered.

The theory of runs was used for deriving drought characteristics from the runoff time series. This method has been widely used in the field of hydrology. Yevjevich et al. (1967), Rodriguez-Iturbe (1969), Saldarriaga and Yevjevich (1970), Millan and Yevjevich (1971), Guerrero–Salazar and Yevjevich (1975), and Sen (1976, 1977) are among the first few who applied the runs theory in hydrology. A run is defined as a portion of time series of drought variable X_t in which all values are either above

or below a threshold level X_0. Accordingly, it can be called a positive or a negative run. The threshold level may be constant or it may vary with time. Thus, the drought characteristics essentially depend upon the threshold chosen (Mishra and Singh, 2010). Figure 4.2 depicts the properties for a drought event using the theory of runs.

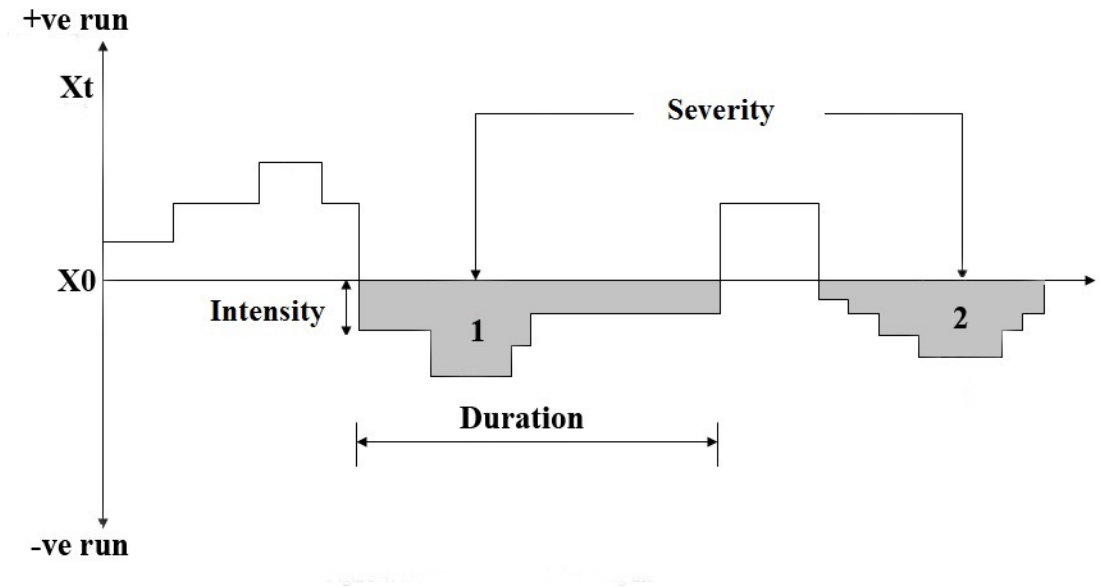


Figure 4.2: Drought Characteristics Using Theory of Runs

The drought classification for MDI values is given in Table 4.1. MDI followed the same drought classification as that of its constituent indices, like SPI, SRI, etc.

Table 4.1: MDI Drought Classification

MDI value	**Classification**
2.0 or more	Extremely wet
1.5 to 1.99	Very wet
1.0 to 1.49	Moderately wet
-0.99 to 0.99	Near normal to mildly dry
-1.0 to -1.49	Moderately dry
-1.5 to -1.99	Severely dry
-2.0 or less	Extremely dry

Following the classification table for MDI values, a threshold of -0.99 was chosen. Any value of MDI below this was considered as a drought event. The derived drought properties based on MDI was then used for an integrated risk assessment.

4.3 Integrated Hazard and Vulnerability Assessment

Although many studies have been carried out on drought risk assessment, only a few of them considered incorporating both hazard and vulnerability factors together (Wilhemi and Wilhite, 2002; Shahid and Behrawan, 2008; Bin et al., 2011; Kim et al., 2013; Zhang et al., 2014). While properties like severity, duration, frequency and spatial extent of the drought are important, one should also consider the socioeconomic ability of the region to cope with the drought event (Shahid and Behrawan, 2008). A comprehensive risk analysis, thus, bridges the gap between impact assessment and policy formulation by directing attention to underlying causes of vulnerability rather than to its result, i.e. the negative impacts which follow triggering events, such as a drought (Ribot et al. 1996). Some of the studies that incorporate socioeconomic elements while assessing droughts are listed below.

Knutson et al. (1998) introduced a step–by–step process for identifying actions that can be taken to reduce potential drought–related impact, taking into consideration the underlying environmental, economic, and social causes of the impacts.

Wilhelmi and Wilhite (2002) conducted a GIS based agricultural drought vulnerability study considering key factors like soil and land use, irrigated cropland, and agroclimatic data. Fontaine and Steinmann (2009) integrated stakeholder data in the vulnerability assessment method. Shahid and Behrawan (2008) introduced a systematic three step methodology for meteorological drought risk assessment framework that incorporates hazard and vulnerability. Bin et al. (2011), Kim et al. (2013), and Zhang et al. (2014) are some of the application based studies which followed the framework introduced by Shahid and Behrawan (2008). Cheng and Ping–Tao (2010), Yuan et al. (2013), and Wu et al.(2011) improved upon the weighting scheme for vulnerability assessment through methods like Analytic Hierarchy Process (AHP) and fuzzy clustering algorithms.

In most of the above mentioned studies, the focus was either on agricultural drought or meteorological drought. Also, only drought severity was considered for hazard assessment. In this study, an integrated approach was followed, which simultaneously considers several drought forms like meteorological, agricultural, and hydrological droughts. The hazard and vulnerability assessment was carried out by considering severity and duration together, so as to capture the joint behavior of droughts.

Risk assessment of drought, thus entails three components: (Singh, 2013): (1) Hazard assessment which is defined as the product of magnitude and the frequency of the event corresponding to that magnitude, (2) Vulnerability assessment which is a measurement of the sensitivity of the exposed system, and (3) Risk assessment incorporating both hazard and vulnerability factors, which is defined by the relationship

$Risk = Hazard \times Vulnerability.$

4.3.1 Hazard Assessment using Drought Hazard Index (DHI)

Hazard can be defined as the probability of occurrence of a potentially damaging phenomenon. It is measured as the product of magnitude and the associated frequency of occurrence of event, and is an indicator of the potential threat.

The drought hazard assessment was carried out using the following steps:

(1) The drought properties of severity and duration were calculated using MDI, as explained in previous section. Monthly MDI values were calculated for the time period 1950–2012 for each $1/8^0$ grid within Texas. The classification of drought into mild, moderate, and severe based on the MDI values is given in Table 4.1.

(2) The joint behavior of drought severity and duration was modeled using a copula. The steps involved in deriving the joint probability of severity and duration using copula is briefly explained below:

   (a) Fit suitable distributions to severity and duration data. Standard goodness of fit statistics may be chosen to decide on the best fitting marginal distributions for severity and duration, respectively. In drought analysis, the two most commonly used continuous distributions are exponential and gamma for fitting the drought duration (Zelenhastic and Salvai, 1987) and drought severity (Shiau, 2006), respectively. This can be used as a simple guide while choosing appropriate marginal distributions for severity and duration.

   (b) Build joint distributions for drought severity and duration using an appropriate copula from among the Archimedean, Elliptical, and Extreme value copula families. Copulas are useful for building joint distributions from marginal distributions belonging to different families. It can be explained using a simple mathematical expression, as shown below. The probability of occurrence

of a drought event having the given severity and duration found from the joint distribution is given as:

$$F_{S,D}(s,d) = C[F_S(s)F_D(d)] \tag{4.19}$$

where $F_D(d)$ is the CDF of drought duration, $F_S(s)$ is the CDF of drought severity, and $F_{S,D}(s,d)$ is the joint CDF of drought severity and duration which was derived using the unique copula, C. Visualization of observed and simulated data through scatter plots were used as a guide for choosing the appropriate copula. Use of standard goodness of fit statistics may also be employed for choice of best fit copula, if necessary.

(3) The joint probability, p, was then standardized by taking an inverse normal, thus obtaining Standardized Joint Probability of occurrence (SJP):

$$SJP = \psi^{-1}(p) \tag{4.20}$$

where ψ is the standard normal distribution function.

(4) A weighting system based on the cumulative distribution function of SJP is given in Table 4.2, and the classification followed by SJP was the same as all other standard drought indices (McKee et al., 1993). Based on this, the "mild" (M) classification will have an SJP value ranging from 0 to -0.99 and was given a weight of 1; the "moderate" (MO) category which ranges from -1 to -1.49 was given a weight of 2; the "severe" (S) category which ranges from -1.5 to -1.99 was weighted as 3; and the "extreme" (E) category that ranges from -2 and was weighted as 4.

Table 4.2: Weighting Scheme for Hazard Assessment

SJP value	**Classification**	**Weights assigned**
-0.99 to 0.99	Near normal or mild (M)	1
-1.0 to -1.49	Moderate (MO)	2
-1.5 to -1.99	Severe (S)	3
-2.0 or less	Extreme (E)	4

(5) After fixing the weights for various categories, each weighted category was further split into four ratings ranging from 1 to 4. This was done using Jenks natural break optimization which divided the actual occurrence probabilities calculated for all the grids that lie within the same planning region identified by Rajsekhar et al. (2012) into four ratings. The Jenks method seeks to minimize each class's average deviation from the class mean, while maximizing each class's deviation from other groups. Hence, the method seeks to minimize within–class variance and maximize variance between classes. Figure 4.3 shows the weight and rating system developed for calculation of an Aggregrated Drought Hazard Index (DHI) based on the cumulative distribution function of SJP.

(6) The DHI was then calculated by combining the weights and ratings of various categories, and is given as:

$$DHI = (M_r \times M_w) + (MO_r \times MO_w) + (S_r \times S_w) + (E_r \times E_w) \tag{4.21}$$

where M_r, MO_r, S_r, and E_r represent the ratings of M, MO, S, and E categories, respectively, and M_w, MO_w, S_w, and E_w represent the weights of M, MO, S, and E categories. The DHI values were then rescaled to a 0–1 range, and evenly classified into four groups as given in Table 4.3.

Table 4.3: DHI Classification for Hazard Assessment

DHI value	Classification
0 to 0.25	Low
0.25 to 0.50	Moderate
0.50 to 0.75	High
0.75 to 1.00	Very High

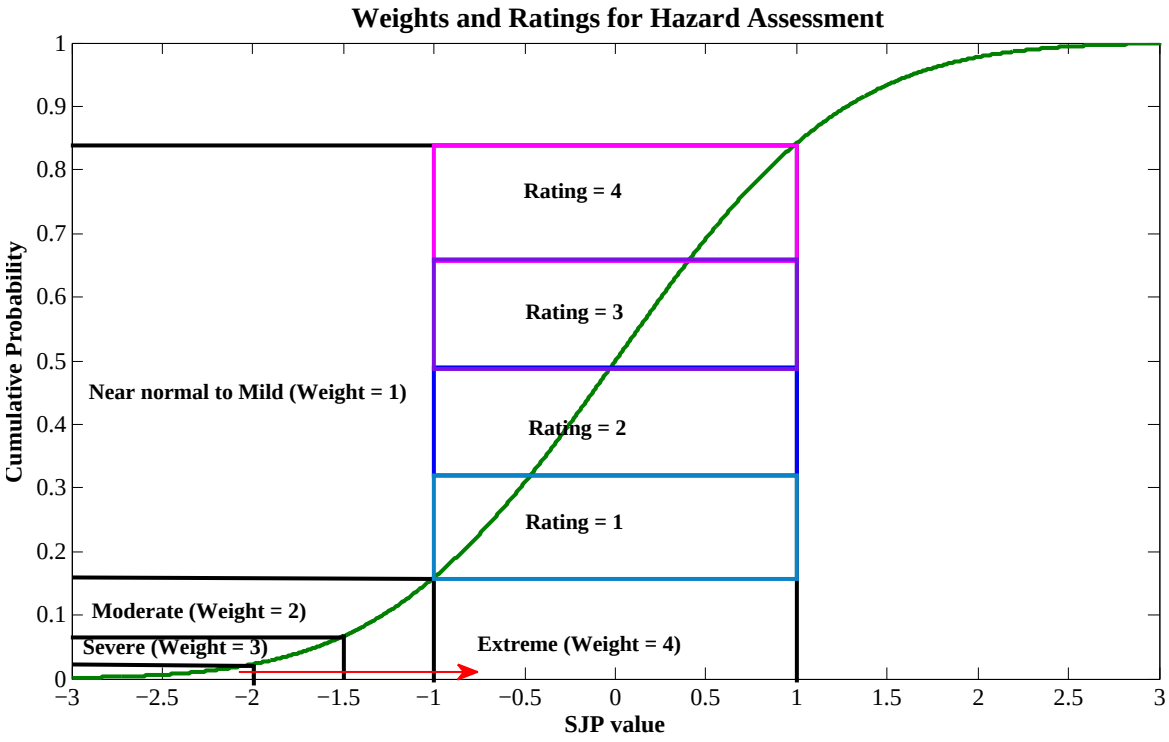


Figure 4.3: Weight and Rating Scheme Based on Cumulative Probability Function for SJP

4.3.2 Drought Vulnerability Assessment

Indicators of vulnerability depends on the region under consideration and are generally complex to objectively assess. Several studies have listed potential vul-

nerability indicators for studies related to climate change impact on water resources (Brooks et al., 2005; Metzger et al., 2006). Depending on the study area, the factors relevant to the location must be chosen. For this study, the vulnerability indicators were chosen keeping in mind that it is a developed economy. Hence, after careful consideration of the availability of reliable data on future projections of vulnerability indicators, the following were chosen: Population Density (PD), Municipal Water Demand (MW), Irrigation Water Needs (IW), Livestock Water Demand (LW), Industrial Water Demand (INW), Water Demand from Mining sector (MNW), and Water needs for Steam electric plants (SW). Population Density (PD) defines the number of persons per km^2. From a human and economic perspective, this is a very important factor determining the degree of calamity associated with a disaster. Higher the PD of the affected area, higher will be the vulnerability of the region, since more people will be affected. The rest of the factors were chosen in light of the most active and important economic sectors in Texas (Agriculture, Mining, Manufacturing Industries). These vulnerability indicators were divided into four classes using the Jenks natural break method. Then, each class of the indicators was given a rating on a scale of 0–1, with lower class values having a lower rating and vice versa. The Aggregate Drought Vulnerability Index (DVI) was then calculated as:

$$DVI = \frac{PD + MW + IW + LW + INW + MNW + SW}{7} \tag{4.22}$$

The DVI which consists of 7 component indicators lie within the range of 0–1. Each of the components was given equal weightage and the aggregate DVI was obtained as a simple average of individual factors. Based on the value of DVI, vulnerable regions were classified under four classes. Table 4.4 shows the four vulnerability classes and the corresponding values of DVI. Areas that fall under high vulnerability group

typically sustain more damage due to drought than other regions, thus entailing the need for a more careful crop and water management scheme.

