CLIMATE VARIABILITY AND EXTREME EVENTS ANALYSIS IN THE ZAMBEZI RIVER BASIN USING STANDARDIZED PRECIPITATION EVAPORATION INDEX AND L-MOMENTS

INCEPTION REPORT

DECEMBER 2017

PREPARED BY

PROFESSOR P.K. KENABATHO & PROFESSOR B.P. PARIDA

UNIVERSITY OF BOTSWANA

GABORONE, BOTSWANA

An Inception report submitted to the Joint Research Centre, European Commission, Ispra, Italy.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 BACKGROUND AND SIGNIFICANCE

The Zambezi River is a very important water resources with its catchment area covering most parts of Southern Africa. It is a habitat to a wide range of plant and animal species. Humanity in this region just like other animals depend on water from the Zambezi River and its tributaries. Its dependence ranges from provision of potable water, agriculture, power, manufacturing, mining, tourism and many other sectors (World Bank 2010). With such benefits from the river, the human population is proved to be rapidly increasing. For example, an annual increase of 3.9% in Africa (the highest in the world) has been recorded with most of the increase in the Southern part of the continent, which is mostly covered by this basin (World Bank, 2010). Despite the population boom, industrialization and urbanization, not all Africans have access to clean water and sanitation. Water availability varies from country to country as some parts of Southern Africa receive very low mean annual rainfall (Namib and Kalahari deserts) and low river flows with others receiving very high precipitation (areas in the sub tropics) hence high river flows. With this in mind, more water is needed to meet the increasing demands in clean water, sanitation, irrigation, power (Hydroelectric) and many for factors increasing the demand (O Serdeczny et al.,2016).

Climate variability (a variation in world climatic patterns due to effects of human activities) is another phenomena impacting on water availability in the Zambezi River Basin. The climate of an area is closely tied to its location relative to the Intertropical Convergence Zone (ITCZ) often called the 'climate equator'. This is an area around the geographic equator where north and south trade winds converge, rise and circulate back. As the winds converge, moist air rises. It then cools causing water vapor to convert to rainfall. Areas closest to the ITCZ receive the highest, frequent and reliable rainfall with a reduction as you go further south and north of the ITCZ. Because the earth is tilted on its axis relative to its orbit around the sun, the amount of heat received from the sun is not equal (in the north and southern hemisphere) and that gives rise to seasonal climatic changes. In January most of the precipitation is in the southern part of the

equator and shifts to the north as the year progresses (around June). Although these climatic variables show a very close link between the geographical location of an area relative to the ITCZ position, many studies predict a general change in precipitation patterns, temperature and dryness in Southern Africa, most of which is within the Zambezi River Basin.

Studies predict an increase in the mean annual surface air temperature of about 4ºC due to global warming and a reduction in the mean annual rainfall (by up to about 20%) by the end of the 21st Century (Engelbrecht et al (2011)). Similarly, there is evidence of marginal increase in precipitation in some parts of the basin and across the entire continent (Schlosser and Strzepek, 2015; Engelbrecht et al, 2011), with more extreme events (droughts, floods, heat waves and veld fires) occurring very often (Kenabatho et al., 2012). The projected aridity changes indicated by the Aridity Index also show a strong deterioration towards more arid conditions due to low rainfall in the Southern part of Africa (Serdeczny et al., 2016). These factors are going to impact negatively on the already existing water resources and low and unreliable rainfall together with prolonged droughts will result in low surface water flows and low groundwater recharge rate (Kenabatho et al., 2012). Increasing temperatures cause high evaporation rates. In combination with loss of vegetation cover due to overstocking, overgrazing and deforestation there is an accelerated rate of loss of soil moisture that will cause a low agricultural produce impacting negatively on food security in Southern Africa.

Researchers also reveal that in the Zambezi River Basin, a slight climatic change case results in a 32% fall in energy production (Belfuss R, 2012) affecting other sectors such as mining and manufacturing which are much reliant on power. Taking all the issues above into consideration, there is a rapid rise in water demand due to population increase and urbanization which automatically expands the human dependence on the Zambezi River Basin.

Despite this increasing dependence, water availability is decreasing due to climate change, which implies the current abundance of water on most parts of the Zambezi River Basin is not likely to last (Beck and Bernauer, 2011).

Although studies may have assessed the implications of water demand and climate change collectively (Beck and Bernauer T, 2011), and predict dryness across the region (Kenabatho et al., 2012), they do not reveal the dryness distribution across basin's riparian countries. This study

therefore, focuses on the dryness characteristics of the basin as a result of increased water demand, increasing temperatures and change in precipitation patterns due to El Nino.

It is therefore in this context that it is proposed to make a holistic study on the changes in the dryness condition across the Zambezi basin so as to assess the trends for the futuristic water resources management for this basin.

1.2 PROBLEM STATEMENT

The rapidly increasing population and economic growth in Southern Africa has a very large impact on water demand. The impacts are further increased by the climatic variations due to El Nino. This is evident as extreme events (such as droughts and heat waves) have been common in the past three decades with frequent water shortages (portable water) as most water resources have been below normal to almost empty (Dams and rivers) due to unreliable and low precipitation. Scientifically, if the rate of loss of soil moisture exceeds the amount of precipitation, dryness is said to occur, in that case leading to droughts. With most of Southern African population heavily dependent on the Zambezi for critical economic activities such as agriculture, Hydro power, fishing, it is necessary to study the dryness and wetness conditions based on historical data and predict the future so as to inform the relevant policies and decision makers on adaptive measures.

1.3 DESCRIPTION OF STUDY AREA

The area under study shall be the Zambezi River Basin. The river basin drains about 1.4 million km^2 across Angola, Botswana, Malawi, Mozambique, Namibia, Tanzania, Zambia and Zimbabwe. The river starts from a small spring in the Northwestern Zambia and stretches for approximately 3000 km across Zimbabwe, Malawi, through Mozambique into the Indian Ocean, sustaining a population of 40 million people (projected to be 51 million by 2025) (SADC 2013). The river itself rises from north of Zambia through Angola and other riparian countries to feed the Indian Ocean through Mozambique. It hosts numerous urban areas including most of Zambian cities and Harare (Zimbabwean Capital City). The natural environment is characterized by lakes (Nyasa, Malawi), gorges, water falls (including the famous Victoria Falls in Zambia)

and a very diverse flora and fauna species. The ecology and natural environment has a socio-economic influence on humanity since it directly hosts (facilitates) human activities such as mining, fishing, agriculture, forestry, manufacturing and tourism all relying on the hydropower electricity utilizing river water (WMO, 2009). The Zambezi River Basin map showing all riparian countries and catchment area (Highlighted) is shown in figure 1. Figure 2 further details the sub basins of the total catchment.

