

MODELING DRYNESS SEVERITY USING ARTIFICIAL NEURAL NETWORK AT THE OKAVANGO DELTA, BOTSWANA

BYAKATONDA J.^{1,3,*}

PARIDA B.P.¹

KENABATHO P.K.²

MOALAFHI D.B.²

¹Department of Civil Engineering, University of Botswana, Gaborone

²Department of Environmental Science, University of Botswana, Gaborone

³Department of Biosystems Engineering

Gulu University, Gulu, Uganda

Received: 22/07/2015

Accepted: 06/03/2016

Available online: 20/05/2016

*to whom all correspondence should be addressed:

e-mail: byakatondaj@hotmail.com

ABSTRACT

Water balance studies in the Okavango Delta indicate that more than 90% of inflow into the Delta is lost through evaporation. This coupled with high climatic variability threatens the ecohydrology of the Delta. Trends indicate decreasing rainfall amounts and increasing temperature at the area of the Delta. The main aim of this study was therefore to investigate long term trends and variability in rain onset, cessation, number of rainy days and their impact on the dryness index at the Delta. The impact of the above variables is expressed through the standardized precipitation and evaporation index (SPEI) quantified by aggregating the climate water balance and fitting monthly series to a generalized logistic distribution using L-Moments. The SPEI, determined at windows of different time scales of one, three and twelve months, provided an extensive evaluation of dryness severity and its impact on this sensitive ecosystem. Rain onset and cessation dates were generated from cumulative pentad rainfall–evapotranspiration relationships. Analysis of climatic data showed mean rain onset occurring in November and ceding in March with average of 44 rainy days between 1970/71 and 2013/14. The results revealed a decrease in the number of rainy days at a rate of 0.16 days/yr and of the duration of the rainy season at 0.25 days/yr with high variability. Annual rainfall was found to decrease at the rate of 1.60 mm/yr with 6.8% probability of failure in rainfall onset. Analysis further revealed that both extreme dryness and wetness are rare phenomena with probabilities of less than 1% and near normal conditions for 67% of the time for all SPEI time scales. Although gradual increase in dryness in the Delta is attributed to high climatic variability, simulations undertaken using Artificial Neural Networks did not predict any major changes in the next five years. However, vulnerability to severe droughts is not completely ruled out because of the high variability in rainfall and of the location of the Delta in a semi-arid zone.

Keywords: Artificial Neural Network, Ecohydrology, Rain cessation, Rain onset, Standardized precipitation evaporation index

1. Introduction

Dryness is a manifestation of metrological drought which transforms into agricultural and hydrological droughts (Sönmez *et al.*, 2005). Dryness can impact regional water supply, lead to loss in crop yield and even threatens wetland ecosystems. Since the severity of dryness is dependent upon the degree of moisture

deficiency, duration and extent of the affected area, attempts have been made to represent it through various indices like, the Palmer Drought Severity index (PDSI), the Standardized Precipitation Index (SPI) and more recently the Standardized Precipitation Evaporation Index (SPEI) (Mkhabela *et al.*, 2010; Vicente-Serrano *et al.*, 2010; Wang *et al.*, 2015).

The PDSI was intended mainly to monitor agricultural drought through quantification of moisture deficiency fluctuations. It measures dryness as the departure of the moisture supply from its long term normalized mean Palmer (1965) and has fixed temporal scales of nine to twelve months. McKee *et al.*, (1993), in their study observed drought to be multiscale in nature. Studies by Livada and Assimakopoulos (2007) and Potop and Možný (2011) observed that the timescale chosen for quantifying drought is important given the varied response times of different hydrological, agricultural and environmental systems to dryness. Their findings revealed that PDSI is limited in capturing drought severity at shorter time scales of one, three or even six months which is crucial in studying seasonal dryness conditions. To address the multiscale nature of drought, the SPI which was brought in by McKee *et al.* (1993) has been able to quantify dryness at one, three, six, nine, twelve, twenty four, thirty six and forty eight months. The simplicity of SPI is premised on the fact that precipitation is the only input variable. The use of just the precipitation amount has inadequacies in quantifying dryness since there is resounding scientific evidence of global warming as a result of temperature rise (Pachauri and Reisinger, 2007; Tabari *et al.*, 2011a). Studies by Vicente-Serrano *et al.* (2010) have indicated the importance of incorporating temperature in the computation of drought index. In lieu of this, the SPEI was recently developed to combine the response of PDSI to changes in evapotranspiration demands with the simplicity of calculating the multiscale nature of SPI through windows of different timescale (Vicente-Serrano *et al.*, 2010; Yu *et al.*, 2014; Wang *et al.*, 2015). Further studies by Stokes *et al.*, (1997) have shown that increased changes in the moisture regimes in the Kalahari which covers the iconic Okavango Delta of Botswana, have not only influenced the dune reactivation but also changed the extent of summer precipitation.

Predictions from climate models indicate a general decrease in precipitation and increasing temperatures with high variability over Southern Africa (Tsheko, 2004), hence the effect of temperature on drought severity cannot be ignored in this region. Model forecasts indicate remarkable reduction in stream flows making drought an increasing feature of Botswana's hydro-climatic regimes. Climatic studies in Botswana have reported a general decrease in rainfall amounts and increase in temperature since 1979/80 (Parida and Moalafhi, 2008; Batisani and Yarnal, 2010).

Even prediction of future dryness scenarios will facilitate future water management options and assess the vulnerability of the Delta towards climate variability. Drought prediction is an important factor in drought preparedness and mitigation measures. For proper water management, forecasts ranging from a few months to a few years play a vital role (Dastorani and Afkhami, 2011; Mishra and Singh, 2011). Linear parametric autoregressive (AR), moving-average (MA) and autoregressive moving-average (AMA) models introduced by Box and Jenkins (1976) have been widely used in prediction of time series (Machiwal and Madan Kumar, 2012). These models being linear are not able to handle non-stationary signals, and signals of nonlinear mathematical models (Diaconescu, 2008). It is against this background that research has seen increased use of artificial intelligence manifested through Artificial Neural Networks (ANNs) that mimic functioning of biological brain cells to handle complex systems that are characterized by non-linearity. Various ANN model formulations exist and have been tested in Illeperuma and Sonnadara (2009); Ardalani-Farsa and Zolfaghari (2010); Menezes and Barreto (2008) and Maier and Dandy (2000).

It has been demonstrated that Recurrent Neural Networks (RNN) with sufficient number of neurons are a realization of the Nonlinear Autoregressive moving average (NARMA) process (Gao and Meng Joo, 2005; Diaconescu, 2008; Menezes and Barreto, 2008).