Table 4.4: DVI Classification for Vulnerability Assessment

DVI value	**Classification**
0 to 0.25	Low
0.25 to 0.50	Moderate
0.50 to 0.75	High
0.75 to 1.00	Very High

4.3.3 Drought Risk Assessment

In this study, the drought risk assessment is conducted by combining the hazard and vulnerability assessment. Typically, the Drought Risk Index (DRI) is calculated as:

$$DRI = DHI \times DVI \tag{4.23}$$

where DHI and DVI are aggregated drought hazard and vulnerability indices. If either DVI or DHI is 0, there will be no risk associated with that drought event. A higher value of either DVI or DHI will result in increased risk from the drought event. Thus, it can be seen that more information about the risk associated with a drought event can be obtained by linking hazard and vulnerability.

5. RESULTS AND DISCUSSION

The results and discussions presented in the subsequent sections follow the order listed below: (1) Drought quantification using MDI, (2) Drought hazard assessment using DHI, (3) Drought vulnerability assessment using DVI, and (4) Drought risk assessment as a combination of hazard and vulnerability.

5.1 Drought Quantification using MDI

The components required for calculation of the drought index MDI, was obtained using VIC model simulations. The model was calibrated and validated to check that the simulations were a good representation of the observed values. The performance of MDI was also compared against PDSI, which is the most commonly used drought index in United States.

5.1.1 Calibration and Validation of VIC Model

Liang et al. (1994) suggests a set of recommended parameters and the plausible range of values for each of them that can be used for calibrating the model. The recommended parameters and the plausible range of values for each of them are given in Table 5.1.

Table 5.1: Model Parameters for Calibration of VIC Model

Soil parameter	**Unit**	**Range of values**
Infiltration shape parameter (binf)	None	0-0.4
Maximum sub-surface flow rate (Dsmax)	mm/day	0-30
Fraction of Dsmax when non linear flow starts (Ds)	None	0-1
Depth of second soil layer (D2)	meter	0.1-1.5
Depth of third soil layer (D3)	meter	0.1-1.5
Fraction of maximum soil moisture when non linear flow starts (Ws)	None	0-1

In this study, six soil parameters were considered for calibration purposes. As

regards the calibration of the routing model, the suggested parameters for adjustment included velocity and diffusivity. The model developers are less specific about the routing model calibration as compared to the VIC model calibration. Application based studies focusing on monthly discharge from large basins have shown that it does not require high accuracy in the routing model parameters. Hence, while parameters like flow direction and contributing fraction can be obtained from DEM, for other parameters, like flow velocity and diffusivity, physically reasonable values were chosen without further calibration (www.hydro.washington.edu). Since only monthly stream flows were required, diffusivity and velocity values of 800 m^2/s and 1.5 m/s were deemed acceptable (Lohmann et al., 1996, 1998).

The VIC model calibration was performed using a random auto–start simplex method program. The method attempts to minimize the differences between simulated and observed discharge records. The magnitude of the differences was given a value by computing the Nash–Sutcliffe R^2 coefficient where 1 represents a perfect match. The cases wherein the coefficient had negative values were not be considered. Having decided on the model parameters required for calibration, the random autostart simplex method starts by selecting a large number of random parameter sets, solving for the model output at each one. It then selects the best set of parameters from the randomly generated sets and uses those to start the simplex minimization algorithm. To explain the procedure simply, if we have n parameters for calibration, the simplex method tries to corral the minimum within a geometric shape with $n+1$ apexes. Once a minimum has been obtained, the algorithm begins to minimize the volume of the simplex, until all of its apexes are within a specified tolerance of each other. The local minima problem inherent to the simplex method is countered by incorporating the random autostart process, wherein in each case a new set of initial parameters is produced. By running the algorithm several times, the global

parameters are eventually located.

In this study, the simplex method was applied using random autostart populations of 75–100 parameter sets. The entire cycle was repeated from 5–10 times for each sub–basin. Each autostart yielded different R^2 values (usually within +/- 0.1) and different parameter sets. The simplex method stops when the convergence occurs for R^2 values with a tolerance of 0.001. If the calibrated parameter values do not seem to be physically viable even if they give a high R^2 value, the user may employ visualization of the simulated stream flow and check how well it captures the physical response of the basin.

In the case of stream flow, the routing model was used to route the grid–cell runoff to the selected station locations. Results from the routing model were aggregated to a monthly scale and compared with the observed gage data. The model directly simulates evapotranspiration which was compared against the observed ET values. The soil moisture averaged for each month at 40 cm soil depth was compared with the observed soil moisture obtained from SCAN stations and climate reference network. The two performance criteria selected were correlation coefficient and the Nash–Sutcliffe (N–S) efficiency. These are defined as:

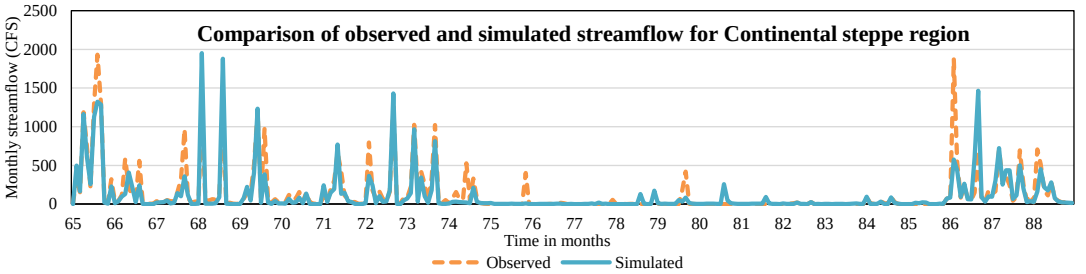
$$r = \frac{M\sum_{i=1}^{M}(S_i O_i) - \sum_{i=1}^{M} S_i \sum_{i=1}^{M} O_i}{\sqrt{(M\sum_{i=1}^{M} {S_i}^2 - (\sum_{i=1}^{M} S_i)^2)(M\sum_{i=1}^{M} {O_i}^2 - (\sum_{i=1}^{M} O_i)^2)}} \tag{5.1}$$

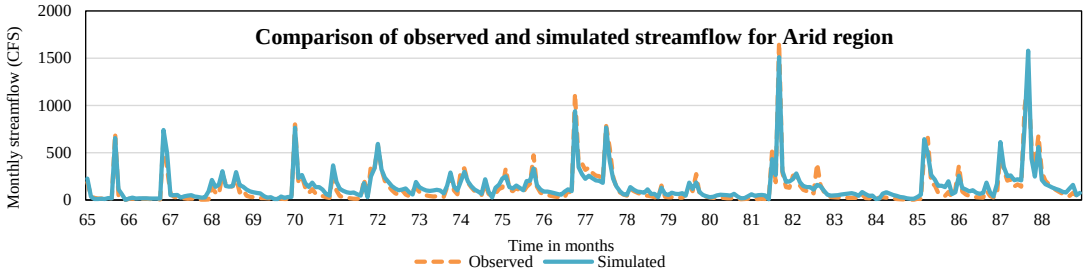
$$\textit{N-S efficiency} = 1.0 - \frac{\sum_{i=1}^{M}(O_i - S_i)^2}{\sum_{i=1}^{M}(O_i - \bar{O})^2} \tag{5.2}$$

where M is the number of months, S_i is the simulated stream flow for the i^{th} month, O_i is the observed stream flow for i^{th} month. The logarithmic transforms of observed and simulated values was used to calculate the N–S value so as to reduce sensitivity to extreme values.

A higher value of correlation coefficient and the Nash–Sutcliffe (N–S) efficiency indicate good performance of the model. The closer the value is to 1, the more accurate the model is.

Table 5.2 gives the locations of validation stations for input variables, and the time period considered for validation. The chosen validation time period was dependent on the availability of observed data, which is sparse particularly in the case of soil moisture. Figures 5.1, 5.2, and 5.3 show the time series comparison between observed and simulated values of stream flow, evapotranspiration and soil moisture, respectively, at the selected locations in various climate zones of Texas. Table 5.3 gives the values of performance statistics at each location considered for validation.





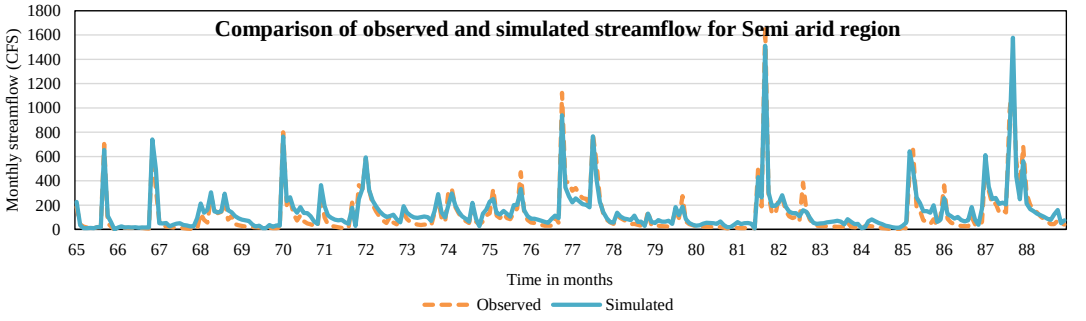


Figure 5.1: Comparison of Simulated and Observed Stream Flows at Selected Stations