Figure 1-1*.Zambezi River Basin Map (Source:ZAMCOM)*

The river is fed by a number of tributaries who receive water from smaller streams in the sub basins. Such rivers in sub basins are detailed in figure 2.

Figure 1-2. *Sub basins and tributaries of the Zambezi River (Source: www.researchgate.com)*

The hydrology of the basin is dependent on a number of variables. The variation in precipitation (table 1.) in summer is due to Inter Tropical Convergence zone between the North east Monsoon and the South east trade winds for the middle and lower Zambezi and the Congo air boundary between the South West Monsoon and the South east Trades for the upper Zambezi (Moore at al., 2007). The further then South the boundaries move, the more the precipitation as seen in table 1. Drier conditions are experienced in winter as the ITCZ shifts to the north, and so most of the flow contribution is from the North (Moore at al., 2007). The mean monthly changes in flow are depicted in table 2.

Table 1-1: Precipitation data for the Zambezi River Basin (Source: WMO, 2009)

Sub basin	Mean annual precipitation (mm)
Kapombo	1211
Upper Zambezi	1225
Lungue Bungo	1103
Luanginga	958
Barotse	810
Cuando/Chobe	797
Kafue	1042
Kariba	701
Luangwa	1021
Mupata	813
Shire River and Lake Malawi/Niassa/Nyasa	1125
Tete	887
Zambezi Delta	1060
Zambezi River Basin, mean	956

Table 1-2: Zambezi River mean monthly flows in m^3s^{-1} (Source: Moore et al., 2007)

Station Period (Years)	Chavuma 1959/60-2001/02	Lukulu 1950/51-2001/02	Katima Mulilo 1967/68-2001/02	Victoria Falls 1951/52-2001/02
Oct	68	271	306	293
Nov	94	310	320	297
Dec	228	468	430	438
Jan	655	803	678	686
Feb	1411	1294	1211	1184
Mar	2031	1645	2374	2175
Apr	1770	1523	3129	3007
May	684	944	2427	2613
Jun	310	575	1326	1621
Jul	188	434	691	845
Aug	124	361	467	519
Sep	83	306	364	376
Mean annual	637	745	1144	1171

1.4 RESEARCH OBJECTIVE

How is the Zambezi River Basin changing in terms of climate variability and extreme events?

1.4.1 SPECIFIC RESEARCH OBJECTIVES

- To analyze dryness/wetness conditions of the Zambezi River Basin based on historical data
- To produce a dryness/wetness severity map for the Zambezi River Basin based on the analysis
    - To project the likely dryness/wetness conditions of the Zambezi River Basin in the next five years.

1.5 RESEARCH QUESTIONS

Table 1-3: Research objectives and questions

Objectives	**Research Questions**
To analyze dryness/wetness conditions of the ZRB based on historical data.	Is the average soil moisture increasing or decreasing?
	What is the rate of increase of the dryness or wetness?
	What is the seasonal outlook of the basin using available data?
To produce a dryness/wetness severity map for the ZRB based on the analysis	Which areas are wetting and which ones are drying?
	What is the rate of drying/wetting per location?
	How severe is the drying/wetting per location?
To project the likely dryness /wetness of the ZRB in the next two decades	What is the likelihood of the basin drying/wetting in the next two decades?

1.6 SCOPE OF RESEARCH

The study shall cover the Zambezi river basin's riparian countries. However, this shall depend on the availability of data as the scope might be scaled down if precipitation data of some riparian

countries is not available. Similarly, if all the data is available, the scope can be increased to include Lesotho, Swaziland and South Africa even though they do not form part of the basin.

The tool to be used for data analysis is the Standardized Precipitation Index/Standardized Precipitation Evaporation (SPEI) as described by McKee et al., 1992 and Vicente-Serrano et al., 2010. In addition, related analysis such as trends, and change point/intervention analyses will be undertaken.

1.7 BENEFITS AND BENEFICIARIES

The study shall reveal the dryness/wetness conditions of the study area in response to climate change. This information is beneficial in the sense that it will inform future decision making for policy makers in various organizations and enterprises on adaptive measures to take within the context of climate change. These are shown in Table 1-4.

Table 1-4: Benefits and beneficiaries of this study

Beneficiary sector	Benefits
Agriculture	Development of climate adaptive technologies in plant and animal production.
	Prediction of likelihood climate borne plant and animal diseases.
	Forecasting of possibility of low agricultural produce so as to take necessary measures.
Political	Development of policies regarding water supply, agriculture and other sectors affected by dryness/wetness of the land.
Manufacturing	Development of better manufacturing techniques in response to raw material quality (for example, the quality of timber may be influenced by the availability of soil moisture where it was grown).
Education	Dissemination of information so as to inform the society on dryness/wetness adaptation measures in different sectors such as agriculture (for example, new planting technologies).
Forestry	Understanding of how forestry resources may change in future.

1.8 COLLECTION AND COLLATION OF EXISTING DATA

Collation of data shall be based on available data which will be collected during the project. It is expected that the NEPAD Southern African Network of Centres of Excellence (SANWATCE) hub at Stellenbosch University will play a key role in facilitating collection of data through relevant authorities within the Zambezi River Basin.

2 METHODOLOGICAL APPROACH

2.1 CLIMATE VARIABILITY

2.1.1 HOMOGENEITY TEST

Data used in hydrological studies is required to be stationary, consistent, and homogeneous while used for frequency analyses or to simulate a hydrological process (Dahmen and Hall, 1990). Climate data collected at a given weather station during a period of several years may be non-homogeneous. A homogeneity test will be carried out on the data series using one base station against the averages of the rest of the stations in the study area through the method of Cumulative Residuals.