To evaluate the possible consequences of climate change processes on the future availability of water resources in the Okavango Delta, it has been considered necessary to investigate the changes in dryness severity under different timescale windows using SPEI. Similarly variability in onset, cessation of rainfall dates and number of rainy days at the Okavango Delta was also investigated in this study. The study also makes a five year prediction of the dryness severity and wetness conditions at the Delta using the Nonlinear Autoregressive with exogenous inputs (NARX) ANN. This study attempts to determine dryness severity at the Okavango Delta using meteorological inputs.

2. Materials and methods

2.1. Study area and data

Okavango Delta is an iconic Ramsar protected area which inhabits a rich ecology and enormous wildlife. The Delta is located in Kalahari desert of Northern Botswana (Figure 1). The Okavango Delta is the world's largest protected Ramsar site sustained by the surplus between inflows and evapotranspiration. The Okavango River does not discharge into a surface water body, but rather in a Delta-like channel network and adjoining wetlands distributed over a gently sloping conical surface (McCarthy, 2006). It covers approximately 2500 km² of permanent wetland and up to 8000 km² of seasonal wetland (Bauer *et al.*, 2006). The wetland and its catchment form a dynamic, integrated system with many interfaces and feedback loops operating on different spatial and temporal scales. The main source of inflow to the wetland is through rainfall with more than 90% of its volume lost through evaporation mainly driven by climate variability (Farquharson *et al.*, 1990). Hence the main focus of this study is to assess the vulnerability of the Okavango Delta under recurrent dryness conditions. Wetland systems are particularly susceptible to hydrological changes, in the context of climate change, warranting particular attention (Burkett and Kusler, 2000; Acreman *et al.*, 2009).

Observations of daily rainfall, minimum and maximum monthly temperature values recorded at Maun have been used in this study. The data set spans a period of 44 years starting from 1970/71 to 2013/14 from Maun meteorological station. Missing values were estimated using the normal ratio method (De Silva *et al.*, 2007).

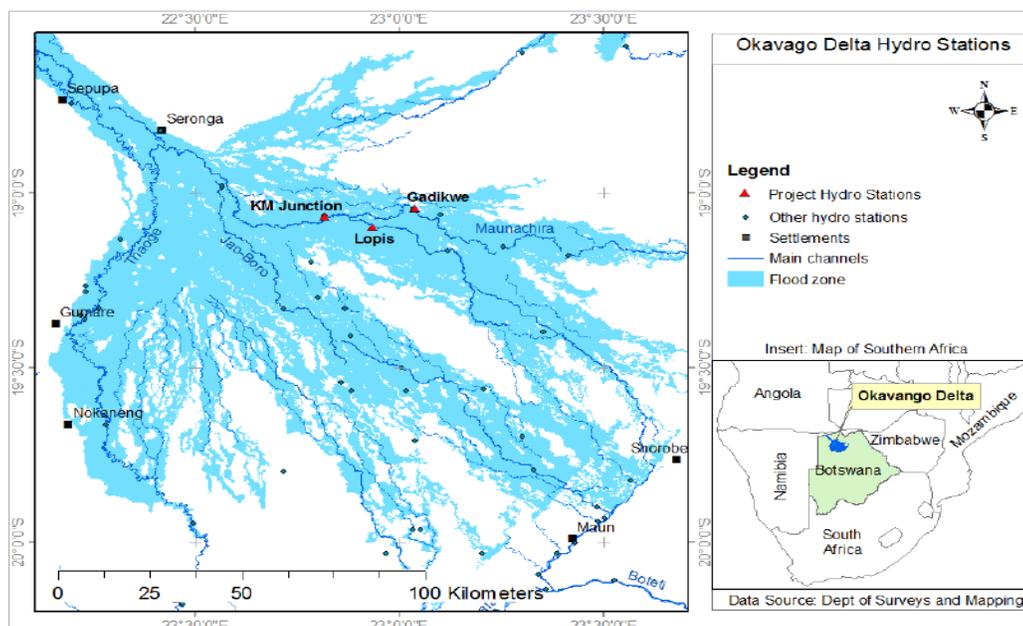


Figure 1. Location of the Okavango Delta (Moalafhi *et al.*, 2014)

2.2 Analysis of rainfall characteristics at Okavango Delta

The long term mean, standard deviation and coefficient of variation of annual rainfall are computed to assess the general rainfall characteristics in the study area. The number of rainy days, onset and cessation of rain are computed. Climatic variability was also assessed based on these properties. Onset and cessation of rain in this study are defined as the start and end of the rainy season respectively.

To determine the onset and cessation dates of rainfall, the criteria stipulated in Araya and Stroosnijder, (2011) have been used. These criteria use cumulative rainfall–evapotranspiration relationships to determine onset and cessation dates. In addition, these criteria consider the probability of failure in rainfall onset which in other words gives an indication of receiving inadequate rainfall amounts. Moeletsi and Walker (2012) used the probability of failure in rain onset as a measure of classifying duration of the rainy season. The probability of onset failure was computed as the ratio of the number of years in which the onset criteria were not realized to the total number of meteorological years under consideration.

Based on these premises, onset occurs when the long-term cumulative 5-day rainfall is greater than or equal to the cumulative half of the 5 days' potential evapotranspiration (Araya and Stroosnijder, 2011). The conditions are that this trend continues for at least two successive pentads and during this period the rainfall sum should not be less than 25 mm, and that no dry spell longer than 10 consecutive days occurs within 30 days of the onset.

Consequently, cessation in this study occurs after onset date when half of the pentad potential evapotranspiration exceeds the cumulative 5 days' rainfall. The conditions for cessation are that the date should occur seven days after this condition provided this period is followed by a ten day dry spell from the start of the deficit.

The length of the rainy season for this study is defined as the number of days between the onset and cessation dates of rain. The length of the rainy season is calculated by subtracting the starting date of the rain in Julian days from the number of days in a year and adding the number of Julian days for cessation of the rain. The number of rainy days in this study is defined as the number of days of the year with a rainfall amount ≥ 1 mm (Batisani and Yarnal, 2010; Ngetich *et al.*, 2014).

The wet/dry year, early/late onset, early/late cessation was determined following Singh (1999) criteria. A wet year was defined as $(R+\sigma)$ or more and a dry year as $(R-\sigma)$ or less where R is the long term mean of the time series for the 44 year period under consideration and σ the standard deviation of the series.