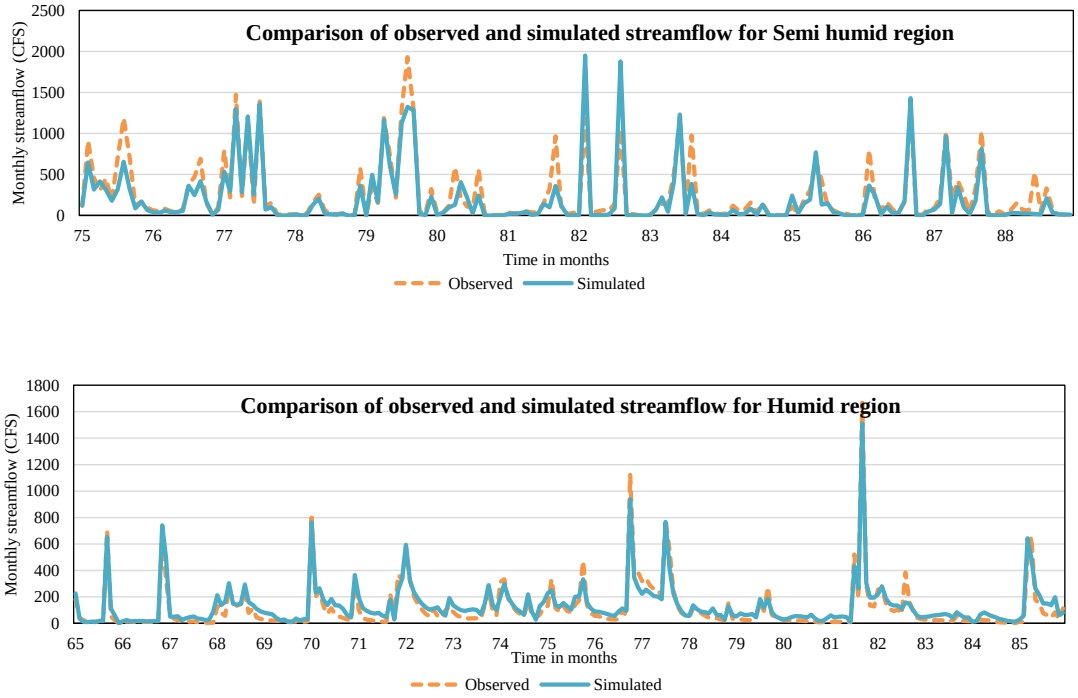


Figure 5.1: Continued

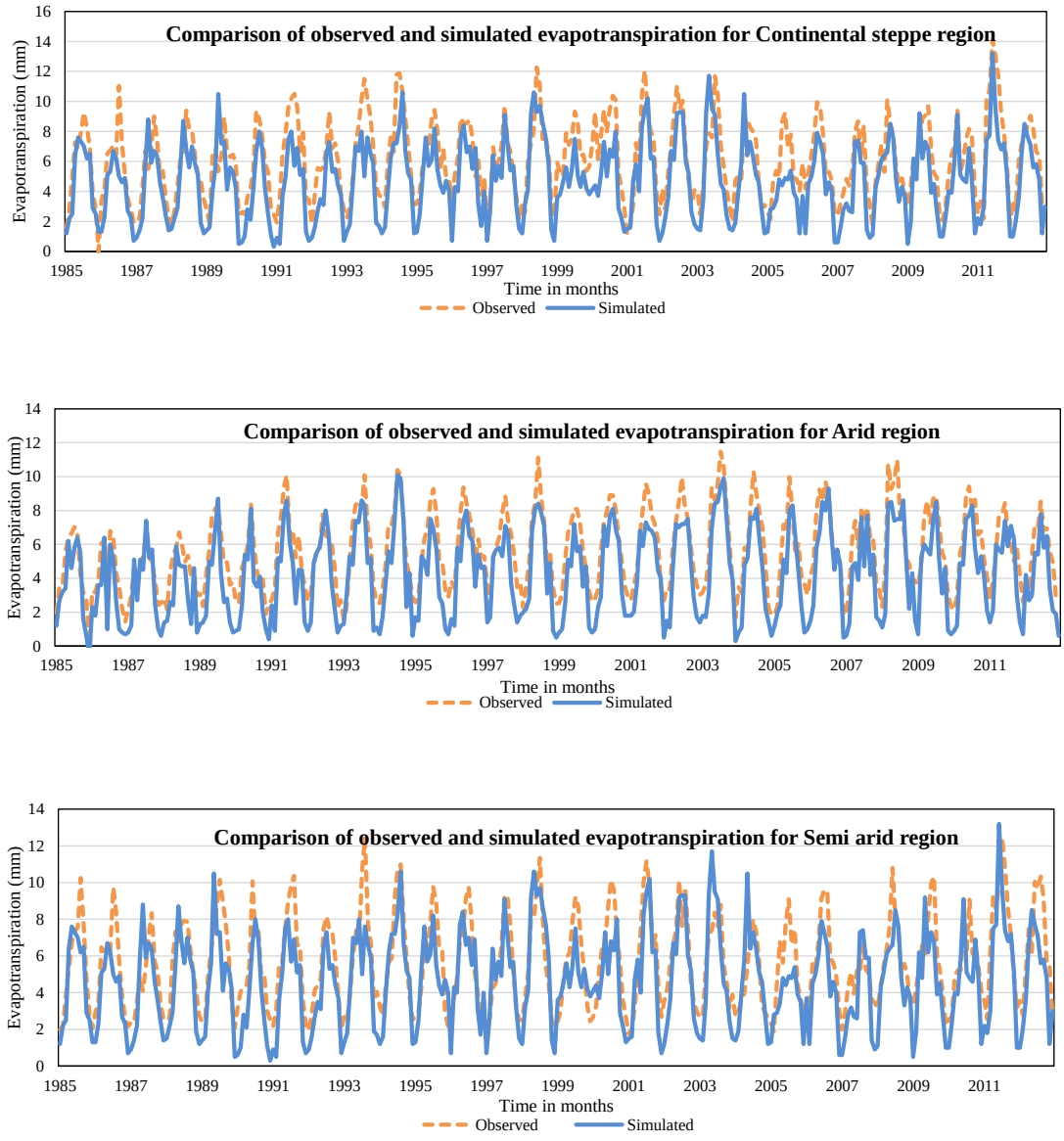
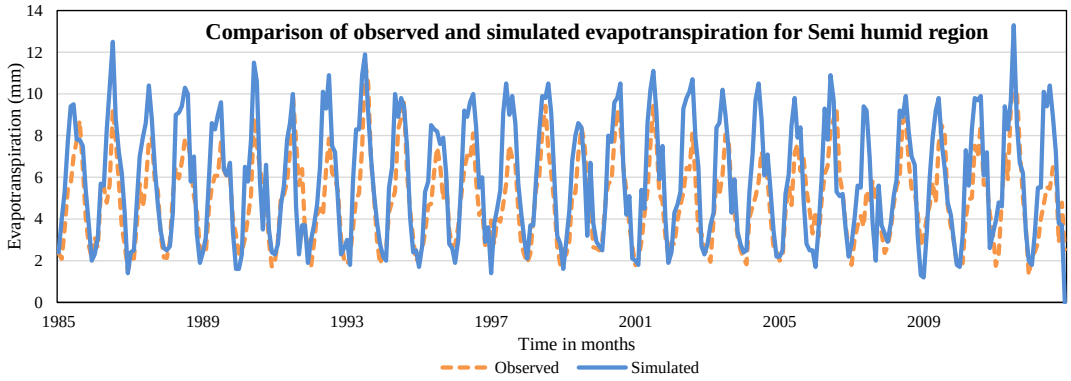


Figure 5.2: Comparison of Simulated and Observed Evapotranspiration at Selected Stations



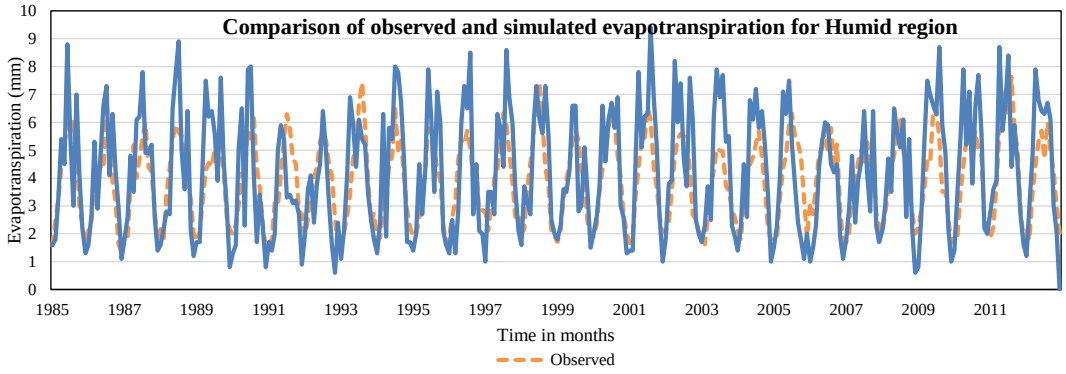


Figure 5.2: Continued

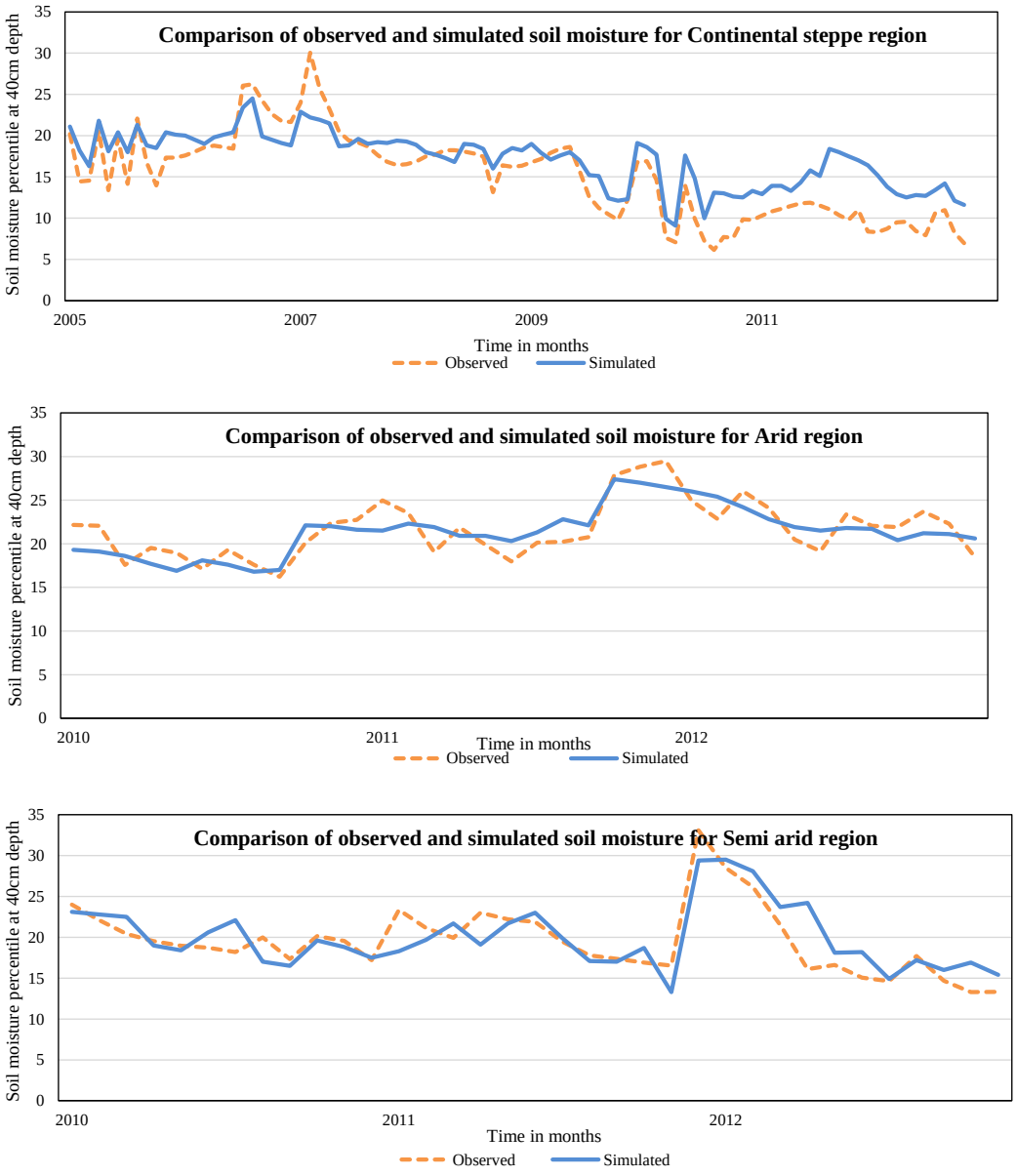
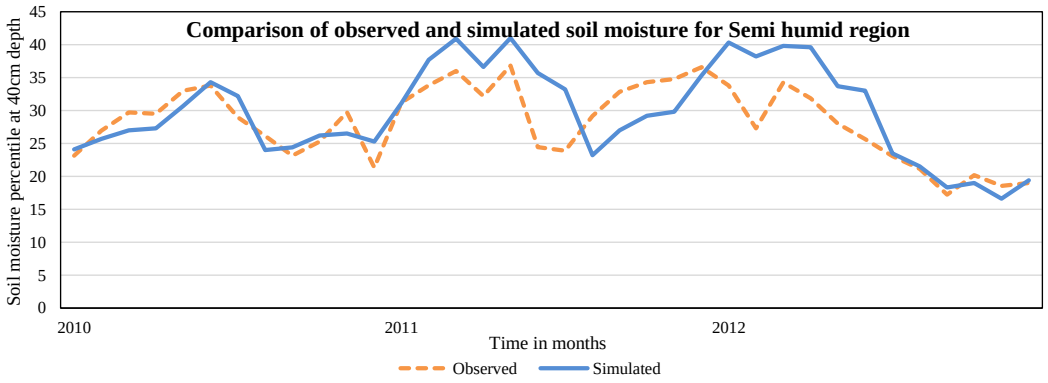


Figure 5.3: Comparison of Simulated and Observed Soil Moisture at Selected Stations