2.1.1.1 CUMULATIVE RESIDUALS

The average climatic data series generated from the rest of the meteorological stations from the study area is assumed to be homogeneous so that any climatic changes observed are as a result of natural occurrence. If data from a base station is homogeneous with respect to other stations, then the cumulative residuals from that data should not be biased (Allen et al. 1998). The bias hypothesis can be tested for a given probability p. This is done by confirming whether the residuals can be contained within an ellipse with axes β and α. The procedure for the analysis to be undertaken in this study is presented as follows (Allen et al. 1998):

(i) Regionalized homogeneous data set will be created by averaging the observations at various synoptic stations within the study area denoted by X_i and the base station as Y_i.
(ii) A regression analysis will be performed on a plot of the base station against the mean of the other stations.
(iii) The regression generated in 2 above will be used to compute estimated values of Y_i (Y_{est}) values in reference to the X_i data set.
(iv) Residuals of the observed Y values to the regression line will be computed, thus (Y_i-Y_{est}) and accumulated.
(v) A probability for accepting the hypothesis of homogeneity will be selected. The value of non-exceedance q = 0.95 (or 95 %) is desirable according to (Allen et al., 1998).
(vi) Compute the parameters α and β as follows:

$$\alpha = \frac{n}{2}$$

where n is the number of years under consideration

$$\beta = \frac{n}{\sqrt{n-1}} Z_p S_{xy}$$

where Z_p and S_{xy} is the standard normal variate for the probability of p and accumulated residual standard deviation, respectively.

The parametric equation of the ellipse is given by:

$$X = \alpha \cos \theta$$

$$Y = \beta \sin \theta$$

With ɵ varying from 0 to 2 Π

(vii) Plots of cumulative residuals and the ellipse against time will be done and if the graph of the cumulative residuals lies inside the ellipse, the hypothesis of homogeneity is accepted at the 95 % level of confidence.

2.1.2 INTERVENTION ANALYSIS

An intervention analysis will be undertaken to identify any possible changes (or point of intervention) in climate regimes for stations across the basin. This will be carried out using the cumulative deviation test. The test is based on rescaled cumulative sum of the deviations from the mean. This technique has been used in detecting changes in rainfall (Parida and Moalafhi, 2008, Kenabatho et al., 2012 and Byakatonda et al., 2018). If the annual time series are represented by $x_1, x_2, x_3, \ldots\ldots, x_n$ over n number of years, the computed CUSUM value Y_i at any time i is given by:

$$\boldsymbol{Y_i = x_i + x_{i-1} + x_{i-2} + \cdots x_1 - i.\frac{\sum_i^n x_i}{n}} \qquad (2.1)$$

The computed CUSUM values are plotted against time and examined for any suspected point of intervention. When the series under test is free from any interventions, the plot should normally oscillate around the horizontal axis. A observable decline or rise of this plot would suggest the possibility of intervention from the year of observation (corresponding to the relevant 'i') of such a change (Parida and Moalafhi, 2008). Positive slopes on these charts indicate a period of above average values and negative indicate otherwise (Byakatonda et al., 2018). Figure 2.1 shows an example of a possible point of intervention/change (shown by a red arrow).

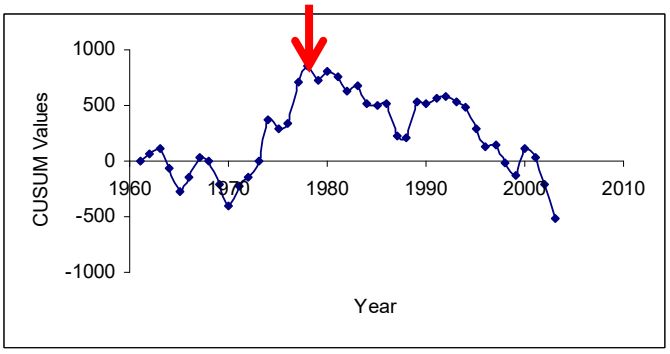

Figure 2-1: CUSUM plots showing a point of intervention/change

2.1.2.1 TEST FOR INTERVENTION ANALYSIS

If any intervention is detected, the two series (i.e. before and after the point of intervention) will be split into two no overlapping samples to be further investigated using split sample analysis and the student t test for investigating the significance of intervention. A suitable statistic for testing the null hypothesis $H_0: \overline{X}_1 = \overline{X}_2$ for the two samples with means $\bar{X}_1$ and $\overline{X}_2$ is given in Parida and Moalafhi (2008) and Kenabatho et al. (2012)

$$t_t = \frac{|\overline{X}_1 - \overline{X}_2|}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{(n_1+n_2-2)}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \qquad (2.2)$$

Where n is the number of the data set and s is the standard deviation. The null hypothesis is usually accepted within the two sided critical region U is:

$\{-\infty, t(v, 2.5\%)\} U \{t(v, 97.5\%), +\infty\}$ With $v = (n_1 - 1 + n_2 - 1)$ degrees of freedom.

2.1.3 TREND ANALYSIS

2.1.3.1 MONOTONIC TREND TEST

Tests for the detection of significant trends in climatologic time series can be categorized as parametric and non-parametric. Parametric trend tests require data to be independent and normally distributed, while non-parametric trend tests require only that the data be independent (Gocic and Trajkovic, 2013). Non-parametric methods such as Mann-Kendall will be used to detect trends in the data for the basin.

The method has been used extensively across different climatic zones to assess the significance of trends in hydro-meteorological time-series data like in (Modarres and de Paulo Rodrigues da Silva, 2007; Kampata et al., 2008; Parida and Moalafhi, 2008; Petrow and Merz, 2009; Wang et al., 2015; Byakatonda et al., 2018).The Mann-Kendall test statistic has been shown to be more

robust than parametric tests when dealing with skewed data and outliers in a data series (Helsel and Hirsch, 1992). The Mann-Kendall statistic S is given by;

$$S = \sum_{i}^{n-1} \sum_{j=i+1}^{n} sgn(x_j - x_k) \quad (2.3)$$

Where n is the number of data points, x_k and xj are the data values in time series k and j (j>k), respectively and $sgn(x_j - x_k)$ is the sign function as:

$$sgn(x_j - x_k) = \begin{cases} +1 & if\ (x_j - x_k) > 0 \\ 0 & if\ (x_j - x_k) = 0 \\ -1 & if\ (x_j - x_k) < 0 \end{cases} \quad (2.4)$$

The test statistic represents the number of positive differences minus the number of negative differences for all the differences between adjacent points in the time series.