2.3 Determination of dryness severity

Dryness severity was quantified using the multiscalar Standardized Precipitation Evapotranspiration Index (SPEI) at timescales of one, three and twelve months. The SPEI was quantified based on the following: Computation of the Potential Evapotranspiration (ET_0), accumulation of climate water balance (D_j) at different time scales ($P-ET_0$) and normalization of the water balance into a generalized Log-logistic probability distribution to obtain the SPEI index series.

2.3.1 Determination of Potential Evapotranspiration (PET)

Data of wind speed, relative humidity and sunshine hours were not readily available for Maun synoptic station. For this reason the alternative approach of Hargreaves was adopted to determine PET for this study as applied in Droogers and Allen (2002) . The following equation was used to compute ET_0 :

$$ET_0 = 0.0023 \times 0.408 Ra (T_{\text{mean}} + 17.8) (T_{\text{max}} - T_{\text{min}})^{0.5} \quad (1)$$

Where T_{mean} , T_{max} and T_{min} are monthly mean, maximum and minimum air temperature ($^{\circ}\text{C}$) respectively, while Ra is the extraterrestrial radiation [$\text{MJ m}^{-2} \text{day}^{-1}$]

2.3.2 Accumulation of climate water balance (D_j) series

With ET_o established, the monthly climatic water balance was calculated as a difference between precipitation (P_i) and evapotranspiration (ET_{0j}) as follows:

$$D_j = P_j - ET_{0j} \tag{2}$$

Where, P is the monthly precipitation and j the month under consideration.

The calculated D_j values were aggregated at different time scales of one, three and 12 months. The homogeneity of the aggregated D_j series will vary from month to month Vicente-Serrano *et al.* (2010), for this reason, D_j series are split into twelve data series corresponding to calendar months. This also minimizes the effect of serial correlation that is introduced in the process of aggregating through moving windows. 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

$$X_{j,i}^K = \sum_{l=13-K+j}^{12} D_{j-1,l} + \sum_{l=1}^j D_{j,l} \text{ if } j < K \text{ and} \tag{3}$$

$$X_{j,i}^K = \sum_{l=j-K+1}^j D_{j,l} \text{ if } j \geq K \tag{4}$$

Where $D_{j,i}$ is $P_j - ET_{0j}$, the difference in the first month of year i , in mm and

$D_{j,i}^K = X_{j,i}^K = D_j$ aggregated series.

2.3.3 Normalization of the water balance series

The water balance series were normalized as twelve independent series. In quantifying SPEI a three parameter distribution was used, since D_j can take on negative values in periods of deficit climate water balance (Vicente-Serrano *et al.*, 2010; Wang *et al.*, 2015).

Although a number of methods exist for estimation of parameters of a probability distribution function fitted to data series, it has always been a challenge in selecting an appropriate probability distribution function. When sample sizes are large, the methods of maximum likelihood and probability weighted moments (PWMs) have both been found to give almost similar results at desired probabilities (Kysely, 2002; Beguería *et al.*, 2014). In fact the method of PWMs has been found to be more robust while dealing with small samples and samples containing outliers (Hoskings and Wallis, 2005; Stedinger and Griffis, 2008). Thus, L-Moments resulting from PWMs were used to estimate the parameters of the chosen statistical model fitted to the D_j series as they produce the least biased estimates. L-moments allow better identification of the original probability distribution function that adequately models the D_i series. Unbiased probability weights as suggested by Hosking *et al.* (1985) and presented in Haktanira and Bozduman (1995) and Beguería *et al.* (2014) are given by;

$$P_r^m = \frac{(r-1)(r-2)\dots\dots(r-m)}{(n-1)(n-2)\dots\dots(n-m)} \tag{5}$$

Where, P_r^m is the probability weight, r is the rank assigned to the data series arranged in ascending order, n is the number of observations and m is the order.

L-moments ratios τ_3 and τ_4 given in equations (6) and (7) were used to identify the possible probability distribution function that is close to the theoretical distribution curves generated by Hosking and Wallis (2005). L-moment ratios (L-Skewness τ_3 and L-Kurtosis τ_4) were calculated as follows;

$$\tau_3 = \frac{\lambda_3}{\lambda_2} \quad (6)$$

$$\tau_4 = \frac{\lambda_4}{\lambda_2} \quad (7)$$

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

$$\lambda_2 = 2M_1 - M_0 \quad (8)$$

$$\lambda_3 = 6M_2 - 6M_1 + M_0 \quad (9)$$

$$\lambda_4 = 20M_3 - 30M_2 + 12M_1 - M_0 \quad (10)$$

The PWMs of order r are given by,

$$M_m = \frac{1}{N} \sum_{i=1}^N P_r^m \cdot X_{j,i}^K \quad (11)$$

From the L-moment ratio diagram, two candidate distributions emerged namely, the generalized extreme value (GEV) and the generalized logistic (GLO) functions. Since statistics oscillate around these curves, Vicente-Serrano *et al.* (2010) suggested that none of these can be rejected. However L-moments ratios of the D_j aggregated series fitted closer to the GLO, hence the latter was judged more appropriate for modelling the series. The GLO is given by following probability density function (Hosking and Wallis, 2005);

$$f(x) = \frac{\alpha^{-1} \exp[-(1-k)y]}{[1 + \exp(-y)]^2} \quad (12)$$

$$y = \begin{cases} -k^{-1} \log \left[1 - \frac{k(X_{j,i}^K - \xi)}{\alpha} \right], & k \neq 0 \\ \frac{(X_{j,i}^K - \xi)}{\alpha}, & k = 0 \end{cases} \quad (13)$$

The probability distribution function is given by;

$$F(x) = \frac{1}{[1 + \exp(-y)]} \quad (14)$$

Where ξ , α , and k are location, scale and shape parameters, respectively, for D ranging from:

$$X_{j,i}^K: \begin{cases} -\infty < X_{j,i}^K \leq \xi + \frac{\alpha}{k} & \text{if } k > 0 \\ -\infty < X_{j,i}^K < \infty & \text{if } k = 0 \end{cases} \quad (15)$$

The Parameters are computed as a function of L-moment ratios as follows;

$$k = -\tau_3 \tag{16}$$

$$\alpha = \frac{\lambda_2 \sin k\pi}{k\pi} \tag{17}$$

$$\xi = \lambda_1 - \alpha \left(\frac{1}{k} - \frac{\pi}{\sin k\pi} \right) \tag{18}$$

The F(x) values were then transformed to a normal variable by means of the following approximation by (Abramowitz and Stegun, 1964)

$$SPEI = W - \frac{C_0 + C_1 W + C_2 W^2}{1 + d_1 W + d_2 W^2 + d_3 W^3} \tag{19}$$

Where

$$W = \sqrt{\ln \left(\frac{1}{P^2} \right)} \tag{20}$$

Where P is the probability of exceeding a determined D_j value. The P value is obtained from $P = 1 - F(x)$. 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 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 $SPEI \leq -1$), was defined through its duration, level of dryness, magnitude and intensity.