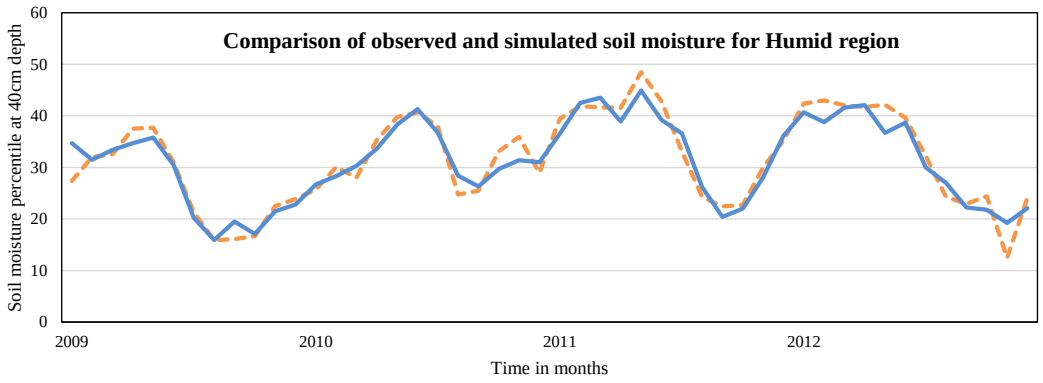


Figure 5.3: Continued

Table 5.2: Details of Validation Stations and Time Periods

Variable	Climate zone	Data source	Station name	Latitude	Longitude	Validation period
Streamflow (CFS)	Continental	HCDN	Spring creek near Spring	35.47	-101.88	1965-1988
Streamflow (CFS)	Arid	HCDN	Frio river near Derby	31.44	-103.47	1965-1988
Streamflow (CFS)	Semi-arid	HCDN	Nueces river below Uvalde	28.5	-99.68	1965-1988
Streamflow (CFS)	Semi-humid	HCDN	Mill creek near Bellville	32.63	-101.29	1964-1974
Streamflow (CFS)	Humid	HCDN	Big cow creek near Newton	32.76	-95.46	1965-1988
Evapotranspiration (mm)	Continental	ET network, TWDB	Lubbock	33.56	-101.88	1985-2012
Evapotranspiration (mm)	Arid	ET network, TWDB	Pebble Hills park	31.78	-106.32	1985-2012
Evapotranspiration (mm)	Semi-arid	ET network, TWDB	Uvalde	29.21	-99.78	1985-2012
Evapotranspiration (mm)	Semi-humid	ET network, TWDB	McKinney	33.19	-96.63	1985-2012
Evapotranspiration (mm)	Humid	ET network, TWDB	Overton	32.27	-94.97	1985-2012
Soil moisture (percentile)	Continental	SCAN	Lehman	32.63	-102.45	2005-2012
Soil moisture (percentile)	Arid	Climate reference network	Monahans	31.58	-102.89	2010-2012
Soil moisture (percentile)	Semi-arid	Climate reference network	Edinburgh	26.3	-98.16	2010-2012
Soil moisture (percentile)	Semi-humid	Climate reference network	Austin	30.25	-97.75	2010-2012
Soil moisture (percentile)	Humid	Climate reference network	Palestine	31.75	-95.64	2009-2012

Table 5.3: Details of Validation Stations and Time Periods

Variable	Climate region	Correlation coefficient	N-S efficiency
Streamflow (CFS)	Continental	0.8	0.54
Streamflow (CFS)	Arid	0.93	0.77
Streamflow (CFS)	Semi-arid	0.96	0.81
Streamflow (CFS)	Semi-humid	0.85	0.75
Streamflow (CFS)	Humid	0.88	0.74
Evapotranspiration (mm)	Continental	0.85	0.71
Evapotranspiration (mm)	Arid	0.92	0.79
Evapotranspiration (mm)	Semi-arid	0.82	0.65
Evapotranspiration (mm)	Semi-humid	0.82	0.68
Evapotranspiration (mm)	Humid	0.77	0.64
Soil moisture (percentile)	Continental	0.89	0.58
Soil moisture (percentile)	Arid	0.82	0.6
Soil moisture (percentile)	Semi-arid	0.81	0.8
Soil moisture (percentile)	Semi-humid	0.76	0.72
Soil moisture (percentile)	Humid	0.96	0.87

It can be seen from Table 5.3 that the correlation coefficients for stream flow validation ranges from 0.80 to 0.96, which means the model is capable of explaining 64% to 92% of variability in the observed data. The N–S efficiency values range from 0.54–0.81. Since an N–S value of 1 corresponds to a perfect match and 0 corresponds to the situation where simulated values match the mean of observed values, a value of 0.5 may be considered to represent a mediocre model performance. Hence, from the values obtained for the model at all the 5 stations, it can be seen that the model performance is satisfactory.

In the case of evapotranspiration, the correlation coefficients fall within the range of 0.71 to 0.92, which means the model is explaining 50% to 85% of variability in the observed data. The N–S efficiency values for ET lie within the range of 0.64–0.79. Although the model replicates ET values well within acceptable limits, it seems to overpredict the values slightly, when it comes to humid climate zone, which in reality

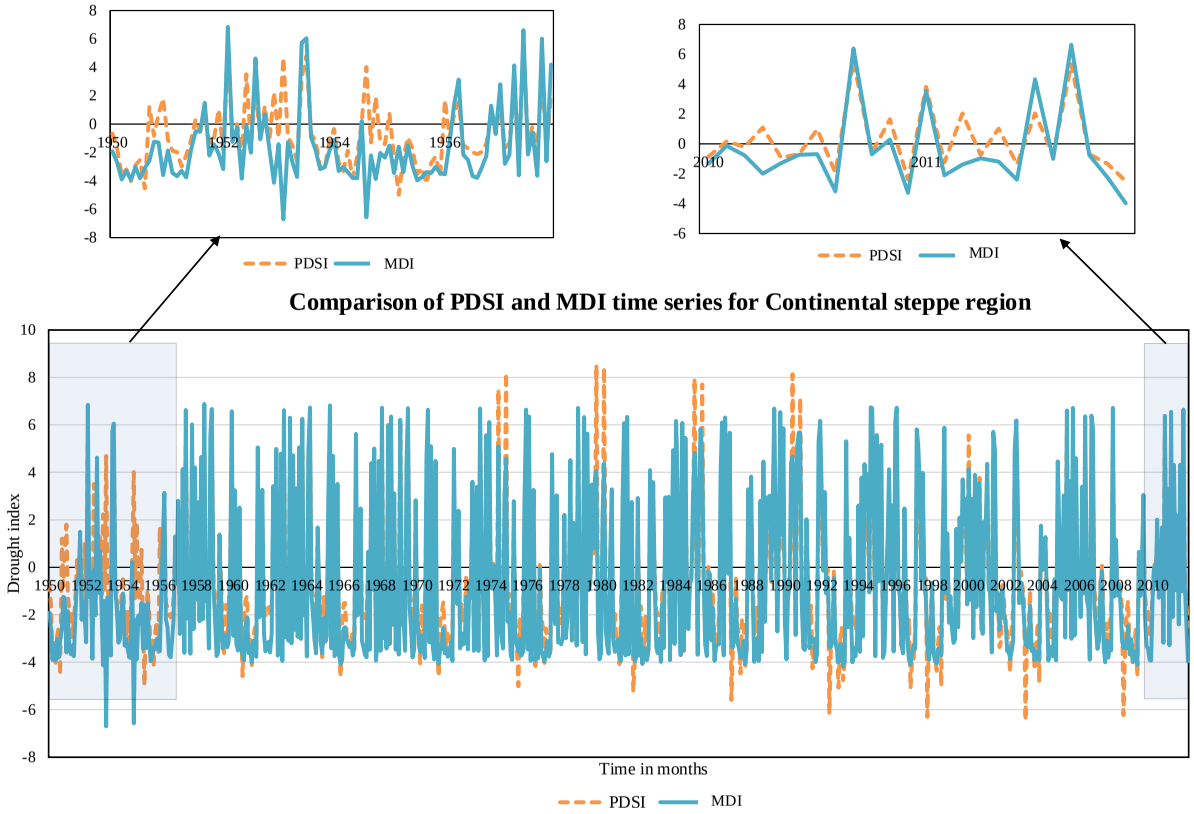
experiences the minimum evapotranspiration in Texas. In the case of soil moisture, the correlation coefficient values fall within the range 0.76–0.96, and thus the model explains 58%-92% of variability in the observed data. The N–S efficiency values for soil moisture lie within the range of 0.58–0.87.

5.1.2 Comparison of MDI With PDSI

In order to understand the performance of MDI in quantifying drought events, we compared it with other existing and established univariate and multivariate drought indices found in the literature. As a preliminary step for assessing the performance of MDI, it was compared against the Palmer drought severity index (PDSI). PDSI was chosen for two reasons: (1) It is widely used in the United States, and (2) it is formulated on the basis of the physical constituents of water balance. MDI also considers precipitation, runoff, evapotranspiration and soil moisture, which form the major components of the water balance. The PDSI values used for comparison were obtained from National Climate Data Center (NCDC).

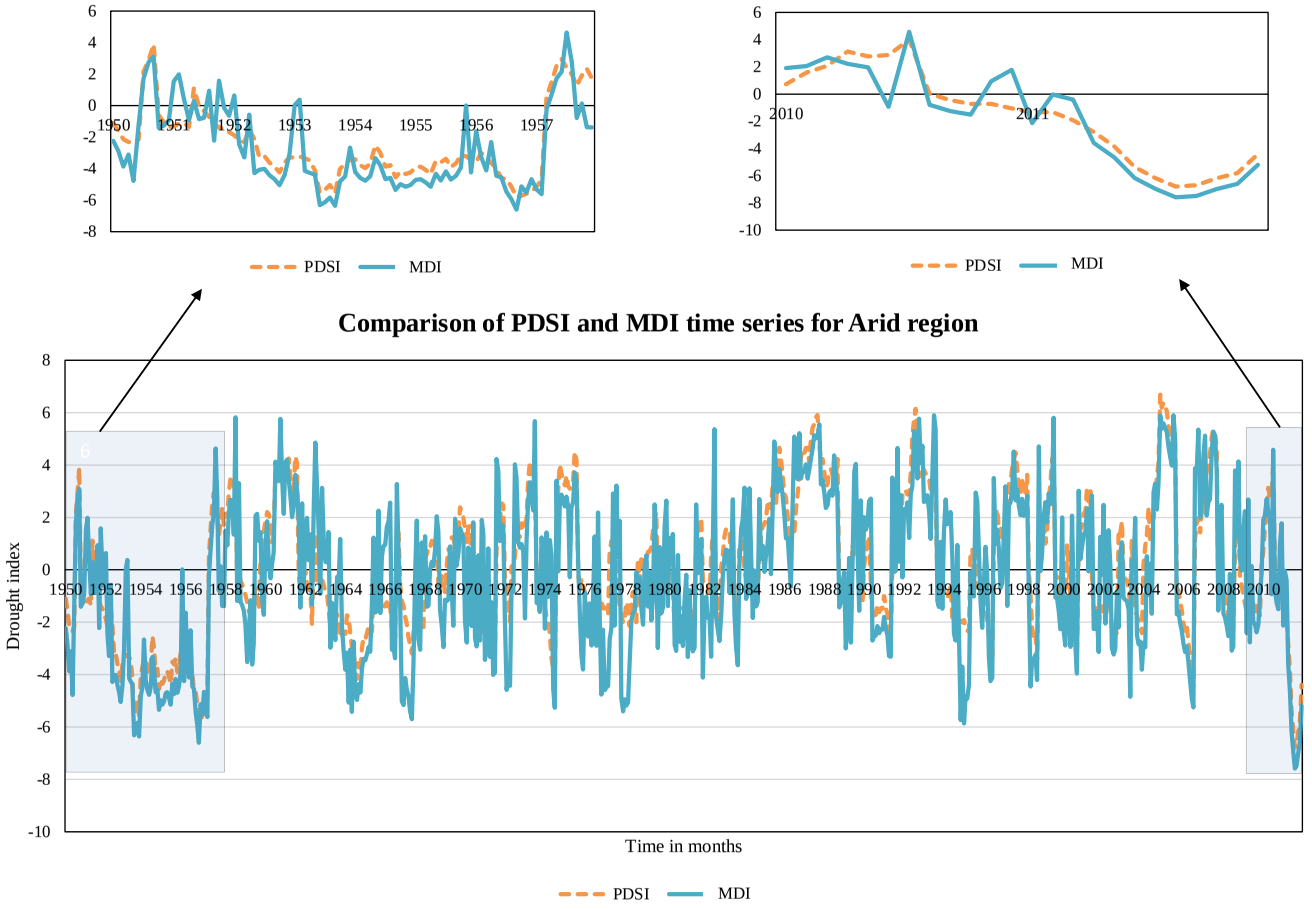
The drought classification for MDI values is given in Table 4.1. MDI followed the same drought classification as that of its constituent indices like SPI, SRI, etc. Comparison of the MDI time series against PDSI for different climate zones of Texas are shown in figures 5.4 a through e. The portions of time series which correspond to two major Texas drought periods: 1950–1957 and 2010–2011 have been enlarged for better visualization. However, although a perfect correlation is not expected between MDI and PDSI, it is natural that they both might follow a general behavioral pattern. Hence, it makes sense to analyze the monotonic relationship between MDI and PDSI. For this purpose, Spearman's rank correlation was used. Spearman's rank correlation (ρ) is a nonparametric statistical dependence measure that has many advantages over the Pearson correlation coefficient, since it does not depend upon

the distribution of data and is specifically designed to study monotonic relationship between variables. Table 5.4 shows Spearman's ρ between PDSI and MDI in all five climate zones of Texas. It can be seen that in all climate zones, a positive monotonic relationship is seen between PDSI and MDI, and the maximum value reaches 0.71. It can be seen from Table 5.4 that the correlation between PDSI and MDI is strongest for the continental steppe climate zone which is characterized by low precipitation and mild winters. Following that, a relatively better correlation between the two indices can be seen for the humid region which is characterized by relatively higher precipitation. The correlation is weak for arid region and transition zones like semi–arid and semi–humid zones. Arid and semi–arid regions, in particular, are characterized by higher rates of evapotranspiration and lower precipitation. It can be seen that MDI and PDSI quantify droughts differently, particularly for transition climate zones and regions with higher rates of evapotranspiration.