For circumstances where the sample size n>10, the standard normal test statistic Z_S is computed using

$$Z_s = \begin{cases} \frac{S-1}{\sqrt{Var(S)}} & if\ S > 0 \\ 0 & if\ S = 0 \\ \frac{S+1}{\sqrt{Var(S)}} & if\ S < 0 \end{cases} \quad (2.5)$$

Where Var(s) is the variance of the sample also given by

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{i}^{m} t_i(t_i-1)(2t_i+5)}{18} \quad (2.6)$$

Where n is the number of data points, m is the number of tied groups and t_i denotes the number of ties of extent i. A tied group is a set of sample data having the same value. Taking a data set of 2, 3, 1, 3, 1 and 3 ,
t_3=1(one set of three values), t_2=1(one set of two value) and t_1=1 (one set of untied values)
Positive values of Z_s designate an upward trend and negative values otherwise.
Testing trends is done at the specific α significance level. When $|Z_S| > Z_{1-\alpha/2}$, the null hypothesis of no trend is rejected and a significant trend exists in the time series. $Z_{1-\alpha/2}$ is obtained from the standard normal distribution table. In this research, significance levels α=0.01 and α=0.05 will be used.

1.1.6 SEN'S SLOPE ESTIMATOR

The Mann-Kendall's trend test is only able to indicate the direction of the trend and as such does not reveal the significance of the trend. The test for significance of a linear dependence between two continuous variables, the meteorological data and time (Y and X) will be investigated by determining whether the regression slope coefficient for the explanatory variable is significantly different from zero by applying the Sen's slope criteria (Helsel and Hirsch, 1992).

Testing the significance of the slope of trend in the sample of N pairs of data will be accomplished using the non-parametric procedure developed by Sen (1968). The slope of N pairs of data sets is given by:

$$\boldsymbol{Q_i = \frac{x_j - x_k}{j-k}\ for\ i = 1, 2, \ldots\ldots, N} \qquad (2.7)$$

Where x_j and x_k are the data values at times j and k (j>k), respectively.

Should there be single measurements for each time period, then $N = \frac{n(n-1)}{2}$;
Where n is the number of time periods. If there are multiple measurements in one or more time periods, then $N < \frac{n(n-1)}{2}$

The N values of Q_i are ranked in ascending order with the median of slope or Sen's slope estimator computed as

$$\boldsymbol{Q_{med} = \begin{cases} Q_{\frac{(N+1)}{2}}\ if\ N\ is\ odd \\ \dfrac{Q_{\left(\frac{N}{2}\right)} + Q_{\left[\frac{(N+2)}{2}\right/_2\right]}}{2}\ if\ N\ is\ even \end{cases}} \qquad (2.8)$$

The Q_{med} sign reflects data trend reflection, while its value indicates the steepness of the trend (Gocic and Trajkovic, 2013). To determine whether the median slope is statistically different than zero, a confidence interval of Q_{med} at specific probability will be obtained.
The confidence interval about the time slope as applied in Gocic and Trajkovic (2013) while they were analyzing significance of trends is given by

$$\boldsymbol{C_\alpha = Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{Var(S)}} \qquad (2.9)$$

Where Var(S) is defined earlier and $Z_{1-\alpha/2}$ is the statistic obtained from the standard normal table. In this study, confidence intervals will be established at two significance levels (α=0.01 and α=0.05).
From C_α above, the $M_1 = \frac{N-C_\alpha}{2}$ and $M_2 = \frac{N+C_\alpha}{2}$ are obtained. The lower and upper limits of the confidence interval, Q_{min} and Q_{max}, are the M_1^{th} largest and the $(M_2+1)^{th}$ largest of the N ordered

slope estimates. The slope Q_{med} statistically differs from zero if the two limits (Q_{min} and Q_{max}) have similar sign (Gocic and Trajkovic, 2013)..

The Sen's slope estimator method has found use in a number of analyses of hydro-meteorological time series, including recent studies (Petrow and Merz, 2009; Tabari et al., 2011; Gocic and Trajkovic, 2013).

2.2 DRYNESS/WETNESS INDICES

Two indices will be considered in this report. These are

(i) Standardized Precipitation Index (SPI), or
(ii) Standardized Precipitation Evaporation Index (SPI).

If there is sufficient climate data representative of the catchment, including temperature (minimum and maximum), the SPEI will be preferred. Otherwise, the SPI will be used as an alternative method. These two methods are discussed below.

2.2.1 STANDARDIZED PRECIPITATION INDEX

For this task, the method for computing standardized precipitation index (SPI) will be based on the approach by McKee et al. (1993) and Edwards et al., (1997) to study relative departures of precipitation from normality. This method has been widely applied in many studies across different climate regions (Vicente-Serrano, 2006; Vicente-Serrano et al., 2010; Guenang and Mkankam Kamga, 2014). Monthly precipitation will be aggregated at various time scales (1,3,6,12,18, and 24 months). An illustration is provided by Guenang and Mkankam Kamga, (2014), e.g. for a 3-month time scale, the precipitation accumulation from month j-2 to month j is summed and attributed to month j (Guenang and Mkankam Kamga, 2014). At this time scale, the first two months of the data time series are missing. This is followed by a normalization procedure, in which an appropriate probability density function is first fitted to the long term time series of aggregated precipitation. Then the fitted function is used to calculate the cumulative distribution of the data points, which are finally transformed into standardized normal variates. This procedure is repeated for the desired time scales (Guenang and Mkankam Kamga, 2014).

For the case study under consideration, various distributions will be tested and the L-Moments and Probability Weighted Moments (PWM) will be used to fit the chosen probability distribution (whose choice will depend on the results of the goodness-of-fit test statistic used).

2.2.2 STANDARDIZED PRECIPITATION EVAPORATION INDEX (SPEI)

The SPEI ((Vicente-Serrano et al., 2010) method will be used. It has been credited for incorporating evaporation into the standard SPI thereby allowing for inclusion of water balance. The method has found wide application across many areas of different climatic zones (Vicente-Serrano et al., 2010; Yu et al., 2014; Wang et al., 2015, Byakatonda et al., 2016). The following procedure is adopted for computation of SPEI:

i) Determination of the Potential Evapotranspiration ETo
ii) Accumulation of climate water balance (D_i) at different time scales (i.e. P_i-ETo_i)

iii) Normalization of the water balance into a probability distribution function to obtain the SPEI index series.