2.4 Artificial neural network model

Artificial neural networks (ANNs) have been applied in hydrological modeling and prediction of nonlinear and randomized time series due to their flexibility (Diaconescu, 2008; Nasr and Zahran, 2014). Some of their merits as articulated by Mishra and Desai (2006) are; 1) The ease to recognize the relation between the input and output variables without explicit physical consideration, 2) The fact that the chaotic time series are easy to train regardless of any measurement errors, 3) The fact that they are easily adapted to solutions over time to compensate for varying conditions and 4) The fact that they possess other memory characteristics and once trained are easy to use.

The tested models are used in prediction of meteorological variables and other chaotic time series. In this study the nonlinear autoregressive with exogenous input (NARX) prediction neural network model is adopted.

2.4.1 NARX model Topology

The NARX model applied in prediction of time series uses a feedforward neural network with memory in form of tapped delay lines for the input variables propagated back at respective concurrent unit time delays and for the respective unit time predictions. For this, the network learns in the parallel-series (open loop) and makes a long term prediction in the parallel connection (closed loop) where the one step prediction ($y(n)$) is fed back into the network. NARX is a class of discrete-time nonlinear systems that can be mathematically represented as (Menezes and Barreto, 2008; Ardalani-Farsa and Zolfaghari, 2010)

$$y(n+1) = \{ [u_1(n), \dots, u_1(n-d+1); u_2(n), \dots, u_2(n-d+1); u_3(n), \dots, u_3(n-d+1); u_4(n), \dots, u_4(n-d+1)]; [y(n), \dots, y(n-d+1)] \} \tag{21}$$

Where d is the delay of the model at discrete time step n. $u_1(n)$, $u_2(n)$, $u_3(n)$ and $u_4(n)$ are predictors at the input regressor and $y(n)$ is the target output at the output regressor. The delay is usually selected based on the option giving the best prediction performance.

2.4.2 NARX model design, training and learning

In this study, the network uses four input variables (precipitation, evaporation, minimum and maximum temperatures) and thus four corresponding tapped delay lines for these inputs/ predictors are used at the input regressor as shown in Figure 2. One tapped delay line is available for the output regressor at each corresponding unit time delay. The number of neurons in the hidden layer chosen is selected by trial and error, although the final number decided upon should be closer to the one based on the procedure developed by Hecht-Nielsen (1987) applied in Maier and Dandy (2001) and Stathakis (2009). In their studies they stated that $N_H \leq 2u+1$ where N_H is the number of hidden layer neurons and u is the number of input variables. The number of hidden neurons used in this study is 9 embedded in one hidden layer. This feedforward backpropagation network uses a tan-sigmoid transfer function and a linear transfer function for the hidden and output layer neurons respectively. The network training function updates the weight and bias values according to Levenberg-Marquardt optimization function. It minimizes a combination of squared errors and weights, then determines the correct combination so as to produce a network which generalizes well (Demuth *et al.*, 2009). This training algorithm uses the regularization technique that involves modifying the performance function for reducing the number of parameters. Different training algorithms were used and the one adopted here (Levenberg-Marquardt) gave the best performances with the lowest Root Mean Squared Error (RMSE) and correlation coefficient (R). Through trial and error, the required number of tapped line delays and hidden layer neurons were then arrived at. During the training and learning period, input and target series are divided into two groups with the first group used during the training phase and the latter for prediction. The first group of data set includes historical data less the most recent data points of length equivalent to the prediction horizon. This arrangement allows the model to learn the input-output relationship without recorded data. The second group of input data set is equivalent to N used for simulation of the model and generating of new target series, whereas the actual target series is responsible for validation of the predicted values.

The typical performance function used in training is the Mean Square Error (MSE) presented as follows (Diaconescu, 2008):

$$MSE = \frac{1}{N} \sum_{i=1}^N e_i^2 = \frac{1}{N} \sum_{i=1}^N (t_i - y_i)^2 \quad (22)$$

Where t_i and y_i are the observed and predicted values respectively and N is the horizon of prediction. During the training and learning phase the network is running in a series-parallel connection and making a one step prediction.

2.4.3 NARX model prediction criteria

The performance of a trained network was measured by the errors on the training, validation, and test sets. In this study, the correlation coefficient R between the outputs and targets of neural network is used together with the autocorrelations. R was used as a measure of the variation between the generated outputs and targets.

During the prediction stage, the network runs in parallel mode for the long term predictions (Figure 2). The one step prediction at the output is fed back into the network to perform multistep predictions. As training continues, the y_i estimates become more similar to the actual values t_i of the time series, indicating convergence of the training process. These are presented in regression plots, autocorrelation plots and mean square error plots.