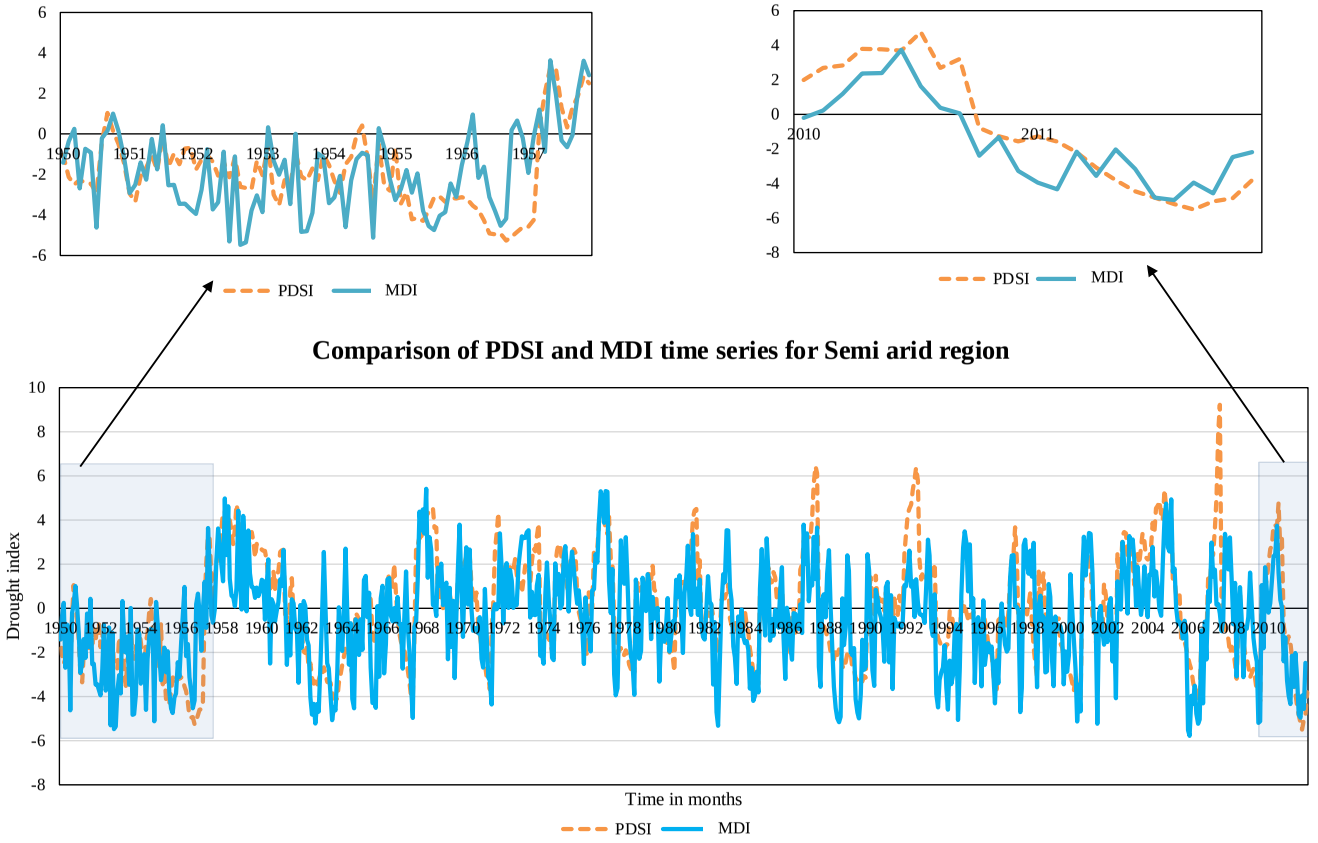
(a) Continental-Steppe Region

Figure 5.4: PDSI and MDI Time Series for Different Climate Regions in Texas During 1950-2012



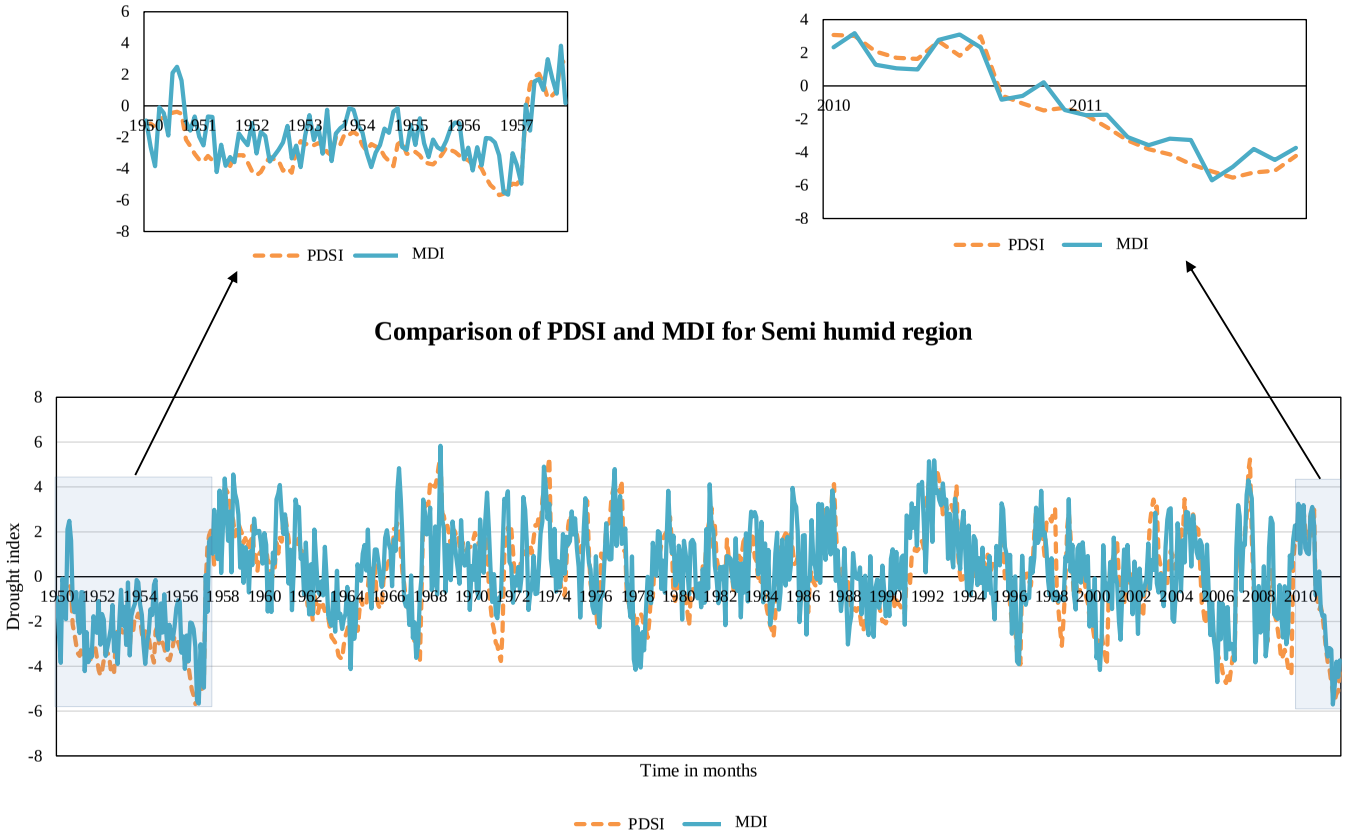
(b) Arid Region

Figure 5.4: Continued



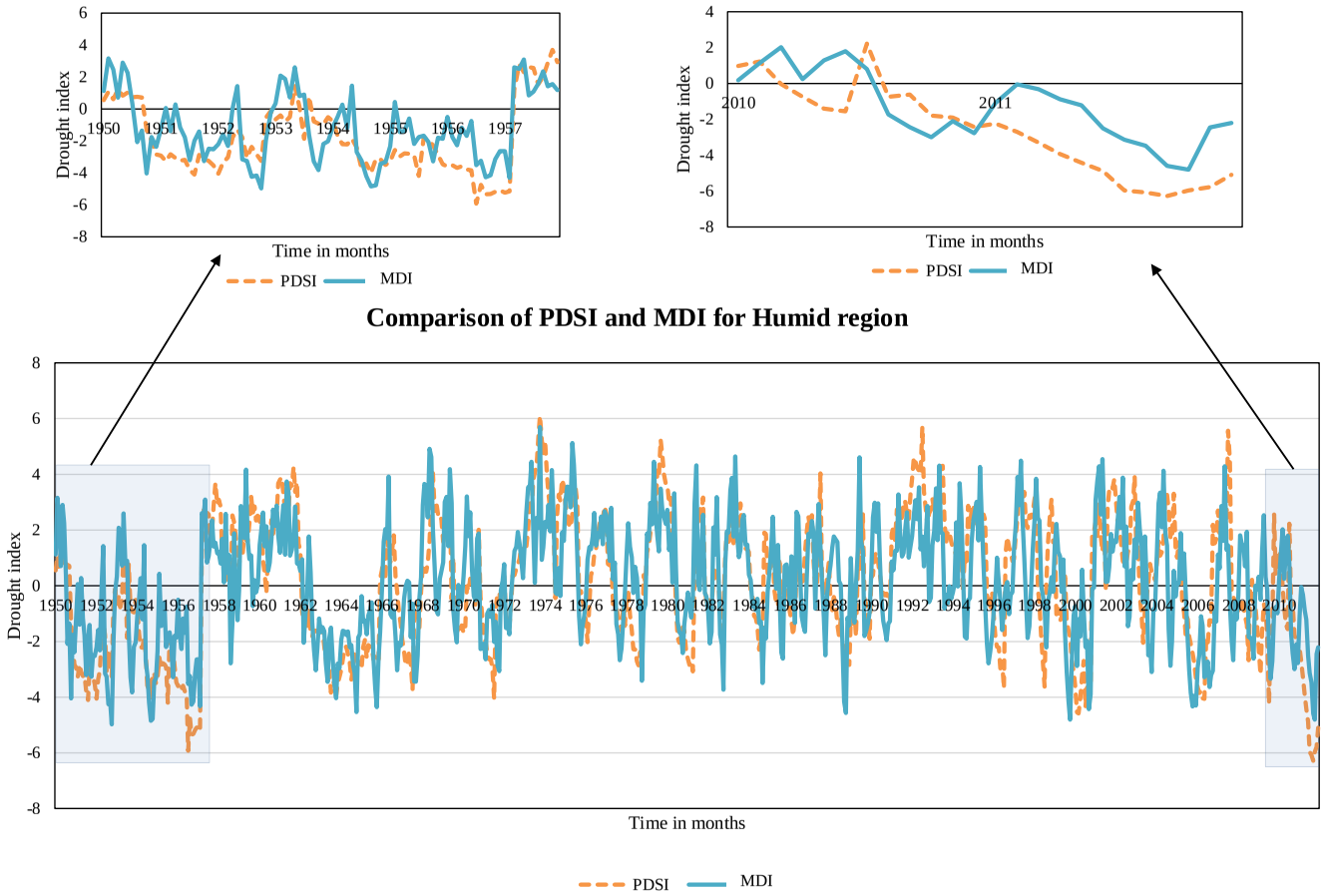
(c) Semi-Arid Region

Figure 5.4: Continued



(d) Semi-Humid Region

Figure 5.4: Continued



(e) Humid Region

Figure 5.4: Continued

Table 5.4: MDI Drought Classification

MDI value	Classification
2.0 or more	Extremely wet
1.5 to 1.99	Very wet
1.0 to 1.49	Moderately wet
-0.99 to 0.99	Near normal
-1.0 to -1.49	Moderately dry
-1.5 to -1.99	Severely dry
-2.0 or less	Extremely dry

Table 5.5: Spearmans Rank Correlation Between PDSI and MDI

Climate zone	Spearmans rank correlation
Continental	0.71
Arid	0.57
Semi-arid	0.51
Semi-humid	0.54
Humid	0.61

5.2 Drought Hazard Assessment

For the quantification of drought hazard using the joint occurrence probability of severity and duration, a weight and rating system as shown in Figure 4.3 was followed. The joint probability of severity and duration was obtained using appropriate copula functions. Previous drought studies conducted in Texas advocates the use of log–normal and exponential distributions for fitting severity and duration, respectively (Rajsekhar et al., 2012, 2014). A Maximum Likelihood method was used for parameter estimation. Even though the marginals belong to different families, copulas permit the modeling of their dependence structure. A two stage Maximum Likelihood method known as Inference Function for Margins (IFM; Joe, 1997) was

used for the parameter estimation of the following families of copulas considered in the study: Elliptical family (Gaussian), and Archimedean family (Clayton, Gumbel, and Frank). This is a two stage approach. Suppose we have a bivariate distribution with n observations for each margin, the first step of IFM is to use the maximum likelihood estimation (MLE) method to find the vector of marginal parameters β which maximizes the likelihood function:

$$logL(X_{ij};\alpha,\beta) = \sum_{i=1}^{n}\sum_{j=1}^{2} logf_i(X_{ij};\beta_j) \tag{5.3}$$

where $f(.)$ is the marginal probability density function. Section 2 briefly discusses parameter estimation using the MLE method. Then, the estimated $\hat{\beta}_{IFM} = (\hat{\beta_1^T}, \hat{\beta_2^T})^T$ from step 1, along with the sample data, was used to estimate the copula parameter α, which maximizes the likelihood function:

$$logL(X_{ij};\alpha,\beta) = \sum_{i=1}^{n} logc(F_1(X_{i1};\hat{\beta}_1), F_2(X_{i2};\hat{\beta}_2)) \tag{5.4}$$

where $F(.)$ is the marginal CDF. Again, iterative methods were applied to optimize the likelihood function to obtain the copula parameters $\hat{\alpha}_{IFM}$. The most suitable copula was chosen through a visual assessment of the scatter plots of observed and simulated data, and distance statistics like Anderson–Darling statistic. Table 5.6 lists the chosen copula for each planning region. The Gumbel copula was chosen for modeling the dependence structure in arid and humid climate regions. The Gaussian copula was found to be appropriate for semi-humid and semi-arid regions, whereas the Frank Copula was chosen for continental-steppe region. Figure 5.5 shows the visualization of the observed severity–duration data and the simulated severity–duration data from the best fit copula chosen for each planning region.