2.2.2.1 DETERMINATION OF THE POTENTIAL EVAPOTRANSPIRATION (ET_0)

From the numerous existing ETo equations, the FAO-56 application of the Penman-Monteith (PM) equation (Allen et al., 1998) has received wide application. The method is widely used and recommended by the Food and Agricultural organization FAO Statistics (1998), the International Commission on Irrigation and Drainage (ICID) and the American Society of Civil Engineers (ASCE) as a robust procedure because it is predominately a physically based method (Vicente-Serrano et al., 2010). A major limitation to the application of the PM however, is the relatively high data demand in the form of air temperature, wind speed, relative humidity, and solar radiation. These datasets however, are not always available for the desired case studies.

An alternative approach was developed by Hargreaves (1994) where only mean maximum and mean minimum air temperature and extraterrestrial radiation are required. The extraterrestrial radiation in the computation can be calculated for a certain day and location. Due to the limited data availability, the Hargreaves (1994) method offers a better alternative. The study will follow the approach as implemented by Droogers and Allen (2002),, and Byakatonda et al (2018). The following equation will be used to compute ET_o:

$$\boldsymbol{ET_0 = 0.0023(T_{mean} + 17.8)(T_{max} - T_{min})^{0.5} * 0.408R_a} \quad (2.2)$$

Where,

T_{mean} is the mean air temperature (^0C),

T_{max} is the maximum air temperature (^0C),

T_{min} is the minimum air temperature (^0C)

R_a is extraterrestrial radiation [MJ m^{-2} day^{-1}] and is given by:

$$\boldsymbol{R_a = \frac{24(60)}{\pi} G_{sc} d_r [\omega_s \, sin(\varphi) \, sin(\delta) + cos(\varphi) \, cos(\delta) \, sin(\omega_s)]} \quad (2.2)$$

Where,

G_{sc} is solar constant = 0.0820 MJ m^{-2} min-1,

d_r is inverse relative distance Earth-Sun (Equation 2.3),

ω_s is sunset hour angle (Equation 2.6) [rad],

φ is latitude [rad],

δ is solar declination (Equation 2.4) [rad].

The inverse relative distance Earth-Sun, d_r, and the solar declination, δ, are given by:

$$\boldsymbol{d_r = 1 + 0.033 \, cos\left(\frac{2\pi}{365} J\right)} \quad (2.3)$$

$$\delta = 0.409\, sin\left(\frac{2\pi}{365}J - 1.39\right) \quad (2.4)$$

Where J is the average Julian day of the month given by

$$J = \begin{cases} \left(\frac{275M}{9} - 30 + D\right) - 2 & M > 3 \\ \left(\frac{275M}{9} - 30 + D\right) & M < 3 \\ \left(\frac{275M}{9} - 30 + D\right) + 1 & M > 2\ for\ leap\ year \end{cases} \quad (2.5)$$

D=15 for average month

The sunset hour angle, ω_s, is given by:

$$\omega_s = cos^{-1}[-tan(\varphi)tan(\delta)] \quad (2.6)$$

2.2.2.2 ACCUMULATION OF CLIMATE WATER BALANCE (D_I) SERIES

With ETo established, the monthly water balance will be calculated as a difference between Precipitation (P_i) and evapotranspiration (ET_{oi}) as follows:

$$D_i = P_i - ET_{0i} \quad (2.7)$$

Where,
P is monthly precipitation
i is month under consideration

The calculated D_i values will be aggregated at different time scales, following the same procedure as that for the SPI. The difference $D_{j,i}^k$ in a given month j and year i depends on the chosen time scale k. For example, the accumulated difference for one month in a particular year i with a 12-month time scale is calculated using

$$X_{j,i}^k = \sum_{l=13-k+i}^{12} D_{i-1,l} + \sum_{l=1}^{j} D_{i,l} \quad if\ j < k\ and \quad (2.8)$$

$$X_{j,i}^k = \sum_{l=j-k+1}^{j} D_{i,l} \quad if\ j \geq k \quad (2.9)$$

Where $D_{i,l}$ is $P_i - ET_{0i}$ the difference in the first month of year j, in millimeters and $D_{j,i}^k = X_{j,i}^k = D_i$ series.

2.2.2.3 NORMALIZATION OF THE WATER BALANCE SERIES

In quantifying SPEI a three parameter distribution will be used, since in the two parameter distributions the variable (D_i) has a lower boundary of zero ($0 > D < \infty$) which is the case in SPI that uses precipitation series with a lower minimum of 0.0 mm (Potop et al., 2010; Vicente-Serrano et al., 2010), whereas in three parameter distributions x in this case D_i series can take values in the range $\gamma > D < \infty$, where γ is the parameter of origin of the distribution, consequently D can have negative values, in an event of a deficit climate balance. The water balance series were normalized as twelve independent series. The method of L-Moments in combination with

probability weighted moments (PWMs) will be used for parameter estimation of the probability distribution functions fitting the D_i series. Parameters resulting from this procedure are more stable against possible outliers in the series data (Haktanira and Bozduman, 1995). The unbiased probability weights as suggested by (Hosking et al., 1985) and presented in Haktanira and Bozduman (1995) is given by;

$$\boldsymbol{P_i^r} = \frac{(\boldsymbol{i-1})(\boldsymbol{i-2})\ldots\ldots(\boldsymbol{i-r})}{(\boldsymbol{n-1})(\boldsymbol{n-2})\ldots\ldots(\boldsymbol{n-r})} \quad \textit{(2.10)}$$

Where,

P_i^r is the probability weight,
i is the rank assigned to the data series arranged in ascending order,
n are number of observations,
r is the order

The L-moments permit the comparison of various candidate distributions frequency (Hosking and Wallis, 2005). To identify the candidate distributions, L-moment ratios (L-Skewness τ_3 and L-Kurtosis τ_4) will be calculated as follows;

$$\tau_3 = \frac{\lambda_3}{\lambda_2} \quad \textit{(2.11)}$$

$$\tau_4 = \frac{\lambda_4}{\lambda_2} \quad \textit{(2.12)}$$

and note that the L-coefficient of variance, $L - C_v(\tau_2)$; is given by:

$$\tau_2 = \frac{\lambda_2}{\lambda_1} \quad \textit{(2.13)}$$

λ_2, λ_3 and λ_4 are L-moments of the of the D_i series computed from probability weighted moments (PWMs) as indicated in equations below:

$$\lambda_1 = M_0 \quad \textit{(2.14)}$$

$$\lambda_2 = 2M_1 - M_0 \quad \textit{(2.15)}$$

$$\lambda_3 = 6M_2 - 6M_1 + M_0 \quad \textit{(2.16)}$$

$$\lambda_4 = 20M_3 - 30M_2 + 12M_1 - M_0 \quad \textit{(2.17)}$$

These four moments are analogous to the first four conventional moments of X (i.e. mean, variance, skewness and kurtosis).