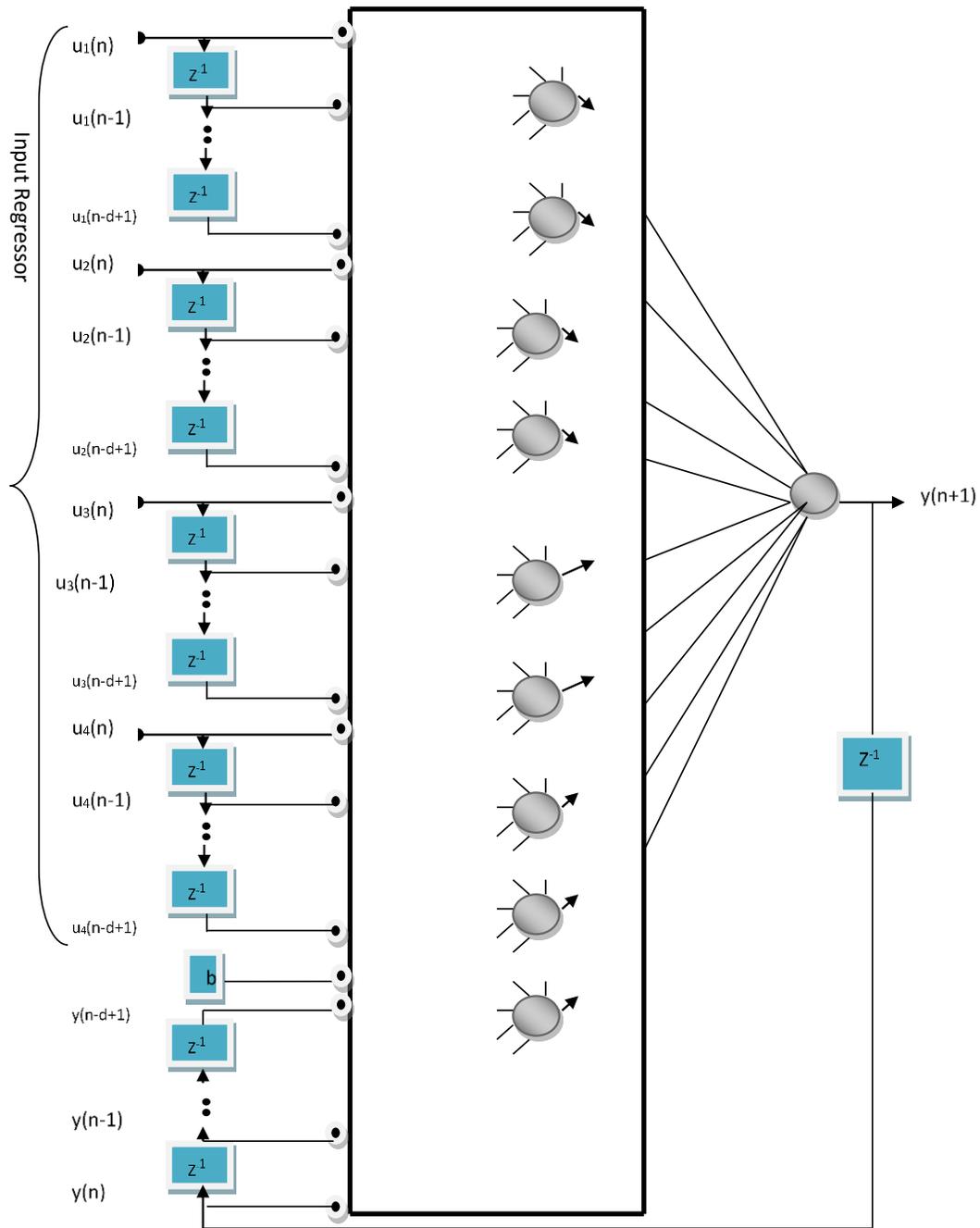


Figure 2. Optimum parallel NARX network prediction with Z^{-1} unit time delays

3 Results

3.1 Climatic variability

Results of analysis of rainfall characteristics are presented in form of annual rainfall, number of rainy days, rain onset day, cessation day and duration of the rain season (Table 1). The respective descriptive statistics are also presented. Inter annual variability in the rainfall characteristics are represented through the

coefficients of variation (CVs). The CVs for annual rainfall, number of rainy days, rain onset day, cessation day and duration of the rainy season are 42%, 17%, 29% and 28% respectively. Tuerkes (1996), Gocic and Trajkovic (2013) and Nsubuga *et al.* (2014) characterized stations with low rainfall reliability as those with coefficient of variation greater than 20%. This implies that there is a low chance of dependable rainfall at the Okavango Delta, making it prone to frequent droughts.

3.2 Rain Onset and cessation of rain analysis

Table 1 presents a long term analysis of onset and cessation of rain ranging from the earliest onset on the 27th September and the latest date of 7th January with an average on the 22nd November. The cessation dates range between 11th February and 14th May for the early and late cessation respectively with an average on the 25th March. The number of rainy days ranges from 23 to 85 exhibiting a long term mean of 44 days. The duration of the rainy season also ranges from 59 days to 169 days with an average of 113 days.

There was no rain onset date recorded in the years 1979/80, 1992/93 and 1994/95 (Table 1) with the probability of 6.8% being registered. Figure 3a reveals a near normal onset for all the years on record except for the three years where the onset criterion was not met. Figure 3b shows early cessation for 1970/71, 1982/83, 1991/92, 1993/95, 1996/97, 1997/98, 1998/99, 2010/11, 2011/12 and 2012/13. Late onset was reported in 1971/72, 1974/75, 1983/84, 1987/88, 2000/01, 2003/04, 2005/06 and 2009/10.

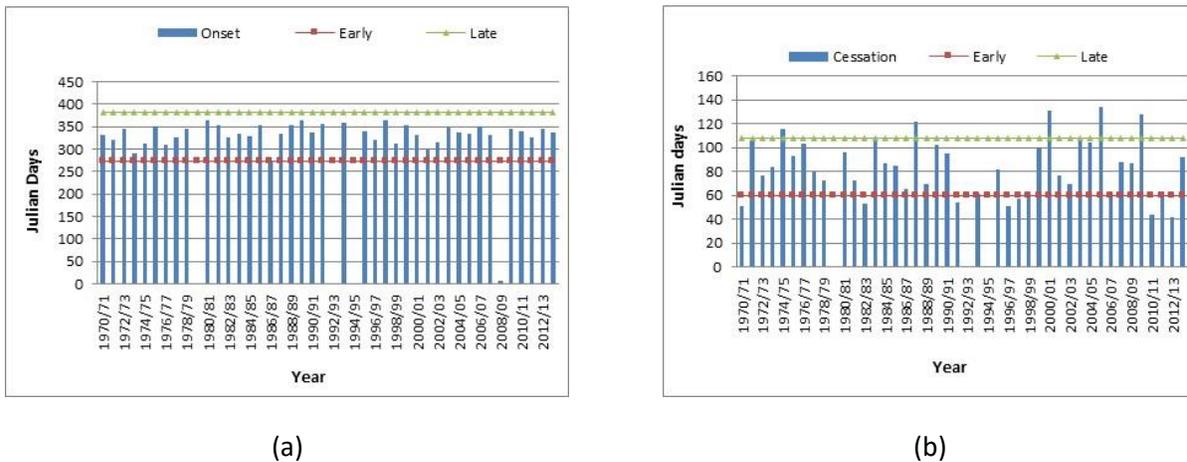


Figure 3. Analysis of (a) rain onset dates and b) rain cessation dates

3.3 Analysis of dryness categories

The analysis was expressed through the probability of occurrence of drought of a given threshold (classified in Tables 2 and 3). Probabilities of experiencing dryness/wetness were determined as a ratio of frequency of occurrence of the SPEI category to the total SPEIs on record for the study period. A drought event starts when the SPEIs are continuously negative and ends when SPEIs are positive. In this case drought can be presented in the form of severity and duration (Figure 4). Probabilities of experiencing dryness or wetness classified as extreme, severe, moderate and near normal conditions are presented in Table 2 for SPEI computations of one, three and twelve months. The probability of extreme dryness for all the three timescales is less than 1.0% while moderate dryness is encountered with 11.0%, 12.0% and 9.5% respectively. Near normal conditions are more frequent with 67.4%, 64.8% and 66.2% for one, three and twelve month's scales respectively. The moderately wet and moderately dry conditions have approximately equal chance of occurrence with probabilities of 9.5%, 11.4% and 12.6% for the three time scales respectively. Extreme wetness is also a rare phenomenon with only 1.9% probability of occurrence for SPEI-1 and SPEI-3 and 1.5% for the SPEI-12. The linear trend for SPEIs at the three time scales, show an increase in dryness of 0.06% for both the three and twelve months and of 0.03% for the one month period.