Figure 5.6 shows the joint probability plots of severity and duration for different planning regions. This would be used later for the calculation of Drought Hazard Index (DHI).

Table 5.6: Copula Chosen for Development of Joint Distribution in Each Planning Region

Region	**Region 1 Arid**	**Region 2 Continental**	**Region 3 Semi-Humid**	**Region 4 Semi-Arid**	**Region 5 Humid**
Best Copula	Gumbel	Frank	Gaussian	Gumbel	Gaussian

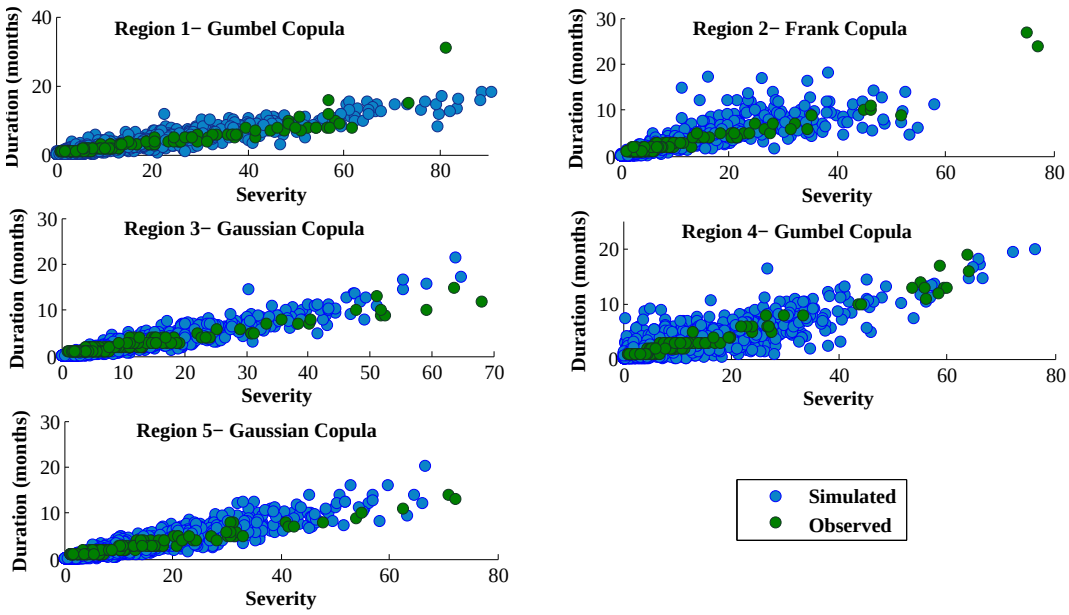


Figure 5.5: Scatter Plots Between Observed and Simulated Severity-Duration Values From Best Fit Copula For Each Planning Region

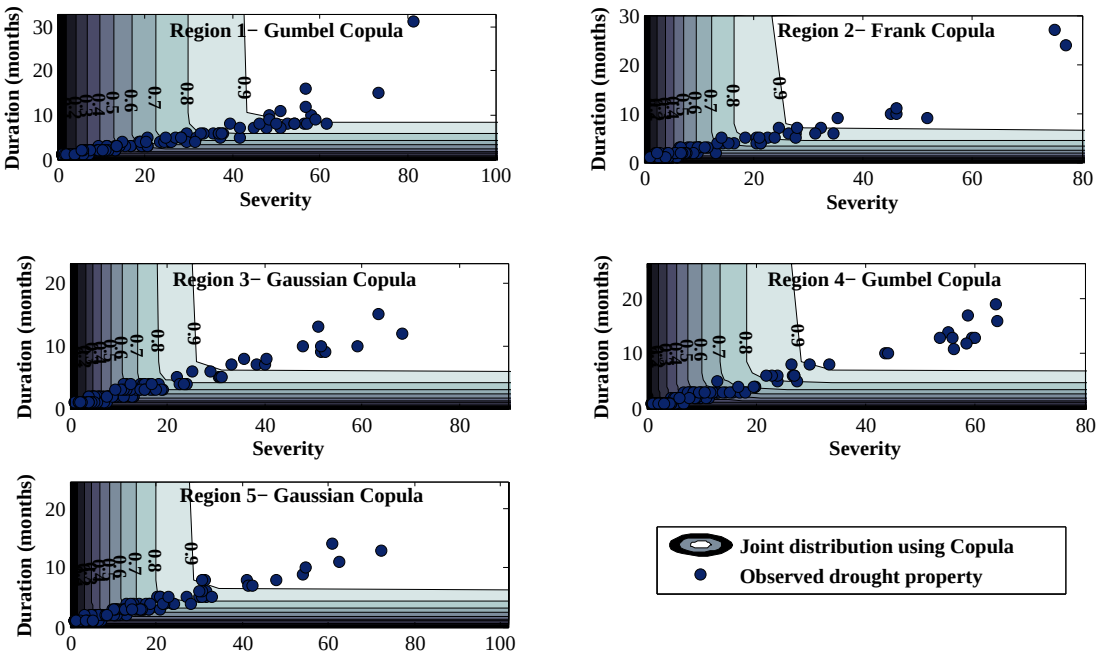


Figure 5.6: Severity and Duration Joint Probability Plots For Each Planning Region

The joint probability of severity and duration obtained from the plots shown in Figure 5.6 was transformed to Standardized Joint Probability (SJP), which lies in the normal probability space and follows the same classification system for standard drought indices like SPI. The weights were assigned for SJP values based on the system given in Table 4.2. Each of these weights was then assigned ratings from 1 through 4 on the basis of actual joint occurrence probability at the location of interest.

In order to understand the weights and ratings system to calculate the DHI, an example is given below. A random grid in Texas was chosen, and the joint probabilities of severities and durations were obtained using the copula probability plot. These probabilities were standardized to SJP. Based on the observed SJP values at this location, each SJP class (M, MO, S, and E) was further divided into

four sub–classes through the Jenks Natural Breaks method. These sub–classes were assigned ratings from 1 through 4. Table 5.7 gives the weights assigned, based on the SJP classes, and the ratings assigned through sub–classification, based on observed occurrence probabilities. Hence, if the location of interest is experiencing a series of drought events that has the following SJP values during the time period considered: -2.69, -1.54, -0.89, 0.14; then the weights and ratings assigned for these specific events based on Table 5.7 would be: 4 and 1, 3 and 2, 1 and 4, 1 and 2, respectively. Thus, the DHI based on these weights and ratings would be $4\times1 + 3\times2+1\times4+1\times2 = 16$. This same procedure was followed for all the grids for the time period of interest, and then the DHI values were rescaled to 0–1 range.

Table 5.7: Weight and Rating System for DHI Formulation at a Randomly Chosen Location

SJP Value	**Class**	**Weight**	**Jenks Sub-Classification of SJP**	**Rating**
-0.99 to 0.99	Near normal or mild (M)	1	<-0.724	4
			-0.723 to -0.215	3
			-0.214 to 0.821	2
			>0.822	1
-1.0 to -1.49	Moderate (MO)	2	<-1.84	4
			-1.839 to -1.474	3
			-1.473 to -1.065	2
			>-1.065	1
-1.5 to -1.99	Severe (S)	3	<-1.984	4
			-1.983 to -1.723	3
			-1.722 to -1.345	2
			>-1.344	1
-2.0 or less	Extreme (E)	4	<-10.171	4
			-11.08 to -6.763	3
			-6.762 to -2.915	2
			>-2.915	1

Figure 5.7 shows the drought hazard map for Texas during 1950–2012. Table 5.13 gives the percentage area under various hazard classes ranging from low to very

high. It can be seen that the percentage area under "Very High" and "High" drought hazard increases with time. "Very High" and "High" hazard zones are limited to arid, continental, and parts of semi-arid and semi-humid. A combination of several factors, like high evapotranspiration rates, low precipitation, and obvious lack of perennial rivers in these regions, contribute towards a higher drought hazard. These conditions are expected to worsen in future due to rising trends in temperature and decreasing precipitation in arid and semi–arid climate regions.

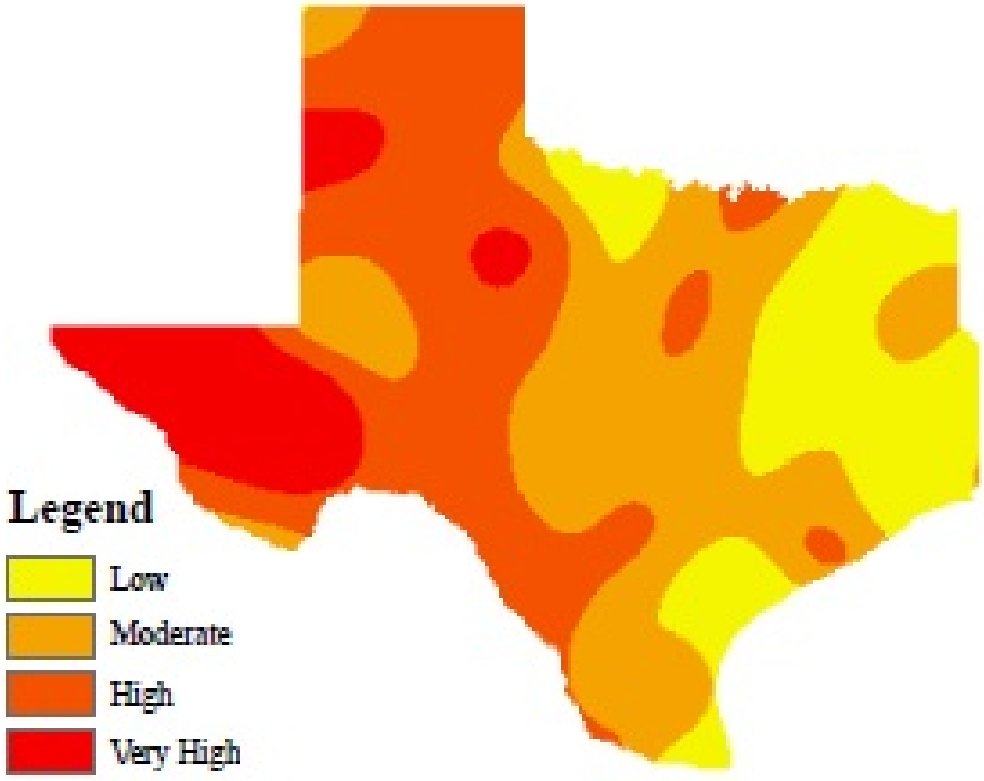


Figure 5.7: Drought Hazard Map For Texas During 1950-2012

Table 5.8: Percentage Area Under Different Classes of Drought Hazard During Various Time Periods

Region	**Classification Level**	**Percentage Area**
Arid	Low Hazard	1
	Moderate Hazard	17.12
	High Hazard	62.84
	Very High Hazard	19.04
Continental-Steppe	Low Hazard	8.64
	Moderate Hazard	21.27
	High Hazard	59.58
	Very High Hazard	10.51
Semi-Humid	Low Hazard	17.74
	Moderate Hazard	69.07
	High Hazard	12.77
	Very High Hazard	0.45
Humid	Low Hazard	68.2
	Moderate Hazard	30.37
	High Hazard	1.43
	Very High Hazard	0
Semi-Arid	Low Hazard	29.09
	Moderate Hazard	46.79
	High Hazard	24.12
	Very High Hazard	0

However, the actual risk posed by these drought events will depend on the socio–economic status of the area under consideration too. That will determine how well the region can cope when a drought event occurs. Thus, a joint hazard and vulnerability assessment is required to get a true picture of how the region under consideration will be affected in the event of drought. The following subsection discusses the results of vulnerability assessment.