The PWMs of order r are given by,

$$M_r = \frac{1}{N}\sum_{i=1}^{N} P_i^{\,r}.X_i \qquad (2.18)$$

$$SPEI = W - \frac{C_0+C_1W+C_2W^2}{1+d_1W+d_2W^2+d_3W^3} \qquad (2.19)$$

Where

$W = \sqrt{-2\ln(P)}$ for $P \leq 0.5$ and P is the probability of exceeding a determined D value. The P value is obtained from $P = 1 - F(x)$. If P>0.5, then P is replaced by 1-P and the sign of the resultant SPEI is reversed. The constants are C_0=2.515517, C_1=0.802853, C_2=0.010328, d_1=1.432788, d_2=0.189269, d_3=0.001308

The average value of SPEI is 0, and the standard deviation is 1. The SPEI is a standardized variable, and it can therefore be compared with other SPEI values over time and space. For each time scale, each drought event (period in which SPEI is continuously negative and SPEI ≤ -1), will be defined through its *(i) duration* (time from the beginning to the end), (ii) *severity* (SPEI value for each month following a given classification), *(iii) magnitude* (SPEI sum for each month and for the duration of the severity), *(iv) intensity* (magnitude/duration ratio of the event).

2.2.3 ARIDITY INDEX (AI)

Among the indices defined for the determination of agricultural water demands (World Meteorological Organization, 1975) is the aridity index introduced by (de Martonne 1926). Its monthly values are described by the following equation as presented in Livada and Assimakopoulos (2007):

$$\boldsymbol{I_i = \frac{12P_i}{(10+T_i)}} \qquad (2.20)$$

Where, P_i is the monthly precipitation amount and T_i is the respective mean monthly air temperature.

The purpose of this index (de Martonne 1926) is to identify the months for which actual evapotranspiration starts to drop below the potential evapotranspiration leading to shortness of water for optimal plant growth, which in turn depends both on rainfall and ambient temperature. Thus, irrigation according to this index becomes necessary when I_i<20. AI will there be determined and the months for which I_i<20 will be highlighted.

2.3 MAPPING OF DRYNESS INDICES

The results obtained for SPI/SPEI and AI will be compiled and mapped to obtain spatially interpolated indices over the Zambezi basin, as well as time series for selected stations representative of the basin. In addition, other variables such as trends, variability will be plotted for the basin.

2.4 EXTREME EVENTS ANALYSIS

2.4.1 REGIONAL FLOOD FREQUENCY ANALYSIS

To study and characterize extreme flow events in the basin, there is need to develop regional growth curves by identifying suitable regional distribution curves for estimation of floods of different return periods for the basin. This will be achieved using L-moments (Hosking, 1990) which has been credited for providing parameter estimates that are nearly unbiased and highly efficient and thus are better suited for use in constructing moment diagrams..

2.4.1.1 L-MOMENTS

Hosking (1990) introduced L-moments as a linear combination of PWMs. The PWMs, defined by Greenwood et al. (1979) for a non-negative integer may be given as:

$$\boldsymbol{\beta_r = E[x\{F(x)\}^r]} \qquad (2.21)$$

which can be written as:

$$\boldsymbol{\beta_r = \int_0^1 x\,(F)F^r dF, r = 0, 1, 2, \ldots} \qquad (2.22)$$

where, $F = F(x)$ is the cumulative distribution function (CDF) for x, $x(F)$ is the inverse CDF of x evaluated at the probability F. When $r = 0$, β_0 is equal to the mean of the distribution $\mu = E[x]$.

Hosking (1990) defined r^{th} L-moments related to the r^{th} PWMs as:

$$\boldsymbol{\lambda_{r+1} = \sum_{k=0}^{r} \beta_\kappa (-1)^{r-\kappa} \binom{r}{\kappa} \binom{r+\kappa}{\kappa}} \qquad (2.23)$$

Using the ranked data (in the ascending order), an unbiased estimate of sample Probability Weighted Moments (PWM) can be computed from Equation (2.10). The L-moments and their PWMs counter parts are given in Equations *(2.14 to 2.17).*

Computed values of τ_3 *and* τ_4 (i.e. L-skewness $L - C_s$ $(\tau_{3)}$ and L-kurtosis, $L - C_\kappa$ (τ_4)) are then plotted on to the Theoretical L-Moment ratio diagram (shown as +) which suggests the likely/possible statistical distribution which could be used for analysis. The theoretical L-moment ration diagram for five commonly used distributions of three parameters is presented in Figure 2.1.