Table 1. Rain onset,cessation, number of rainy days and duration of the rainy season

Year	Annual Rainfall (mm)	Rainy Days	*Onset day	*Cessation day	Duration of the Rainy season (days)	Year	Annual Rainfall (mm)	Rainy Days	*Onset day	*Cessation day	Duration of the Rainy season (days)
1970/71	429.5	36	330	51	86	1993/94	533.1	38	359	61	67
1971/72	715.2	57	319	105	151	1994/95	150.9	24	F	F	F
1972/73	244.7	26	346	77	96	1995/96	515.7	40	339	82	108
1973/74	1190.4	85	289	84	160	1996/97	399.5	43	319	51	97
1974/75	660.2	64	312	116	169	1997/98	361.6	28	363	57	59
1975/76	378.3	49	349	93	109	1998/99	383.1	39	312	59	112
1976/77	539.5	55	310	103	158	1999/00	581.9	59	354	100	111
1977/78	734.0	68	325	80	120	2000/01	371.2	43	331	131	165
1978/79	290.2	30	344	72	93	2001/02	309.0	31	299	77	143
1979/80	508.9	36	F	F	F	2002/03	273.3	42	315	69	119
1980/81	497.7	49	363	96	98	2003/04	523.9	59	348	105	122
1981/82	218.0	34	353	72	84	2004/05	326.7	44	338	104	131
1982/83	391.3	48	326	53	92	2005/06	684.4	62	333	134	166
1983/84	355.8	40	333	108	140	2006/07	283.4	35	351	62	76
1984/85	302.2	35	329	87	123	2007/08	506.4	49	332	88	121
1985/86	386.1	41	352	85	98	2008/09	388.4	47	7	87	80
1986/87	295.0	38	270	65	160	2009/10	748.8	47	344	128	149
1987/88	355.3	42	335	122	152	2010/11	628.5	48	339	44	70
1988/89	645.0	55	352	69	82	2011/12	406.3	38	327	63	101
1989/90	292.1	34	365	102	102	2012/13	355.8	23	344	42	63
1990/91	498.6	41	337	95	123	2013/14	584.1	49	337	92	120
1991/92	275.8	35	355	54	64	Mean	451.4	44	326	84	113
1992/93	355.2	46	F	F	F	SD	187.4	12.3	55.0	24.4	31.9
						Cv	0.42	0.28	0.17	0.29	0.28