5.3 Drought Vulnerability Assessment

In order to account for the socio–economic effects of drought in a developed economy like the study area, a number of factors that reflect the trends in water demand were considered. Values of population densities, municipal water demand, and non–municipal water demands from the primary economic sectors in Texas like agriculture, mining, manufacturing, and steam electric power generation were the vulnerability factors considered in this study. The county–wise values of these factors were obtained from TWDB and each of the above mentioned factors was divided into four classes based on natural break method that uses the average of each range to distribute the data more evenly across the ranges. The values in each class were then rescaled to fall within the ranges mentioned in Table 4.4. Classes with higher values are given a higher rating and vice versa. The composite DVI was then calculated by taking a simple average of the rescaled values of all the factors. A spatial interpolation was done using kriging to obtain the DVI values for grids having $1/8^{th}$ degree resolution. To illustrate the calculation of DVI, consider Anderson county. This location has a PD of 61210. The PD values for various counties in Texas ranges between 82 and 4883007. Based on the Jenks method, the PD value of 62017 belonged to the first class which was assigned a value of 0.25. This is repeated for other factors and the overall average will be taken to get DVI at that location.

Figure 5.8 shows the composite drought vulnerability map for Texas during 1950-2012. Table 5.9 gives the percentage area under various vulnerability classes ranging from low to very high. High vulnerability areas are concentrated in arid, continental, and semi-arid regions. A few hot "spots" were found in semi-humid and humid regions. This is because these regions have high projected population densities and municipal and irrigation water demands. The areas showing high vulnerability indices will typically exhibit higher levels of agricultural damage due to droughts. Hence, identification of highly vulnerable areas is necessary to adopt better crop management tactics and other intensified localized planning measures.

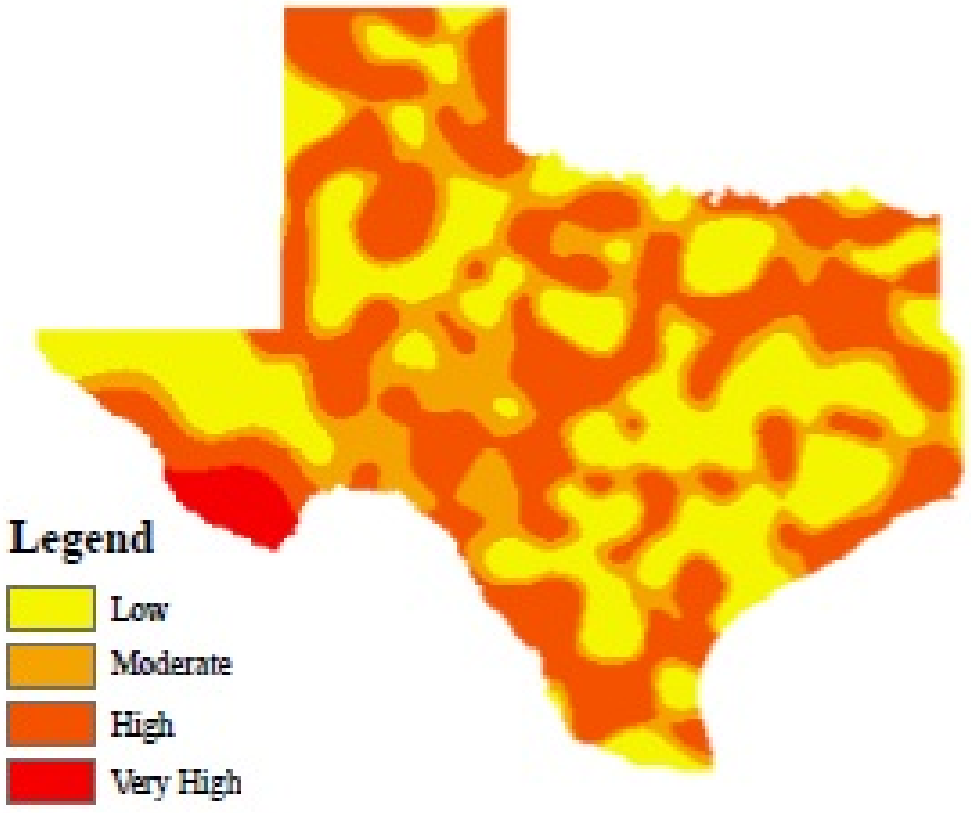


Figure 5.8: Composite Drought Vulnerability Map For Texas During 1950-2012

Table 5.9: Percentage Area Under Different Classes of Drought Vulnerability During Various Time Periods

Region	Classification Level	Percentage Area
Arid	Low Vulnerability	34.56
	Moderate Vulnerability	28.01
	High Vulnerability	32.46
	Very High Vulnerability	4.97
Continental-Steppe	Low Vulnerability	24.78
	Moderate Vulnerability	41.12
	High Vulnerability	34.1
	Very High Vulnerability	0
Semi-Humid	Low Vulnerability	19.87
	Moderate Vulnerability	22.34
	High Vulnerability	57.79
	Very High Vulnerability	0
Humid	Low Vulnerability	63.52
	Moderate Vulnerability	9.37
	High Vulnerability	26.61
	Very High Vulnerability	0
Semi-Arid	Low Vulnerability	29.97
	Moderate Vulnerability	16.19
	High Vulnerability	53.84
	Very High Vulnerability	0

Risk associated with drought events was then calculated as the product of hazard and vulnerability.

5.4 Drought Risk Assessment

The product of DHI and DVI will give the Drought Risk Index (DRI) for Texas. DRI follows the same classification system as DHI and DVI. Figure 5.9 shows the drought risk map for Texas during 1950-2012. Table 5.10 shows the percentage area under different drought risk categories. It can be seen that the "Very High" risk

areas are concentrated within arid, continental-steppe, and semi-arid regions. It can be seen that the spatial distribution of drought hazard and risk are not the same. The reason is the vulnerability of the area that indicates how well it can cope after a drought event. For example, if a region has relatively lower population density, and subsequently lower water demand, it will be less vulnerable to a drought event, and hence even for a higher magnitude drought, the associated risk might be less than that compared to a highly populated region which has a higher vulnerability towards droughts. Thus, the socio–economic scenario of the study area also plays an important role in determining the final impact of the drought event on the inflicted area.

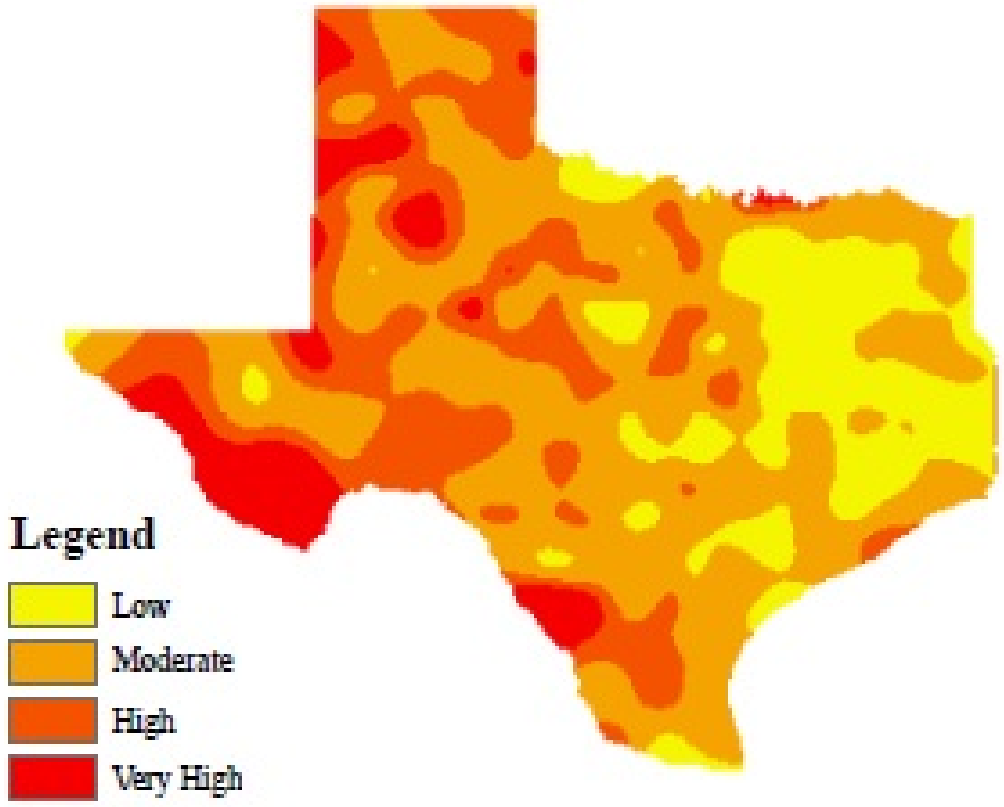


Figure 5.9: Drought Risk Map For Texas During 1950-2012

Table 5.10: Percentage Area Under Different Classes of Drought Risk During Various Time Periods

Region	**Classification Level**	**Percentage Area**
Arid	Low Risk	14.76
	Moderate Risk	18.13
	High Risk	58.74
	Very High Risk	8.37
Continental-Steppe	Low Risk	9.08
	Moderate Risk	40.62
	High Risk	43.85
	Very High Risk	6.45
Semi-Humid	Low Risk	19.98
	Moderate Risk	49.04
	High Risk	30.98
	Very High Risk	0
Humid	Low Risk	43.72
	Moderate Risk	39.87
	High Risk	16.41
	Very High Risk	0
Semi-Arid	Low Risk	29.10
	Moderate Risk	46.77
	High Risk	19.48
	Very High Risk	4.65

6. CONCLUSIONS

In this report, a novel multivariate, non–linear, and multi–scalar drought index named Multivariate Drought Index (MDI) is introduced. MDI uses a feature extraction technique known as Kernel Entropy Component Analysis (KECA), and finds the smallest set of features that maximally preserves the entropy estimate of the input data set, thus preserving all the information content that can be obtained from the input data set of precipitation, evapotranspiration, runoff, and soil moisture. Potential drought hazard regions were identified using a Drought Hazard Index (DHI) formulated on the basis of multivariate occurrence probabilities of drought properties. The vulnerability of the study region was accounted for by calculating a Drought Vulnerability Index (DVI) based on various socio–economic factors. A combined risk assessment of future droughts using both DHI and DVI was conducted to identify the areas which might fall under high drought risk. Overall, the results obtained from this study are expected to help achieve an effective drought mitigation strategy for the state of Texas. A better understanding of the evolution of drought events, integrated characterization of drought, the magnitude of risk it poses, and knowledge of its frequency of occurrence will help mankind to prevent, or at least better prepared for the many devastating effects droughts can potentially cause.

REFERENCES

Abdulla, F. A., Lettenmaier, D. P., Wood, E. F and Smith, J. A (1996). Application of a macroscale hydrologic model to estimate the water balance of the Arkansas-Red River Basin, J. Geophys. Res., 101(D3), 7449-7459.

Andreadis, K. M., and Lettenmaier, D. P. (2006). Trends in 20th century drought over the continental United States, Geophys. Res. Lett., 33(10), L10403, 1-4.

Beirlant, J., Dudewicz, E. J., Gyrfi, L., and Van der Meulen, E. C. (1997). Nonparametric entropy estimation: An overview, Int. J. Mathematical and Stat. Sci., 6, 17-40.

Benke, A. C., and Cushing, C. E. (2005). Rivers of North America, Burlington, Massachusetts: Elsevier Academic Press, 1-1144.

Bin, H., L, A., Wu, J., Zhao, L., and Liu, M. (2011). Drought hazard assessment and spatial characteristics analysis in China, J. Geographical Sci., 21(2), 235-249.

Borg, I., and Groenen, P. J. (2005). Modern multidimensional scaling: Theory and applications, Springer Series in Statistics, Springer, New York, 1-614.

Bowling, L. C., Lettenmaier, D. P., Njissen, B., Polcher, J., Koster, R. D. and Lohmann, D. (2003a). Simulation of high latitude hydrological processes in the Torne-Kalix basin: PILPS Phase 2(e) 3: Equivalent model representation and sensitivity experiments, Global and Planetary Change, 38, 55-71.

Bowling, L. C., Lettenmaier, D. P., Njissen, B., Polcher, J., Koster, R. D. and Lohmann, D. (2003b). Simulation of high-latitude hydrological processes in the Torne-Kalix basin: PILPS Phase 2(e) 1: Experiment description and summary intercomparisons, Global and Planetary Change, 38, 1-30.

Brooks, N., Neil Adger, W., and Mick Kelly, P. (2005). The determinants of vulnerability and adaptive capacity at the national level and the implications for

adaptation, Global Env. Change, 15(2), 151-163.