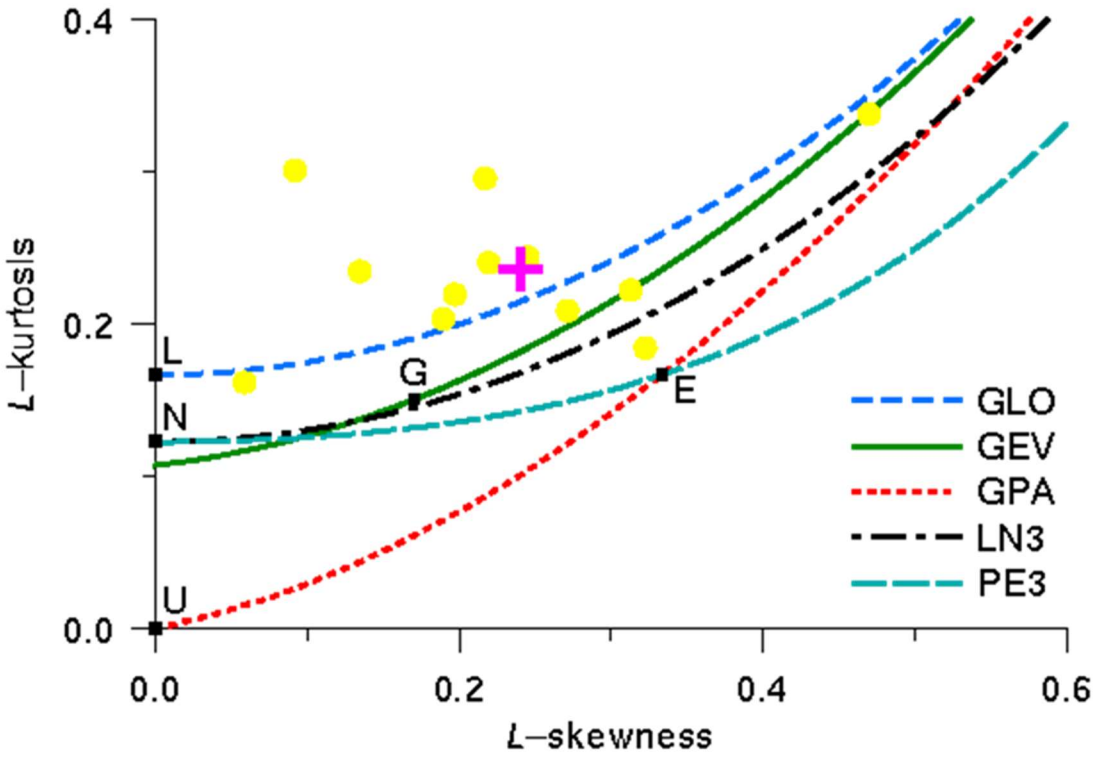

Figure 2-2. Theoretical plots of L-skew versus L-kurtosis diagram for some common statistical distributions (viz: Generalised Pareto (GPA), Generalised Extreme Values (GEV), Generalised Logistic (GLO), 3 Parameter Log-normal (LN3), Pearson Type 3 (PE3) Distribution) (Hosking and Wallis, 1996)

The flood quantiles for each return period for the basin (*or for regions identified as homogeneous within the basin*) will be computed based on the identified suitable probability distribution to produce regional growth curve.

3 DATA NEEDED

3.1 OBSERVED DATA

In order to implement the proposed methodology, the following data is needed (Table 3-1):

Table 3-1: Milestones and dates

Method	Data Needed	Temporal Resolution	Data format	Period of Analysis
SPI	Rainfall data/Remote sensed data (eg TRMM)or Reanalysis data (eg NCEP)	Monthly	Raster or Point data	1980-2016
SPEI	Rainfall, Temperature (Mean, Maximum and Minimum), Sunshine hour data/Remote sensed data (eg TRMM)or Reanalysis data (eg NCEP)	Monthly	Raster or Point data	
AI	Rainfall, Temperature (Mean) /Remote sensed data (eg TRMM)or Reanalysis data (eg NCEP)	Monthly	Raster or Point data	
RFFA	Streamflow data & Flood data/ Reanalysis data (eg NCEP)	Daily/monthly	Point data/Time/ Gridded	
Intervention Analysis	Rainfall, Temperature (Min, Max), Relative Humidity/Remote sensed data (eg TRMM)or Reanalysis data (eg NCEP)	Monthly	Point data/Time Series/Raster/Gridded	

3.2 ALTERNATIVE DATA PRODUCTS AND SOURCES

3.2.1 REMOTE SENSING AND SATELLITE DATA

In the absence of observed datasets, remote sensing data will be used. This includes, for example the Tropical Rainfall Measuring Mission (TRMM) rainfall products (https://pmm.nasa.gov/data-access/downloads/trmm).

3.2.2 GRIDDED AND REANALYSIS DATA

Gridded climate data from the Climatic Research Unit (CRU) will be used. This includes monthly temperature and rainfall from 1901 to 2015, available at https://climatedataguide.ucar.edu/climate-data/cru-ts-gridded-precipitation-and-other-meteorological-variables-1901.

In addition, reanalysis data (i.e. the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) known as the NCEP/NCAR data, https://www.esrl.noaa.gov/psd/data/reanalysis/reanalysis.shtml) will be used.

These datasets will be validated using data from neighboring areas or areas with similar environmental setting as a form of ground truthing. .

4 EXPECTED OUTPUTS AND TIME LINES

Table 4-1 below shows the milestones and dates as per the contract.

Table 4-1: Milestones and dates

Milestone	Activity Description	Milestone Date
M-1.0	**Inception Report**	**Dec 2017**
M-1.1	*PowerPoint Presentation on CV and Extreme Events*	*Nov 2017*
M-2.0	**Draft Report and Database**	**Sept 2018**
M-2.1	*Data Collection and Collation*	*Jan-Feb 2018*
M-2.2	*Data Quality checks*	*Feb-2018*
M-2.3	*Data analysis and model development*	*Mar-Aug 2018*
M-3.0	Report, database and model on Climate Variability and Extreme Events Analysis	**Nov 2018**
M-3.1	*Report Writing-Draft*	*Sept-Oct 2018*

5 REFERENCES

1) Allen, R. G., Pereira, L., Raes, D. and Smith, M. (1998). FAO Irrigation and drainage paper No. 56. Rome: FAO.
2) Beck, L., Bernauer, T (20110. How will combined changes in water demand and climate affect water availability in the Zambezi river basin? Global Environmental Change 21 (2011) 1061–1072.
3) Beilfuss (2012). A Risky Climate for Southern African Hydro: Assessing hydrological risks and consequences for Zambezi River Basin dams. International Rivers. Berkeley, CA 94704, USA.
4) Byakatonda, J., Parida, B., Kenabatho, P. and Moalafhi, D. (2016). Modelling dryness severity using artificial neural networks at the Okavango delta, Botswana. GLOBAL NEST JOURNAL, 18(3), 463-481.
5) Byakatonda, J., Parida, B., Kenabatho, P. K. and Moalafhi, D. (2018). Influence of climate variability and length of rainy season on crop yields in semiarid Botswana. Agricultural and Forest Meteorology, 248, 130-144.
6) Dahmen, E.R., Hall, M.J. (1990). Screening of Hydrological Data: Tests for Stationarity and Relative Consistency Volume 49 of ILRI publication, ISSN 0167-4072 Issue 49 of Publication, International Institute for Land Reclamation and Improvement. 58 p.
7) De Martonne, E. (1926). A new climatological function: The Aridity Index. Gauthier-Villars, Paris, France.
8) Droogers, P. and Allen, R. G. (2002). Estimating reference evapotranspiration under inaccurate data conditions. Irrigation and drainage systems, 16(1), 33-45.
9) Edwards, C. D. C., T. B. McKee, N. J. Doesken, and J. Kleist (1997). "Historical analysis of drought in the United States." In 7th conference on climate variations, 77th AMS annual meeting, vol. 27. 1997.
10) Engelbrecht, F.A., Landman, W.A., Engelbrecht, C.J., Landman, S., Bopape, M.M., Roux, B., McGregor, J.L., Thatcher, M. (2011). Multi-scale climate modelling over Southern Africa using a variable-resolution global model. Water SA Vol. 37 No. 5. 647-658.
11) FAO Statistics. (2009). Statistical data base of Food and Agriculture Organization of the United Nations.