*Julian days were used in the computation of onset and cessation of rainfall

F=Onset Failure

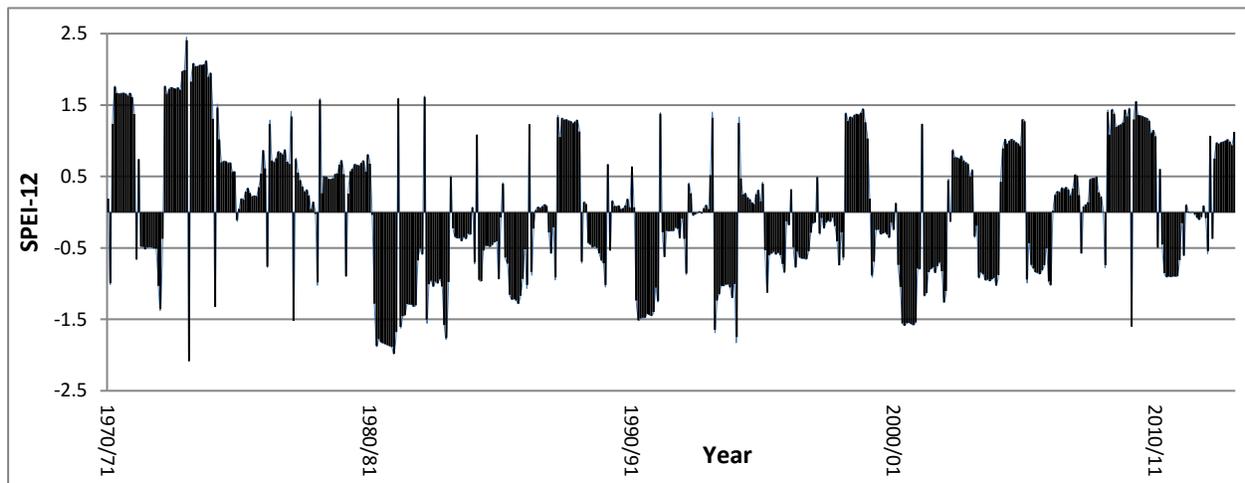
SD= Standard deviation

CV=Coefficient of variation

Table 2. Probabilities of experiencing various drought conditions

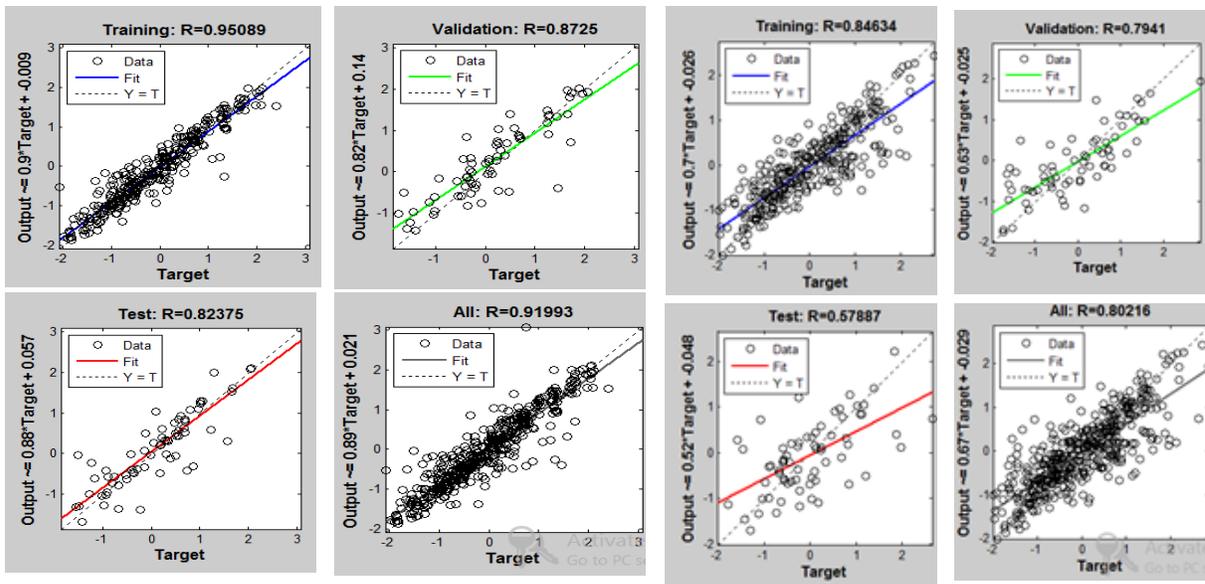
Class	Dryness Classification	SPEI-1		SPEI-3		SPEI-12	
		F	P	f	P	f	P
<-2	Extremely Dry	3	0.006	0	0.000	1	0.002
-1.99 to -1.50	Severe Dryness	19	0.036	27	0.051	26	0.050
-1.49 to 1.00	Moderate Dryness	58	0.110	63	0.120	49	0.095
-1.00 to 1.00	Near Normal	356	0.674	341	0.648	342	0.662
1.00 to 1.49	Moderate Wetness	50	0.095	60	0.114	65	0.126
1.50 to 1.99	Severe Wetness	32	0.061	25	0.048	26	0.050
> 2.00	Extreme wetness	10	0.019	10	0.019	8	0.015

The plot of SPEI-12 in Figure 4 shows drying and wetting patterns clustered together. It indicates the period between 1972/73 and 1979/80 as a wet period rising to a maximum index of 2.4 in 1973/74. The period from 1980/81 to 1987/88 presents itself as a dry spell registering an index of -2.0 in 1981/82. A short wet spell merged between 1987/88 to 1988/89 with conditions of moderate wetness and thereafter, the dry spell continued until 1998/99 but less severe than the one registered between 1972/73 and 1979/80. The Period from 1998/99 to 1999/00 was another moderate wet spell. Within the period from 1999/00 to 2002/03 a moderately dry period was registered with near normal conditions continuing until 2008/09. A moderately wet spell occurred between 2008/09 and 2010/11 with the remainder of the period registering near normal conditions.

**Figure 4.** Progress of twelve months SPEI in the study area from 1970/71 to 2013/14

3.4 Dryness prediction using artificial neural networks

The five years prediction results were validated using historical data from the past five years. The coefficient of correlation (R) between the predicted values and the observed values was 92%, 80% and 43% for SPEI-12, SPEI-3 and SPEI-1 respectively shown in Figures 5a, 5b and 6. The data set was split into two series with 1970/71 to 2008/09 series in training, testing and validation of the one step predictions in the series-parallel network. Data series from 2009/10 to 2013/14 was used for fitting of the multistep prediction in the parallel network mode.



(a) (b)

Figure 5. Regression plots for (a) SPEI-12 and (b) SPEI-3

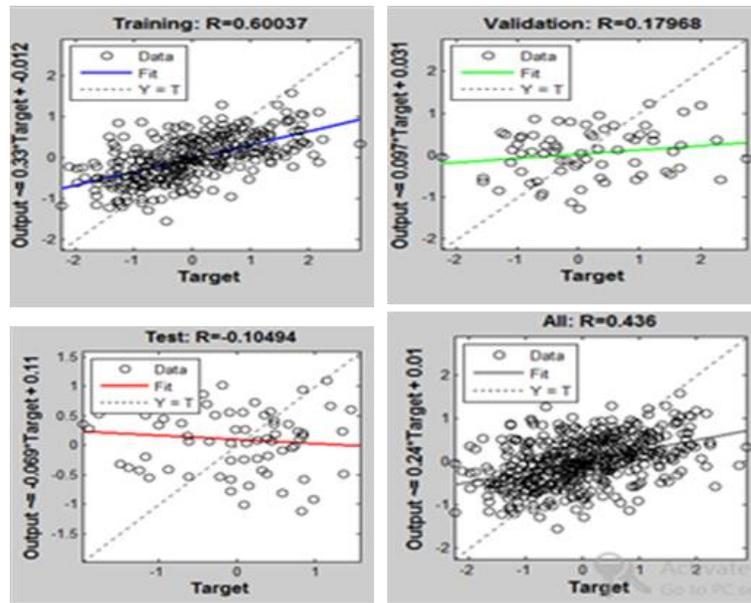
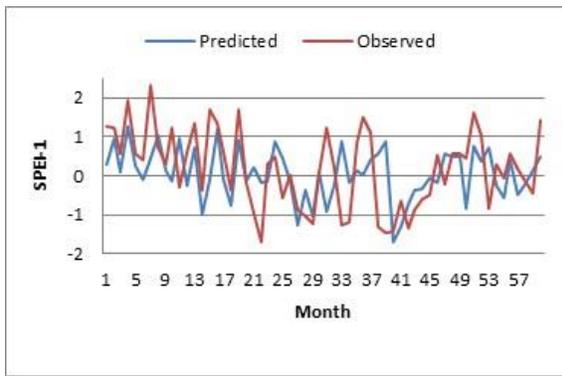


Figure 6. Regression plots for SPEI-1

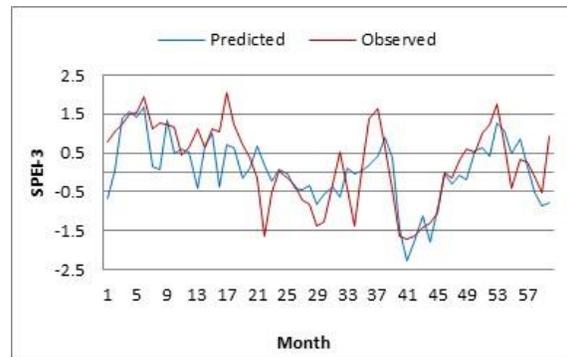
The near normal conditions remained predominant in the period of prediction. Table 3 presents comparisons in frequency analysis for the three dry periods. Near normal conditions for SPEI-12 were noticed with probability of 60% in the period of prediction, compared to 55% in the historical period. At SPEI-3, these were noticed with probability of 80% in the period of prediction, compared to 52% in the historical period. For SPEI-1, almost perfect near normal conditions were noticed with probability of 87% as compared to 60% in the historical period. Figures 7a, 7b and 8 present the fit between five year prediction and historical data covering the period of five years.