Bureau of Economic Geology (1996). River Basin Map of Texas, University of Texas, Austin, 1-2.

Cheng, J., and Tao, J. P. (2010). Fuzzy Comprehensive Evaluation of Drought Vulnerability Based on the Analytic Hierarchy Process: An Empirical Study from Xiaogan City in Hubei Province, Agriculture and Agricultural Sci. Procedia, 1, 126-135.

Cosby, B. J., Hornberger, G. M., Clapp, R. B. and Ginn, T. R. (1984). A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils, Water Resour. Res., 20, 682690.

Dionisio, A., Menezes, R., and Mendes, D. A. (2007). Entropy and uncertainty analysis in financial markets, Applications of Physics in Finance Analysis, 4-7 July, Portugal, 1-9.

Ebrahimi, N., Maasoumi, E., and Soofi, E. S. (1999). Ordering univariate distributions by entropy and variance, J. Econometrics, 90(2), 317-336.

Fontaine, M. M., and Steinemann, A. C. (2009). Assessing vulnerability to natural hazards: Impact-based method and application to drought in Washington State, Nat. Hazards Review, 10(1), 11-18.

Franchini, M., and Pacciani, M. (1991). "Comparative-analysis of several conceptual rainfall runoff models," J. Hydrol., 122(1-4), 161-219.

Guerrero-Salazar, P. and Yevjevich, V. (1975) Analysis of Drought Characteristics by the Theory of Runs, Hydrol. Paper No. 80, Colorado State University, Fort Collins, 1-53.

Hansen, J., Ruedy, R., Glascoe, J., and Sato, M. (1999). GISS analysis of surface temperature change, J. Geophys. Res.: Atmospheres (19842012), 104(D24), 30997-31022.

Hansen, M. C., DeFries, R. S. Townshend, J. R. G. and Sohlberg, R. (2000). Global land cover classication at 1 km spatial resolution using a classication tree approach, Int. J. Remote Sens., 21, 13311364.

Hao, Z., and AghaKouchak, A. (2013). A non-parametric multivariate multi-index drought monitoring framework, Sp. Issue in Adv. Drought. Monitoring., Am. Met. Soc., 15, 89-101.

Hao, Z., and AghaKouchak, A. (2013). Multivariate Standardized Drought Index: A parametric multi-index model, Adv. Wat. Res., 57, 12-18.

Huang, J., van den Dool, H. M., and Georgarakos, K. P. (1996). Analysis of model-calculated soil moisture over the United States (1931-1993) and applications to long-range temperature forecasts, J. Clim., 9(6), 1350-1362.

Ignacio Rodriguez-Iturbe (1969). Applications of Theory of Runs to hydrology, Water Resour. Res. Letters, Vol. 5(6), 1422-1426.

Jenssen, R. (2010). "Kernel Entropy Component Analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(5), 847-860.

Jenssen, R., Erdogmus, D., Hild, K. E., Principe, J. C. and Eltoft, T. (2005). "Optimizing the Cauchy-Schwarz PDF Divergence for Information Theoretic, Non-Parametric Clustering," Proc. Int'l. Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, 34-45, St. Augustine, USA, 1-21.

Jenssen, R., Erdogmus, D., Hild, K. E., Principe, J. C. and Eltoft, T. (2006). "Some Equivalences between Kernel Methods and Information Theoretic Methods," J. VLSI Sig. Processing, 45, 49-65.

Joe, H. (1997). Multivariate Models and Dependence Concepts, Chapman and Hall: New York, 1-395.

Kao, S. C. and Govindaraju, R. S. (2010). A copula-based joint deficit index for droughts, J. Hydrol., 380(1-2), 121-134.

Keyantash, J. A., and Dracup, J. A. (2004). An aggregate drought index: Assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage, Wat. Resour. Res. 40(9), 1-13.

Kim, H., Park, J., Yoo, J., and Kim, T. W. (2013). Assessment of drought hazard, vulnerability, and risk: A case study for administrative districts in South Korea, J. HydroEnv. Res., In Press, Corrected Proof, Available Online 6 August 2013.

Knutson, C., Hayes, M., and Phillips, T. (1998). How to Reduce Drought Risk, Preparedness and Mitigation Working Group of the Western Drought Coordination Council, Lincoln, Nebraska, 1-43.

Kogan, F. N. (1995). Application of vegetation index and brightness temperature for drought detection, Adv. Space Res., 15(11), 91-100.

Lathi, B. P. (1968). An introduction to Random Signals and Information Theory, International Textbook Company, Scanton, Pennsylvania, 1-488.

Liang, X., Lettenmaier, D. P., Wood, E. F. and Burges, S. J. (1994). A Simple hydrologically Based Model of Land Surface Water and Energy Fluxes for GSMs, J. Geophys. Res., 99(D7), 14415-14428.

Lohmann, D., Nolte-Holube, R. and Raschke, E. (1996). A large-scale horizontal routing model to be coupled to land surface parametrization schemes, Tellus, 48(A), 708-721.

Lohmann, D., Raschke, E., Nijssen, B. and Lettenmaier, D. P. (1998). Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model, Hydrol. Sci. J., 43(1), 131-141.

Maurer, E. P., Nijssen, B. and Lettenmaier, D. P. (2000). Use of reanalysis land surface water budget variables in hydrologic studies, GEWEX News, 10 (4), 68.

Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier, D. P. and Nijssen, B. (2002). A Long-Term Hydrologically-Based Data Set of Land Surface Fluxes and

States for the Conterminous United States, J. Clim., 15(22), 3237-3251.

McKee, T. B., Doesken, N. J., Kleist, J. (1993). The relationship of drought frequency and duration to time scales, Eighth conf. App. Clim., 17-22 January, Anaheim, CA, 1-6.

Metzger, M. J., Rounsevell, M. D. A., Acosta-Michlik, L., Leemans, R., and Schrter, D. (2006). The vulnerability of ecosystem services to land use change, Agriculture, Ecosys. Env., 114(1), 69-85.

Millan, J. and Yevjevich, V. (1971). Probabilities of observed droughts, Hydrology Paper No. 50, Colorado State University, Fort Collins, Colorado, 1-23.

Miller, D. A., and White, R. A. (1998). A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling, Earth Interactions, 2.

Mishra, A. K. and Singh, V. P. (2010). A review of drought concepts, J. Hydrol., 391, 202-216.

Mitchell, K. (1999). The GCIP Land Data Assimilation (LDAS) ProjectNow underway, GEWEX News, 9 (4), 36.

Modarres, R. (2007). Streamflow drought time series forecasting, Stoch. Environ. Res. Risk Assess., 22, 223-233.

Myneni, R. B., Nemani, R. R. and Running, S. W. (1997). Estimation of global leaf area index and absorbed PAR using radiative transfer models, IEEE Trans. Geosci. Remote Sens., 35, 13801393.

Nijssen, B. N., Lettenmaier, D. P., Liang, X., Wetzel, S. W. and Wood, E. F. (1997). Streamflow simulation for continental-scale river basins, Water Resour. Res., 33(4), 711-724.

Nijssen, B. N., O'Donnell, G. M., Lettenmaier, D. P. and Wood, E. F. (2001). Predicting the discharge of global rivers, J. Clim., 14, 3307-3323.

Palmer, W. C. (1965). Meteorological drought, US Department of Commerce, Weather Bureau, Washington, D.C., 1-58.

Palmer, W. C. (1968) "Keeping track of crop moisture conditions, nationwide: The new crop moisture index," Weatherwise, 21(4), 156-161.

Rajsekhar, D., Mishra, A. K. and Singh, V. P. (2012). Regionalization of drought characteristics using an entropy approach, J. Hydrol. Eng., 18(7), 870-887.

Rajsekhar, D., Singh, V. P., and Mishra, A. K. (2014). Hydrologic Drought Atlas for Texas, J. Hydrol. Eng., In Press, Corrected Proof, Available Online 27 August 2014.

Rajsekhar, D., Singh, V. P., and Mishra, A. K. (2014). Multivariate Drought Index: An Information Theory based approach for Integrated drought assessment, Special Issue on Droughts, J. Hydrol, In Press, Corrected Proof, Available Online 26 November, 2014.

Rawls, W. J., Ahuja, L. R., Brakensiek, D. L. and Shirmohammadi, A. (1993). Inltration and soil water movement, Handbook of Hydrology, D. Maidment, Ed., McGraw-Hill, 5.15.51.

Reynolds, C. A., Jackson, T. J. and Rawls, W. J. (2000). Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions, Water Resour. Res., 36, 36533662.

Ribot, J. C., Najam, A., and Watson, G. (1996). Climate variation, vulnerability and sustainable development in the semi-arid tropics, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 15-34.

Salathe, E. P. (2003). Comparison of various precipitation downscaling methods for the simulation of stream flow in a rain shadow river basin, Int. J. Clim., 23, 887901.

Saldarriaga, J. and Yevjevich, V. (1970). Application of run-lengths to hydrologic series, Hydrology Paper No. 40, Colorado State University, Fort Collins, Colorado, 1-56.

Saul, L. K., Weinberger, K. Q., Ham, J. H., Sha, F., and Lee, D. D. (2006). Spectral methods for dimensionality reduction, Semisupervised Learning, 293-308.

Schlkopf, B. (2000). Statistical learning and kernel methods, Springer, Vienna, 3-24.

Schlkopf, B., and Smola, A. J. (2002). Learning with kernels: Support vector machines, regularization, optimization, and beyond, MIT press, 1-625.

Scholkopf, B., Smola, A., and Mller, K. R. (1999). Kernel principal component analysis, Adv. Kernel Methods- Support Vector Learning, Artificial Neural Network-sICANN, Springer, Berlin, 583-588.

Sen, Z. (1976). Wet and dry periods of annual flow series, J. Hydraul. Div., American Society of Civil Engineers, Proc. Paper 12457, 102(HY10): 1503-1514.

Sen, Z. (1977). Run-sums of annual flow series, J. Hydrol., 35, 311-324.

Shahid, S., and Behrawan, H. (2008). Drought risk assessment in the western part of Bangladesh, Nat. Hazards, 46(3), 391-413.

Shannon, C. E. (1948). A Mathematical Theory of Communication, Bell System Technical Journal, 27, 379-423.

Sheffield, J., and Wood, E. F. (2008). Global Trends and Variability in Soil Moisture and Drought Characteristics, 19502000, from Observation-Driven Simulations of the Terrestrial Hydrologic Cycle, J. Clim., 21, 432458.

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

Sheffield, J., Goteti, G., Wen, F. and Wood, E. F. (2004). A simulated soil moisture based drought analysis for the United States, J. Geophys. Res., 109D24, 1-19.

Shepard, D. S. (1984). Computer mapping: the SYMAP interpolation algorithm, Spatial statistics and models, Gaile, G. L. and Wilmott, C. J., eds. Dordrecht, Netherlands, 133-145.

Shiau, J. (2006). Fitting Drought Duration and Severity with Two-Dimensional Copulas, Water Resour. Mgmt., 20(5): 795-815.

Shukla, S., and Wood, A. W. (2008). Use of a standardized runoff index for characterizing hydrologic drought, Geophys. Res. Letters, 35(2), L02405, 1-7.

Singh, V. P. (2013). Entropy Theory and its Applications in Water and Environmental Engineering, Wiley and Sons, 559-580.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., and Stephens, S. (2002). The drought monitor, Bul. Am. Met. Soc., 83(8), 1181-1190.

Vicente-Serrano, S. M., Beguera, S., and Lpez-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index, J. Clim., 23(7), 1696-1718.

Wilhelmi, O. V., and Wilhite, D. A. (2002). Assessing vulnerability to agricultural drought: A Nebraska case study, Nat. Hazards , 25 (1), 3758.

Williams, C. K. I. (2002). "On a Connection between Kernel PCA and Metric Multidimensional Scaling," Machine Learning, 46, 11-19.

Wu, J., He, B., L, A., Zhou, L., Liu, M., and Zhao, L. (2011). Quantitative assessment and spatial characteristics analysis of agricultural drought vulnerability in China, Nat. Hazards, 56(3), 785-801.

Yevjevich V., Siddiqui, M. M and Downer, R. N. (1967). Application of Runs to hydrologic droughts, Proc. Int. Hydrol. Sym., Fort Collins, 1(63), 496-505.

Yuan, X. C., Wang, Q., Wang, K., Wang, B., Jin, J. L., and Wei, Y. M. (2013). Chinas regional vulnerability to drought and its mitigation strategies under climate change: data envelopment analysis and analytic hierarchy process integrated approach, Mitigation and Adaptation Strategies for Global Change, 1-19.

Zelenhastic, E. and Salvai, A. (1987). A method of stream flow drought analysis, Water Resour. Res., 23(1), 156-168.

Zhao, R. J. (1980). "The Xinanjiang model," Hydrol. Forecasting Proc., Oxford Symposium, 129, 351-356.

Zhang, Q., Sun, P., Li, J., Xiao, M., and Singh, V. P. (2014). Assessment of drought vulnerability of the Tarim River basin, Xinjiang, China, Theoretical and Applied Clim., 1-11.