12) Gocic, M. and Trajkovic, S. (2013). Analysis of changes in meteorological variables using Mann-Kendall and Sen's slope estimator statistical tests in Serbia. Global and Planetary Change, 100, 172-182.
13) Greenwood, J.A.; Landwehr, J.M.; Matalas, N.C.& Wallis, J.R.(1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. Water Resources Research 15, 1049–1054.
14) Guenang, G.M., Mkankam Kamga, F (2014) Computation of the Standardized Precipitation Index (SPI) and its use to assess drought occurrences in Cameroon over Recent Decades. Journal of Applied Meteorology and Climatology, 53: 2310-2324.
15) Haktanir, T., & Bozduman, A. (1995). A study on sensitivity of the probability-weighted moments method on the choice of the plotting position formula. Journal of Hydrology, 1(168), 265-281.
16) Hargreaves, G. H. (1994). Defining and using reference evapotranspiration. Journal of Irrigation and Drainage Engineering, 120(6), 1132-1139.
17) Helsel DR, Hirsch RM (1992) Statistical methods in water resources, studies in environmental science, vol 49. Elsevier, New York, 522 p.
18) Hosking, J. R. M. and Wallis, J. R. (2005). Regional frequency analysis: an approach based on L-moments: Cambridge University Press.
19) Hosking, J. R., Wallis, J. R., & Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. Technometrics, 27(3), 251-261.
20) Hosking, J.R.M. (1990). L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. Journal of the Royal Statistical Society. Series B (Methodological), Vol. 52 (1), 105-124.
21) Hosking, J.R.M., Wallis, J.R. (1996). Regional Frequency Analysis: An Approach Based on L-moments. Cambridge University Press.
22) Kenabatho, P.K., Parida, B.P., and Moalafhi, D.B. (2012). The value of larger scale climate variables in climate change assessment: the case of Botswana's rainfall. Journal of Physics and Chemistry of the Earth, 50–52 (2012), 64–71. doi:10.1016/j.pce.2012.08.006.

23) Livada, I. and Assimakopoulos, V. (2007). Spatial and temporal analysis of drought in Greece using the Standardized Precipitation Index (SPI). Theoretical and applied climatology, 89(3-4), 143-153.
24) McKee, T. B., Doesken, N. J. and Kleist, J. (1993). The relationship of drought frequency and duration to time scales. Paper presented at the Eighth Conference on Applied Climatology, Anaheim, CA, Boston.
25) Modarres, R. and da Silva, V. d. P., Rodrigues. (2007). Rainfall trends in arid and semi-arid regions of Iran. Journal of Arid Environments, 70(2), 344-355.
26) Moore, A.E., Cotterill, F.P. D, Main, M.P.L., Williams, H.B. (2007). The Zambezi River. In *Large Rivers: Geomorphology and Management*, Chapter 15. Edited by A. Gupta, John Wiley & Sons, Ltd.
27) Parida, B. P. and Moalafhi, D. B. (2008). Regional rainfall frequency analysis for Botswana using L-Moments and radial basis function network. Physics and Chemistry of the Earth, Parts A/B/C, 33(8), 614-620.
28) Petrow, T., Merz, B. (2009): Trends in flood magnitude, frequency and seasonality in Germany in the period 1951 - 2002. - Journal of Hydrology, 371, 1-4, 129-141.
29) Potop, V., Türkott, L., Kožnarová, V. and Možný, M. (2010). Drought episodes in the Czech Republic and their potential effects in agriculture. Theoretical and applied climatology, 99(3-4), 373-388.
30) Schlosser, A., Strzepek, K (2015) Regional climate change of the greater Zambezi River basin: a hybrid assessment. Climatic Change, 130: 9-19.
31) Sen, P.K., (1968). Estimates of regression coefficient based on Kendall's tau: J. Am. Stat. Assoc., 63, 1379-1389.
32) Serdeczny, O., Adams, S., Baarsch, F., Coumou, D., Robinson, A., Hare, W., Schaeffer, M., Perrette, Reinhardt, J (2016). Climate change impacts in Sub-Saharan Africa: from physical changes to their social repercussions. Regional Environmental Change, 15(8). DOI: 10.1007/s10113-015-0910-2.
33) Tabari, H., Somee, B. S. and Zadeh, M. R. (2011). Testing for long-term trends in climatic variables in Iran. Atmospheric Research, 100(1), 132-140.
34) Vicente-Serrano, S. M., 2006: Differences in spatial patterns of drought on different time scales: An analysis of the Iberian Peninsula. Water Resour. Manage., 20, 37–60.

35) Vicente-Serrano, S. M., Beguería, S. and López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. Journal of Climate, 23(7), 1696-1718.
36) Wang, W., Zhu, Y., Xu, R. and Liu, J. (2015). Drought severity change in China during 1961–2012 indicated by SPI and SPEI. Natural Hazards, 75(3), 2437-2451.
37) Wang, W., Zhu, Y., Xu, R. and Liu, J. (2015). Drought severity change in China during 1961–2012 indicated by SPI and SPEI. Natural Hazards, 75(3), 2437-2451.
38) WMO. (2009). Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation (Vol. WCDMP - No. 72. WMO – TD No. 1500): World meteorological Organization.
39) World Bank (2010). The Zambezi River Basin. A multi-sector investment opportunities analysis. Volume 1: Summary Report. Washington DC, USA.
40) Yu, M., Li, Q., Hayes, M. J., Svoboda, M. D. and Heim, R. R. (2014). Are droughts becoming more frequent or severe in China based on the standardized precipitation evapotranspiration index: 1951–2010? International Journal of Climatology, 34(3), 545-558.