Table 3. Probabilities of experiencing various drought conditions for observed (obs) and predicted (Pred) values

Class	Classification	SPEI-1		SPEI-3		SPEI-12	
		P-Obs	P-Pred	P-Obs	P-Pred	P-Obs	P-Pred
<-2	Extremely Dry	0.000	0.000	0.000	0.017	0.000	0.000
-1.99 to -1.50	Severe Dryness	0.017	0.017	0.067	0.033	0.017	0.000
-1.49 to 1.00	Moderate Dryness	0.133	0.067	0.100	0.033	0.000	0.000
-1.00 to 1.00	Near Normal	0.600	0.867	0.517	0.800	0.550	0.600
1.00 to 1.49	Moderate Wetness	0.150	0.050	0.217	0.083	0.417	0.317
1.50 to 1.99	Severe Wetness	0.083	0.000	0.083	0.033	0.017	0.083
> 2.00	Extreme wetness	0.017	0.000	0.017	0.000	0.000	0.000



(a)



(b)

Figure 7. Five years predictions of monthly (a) SPEI-1 and (b) SPEI-3



Figure 8. Five years prediction of monthly SPEI-12

4 Discussion

4.1 Rainfall variability

Rainfall over Botswana is known to vary with a declining trend (Batisani and Yarnal, 2010; Moalafhi *et al.*, 2012). This study found great rainfall variability at the Okavango Delta of CV more than 20% in rainfall amount,

number of rainy days, cessation of rain and duration of the rainy season. From the analysis of rainfall and PET, all the months presented with a high water deficit. The annual PET exceeds the annual rainfall more than four times corroborating the earlier findings that more than 90% of the moisture in the Okavango Delta is lost through evaporation (Farquharson *et al.*, 1990). The results show that annual rainfall totals are well described by the number of rainy days with both rainfall amounts and rainy days showing a declining trend. This finding is consistent with those from Batisani and Yarnal (2010) in a similar study in Botswana. Intervention analysis performed on rainfall amounts time series resulted in 1979/80 being the year at which gradual decrease in rainfall totals starts to be noticeable (Parida and Moalafhi, 2008) over Botswana. The major metrological factors contributing to this changing regime are, however, not well explained. Studies by Nicholson *et al.* (2001) attributed climatic variability to the influence of the Southern Oscillation index (SOI) and of Sea Surface Temperatures (SSTs) of the Atlantic, Pacific, and Indian Oceans during El Niño Southern Oscillation (ENSO) episodes. However similar studies by Manatsa *et al.* (2008) found no association with rainfall variability based on these factors.

4.2 Rain Onset, cessation and length of rain season

To fully describe and understand the climate, regarding the nature and occurrence of drought in Okavango, characteristics of the rainy seasons in terms of the onset and cessation of rain and the duration of the rainy season were assessed. The onset of rain over the study area was less variable than the cessation with respective proportion at 17% and 30%. The higher variability in cessation results in variations in the length of the rainy season. This significantly affects stream flow which in turn affects the aquatic biodiversity (Bunn and Arthington, 2002). The major role of the wet season is to inundate wetlands and recharge ground water (Acharya, 2000). A short rainy season will have a negative impact on the aquatic ecosystems. In this regard, the early cessation is responsible for short rains rather than rainy days or rain onset which was also found true in other semi-arid studies in Ethiopia (Araya and Stroosnijder, 2011).

Determination of the number of years in which the onset criterion is not realized is a significant risk index. In Okavango Delta, the chances of onset failure are still low compared to 10% reported in South Africa's Free state by Moeletsi and Walker (2012). This implies that the onset criterion was not leading to inadequate rain to define the start of a rainy season. Failure of onset between November and January leads to total failure associated with a low number of rainy days and high drought severity. During the years in which the onset criteria was not met in the 44 years of analysis, particular interest is given to 1979/80 which is the year reported for intervention in earlier studies.

4.3 Dryness severity in Okavango

For the Okavango ecosystems, the SPEIs for three and twelve months are of great interest since they describe seasonal dryness patterns. The results have established that near normal conditions are the most predominant in the Okavango Delta followed by moderate droughts for all the three drought categories under investigation. The frequency of severe droughts is very low. This finding is in agreement with Batisani (2011) who found out similar characteristics over Botswana while using the standardized precipitation index (SPI). The agreement in the results could be attributed to the increasing dryness resulting from decreasing rainfall rather than raising temperatures over the study area. The low possibilities of severe and extreme droughts signify a low vulnerability for the period under consideration. However the risk of extreme droughts cannot be ruled out in this Delta due to the high variability in rainfall of 42%. From the findings of Romero *et al.* (1998) and Modarres and de Paulo Rodrigues da Silva (2007), rain variability is of common occurrence in semi-arid regions leading to inter annual, and seasonal variability. This variability results in difficulties in developing coping mechanisms to reduce the tendency towards drought vulnerability. This assertion has also been confirmed by studies on climate change conducted in Botswana (Batisani, 2011; 2012; Ministry of Environment Wildlife and Tourism of the Republic of Botswana, 2012). The results have confirmed gradually increasing dryness for the one, three and twelve months though not substantial; these findings indicate that

the risk of extreme dryness still exists due to the aridity of the Delta. The multistep ahead NARX artificial neural network model shows its ability to predict dryness conditions. The accuracy of the model predicting capacity was highest for SPEI-12 with a coefficient R of 92% and lowest for SPEI-1 with a coefficient R of 43%. The same trend was registered by Mishra and Desai (2006) and Illeperuma and Sonnadara, (2009) in their studies of drought prediction using SPI. They recorded R values of up to 20% for the 6 month prediction. The coefficient R of 43% was reasonable since it was found to be statistically significant at both 5% and 1% levels. In this regard, the 60 months SPEI-1 prediction is realistic. The near normal conditions are predominant in the next five years prediction, but more pronounced with the SPEI-1.

5. Conclusions

Okavango Delta, being an important world heritage site, requires an investigation of the long term trends in rain onset, cessation and number of rainy days and their impact on the dryness severity at the Delta. The results confirm a highly variable rainfall regime of 42% with decreasing number of rainy days and high variability in cessation and duration of the rainy season. This makes Okavango more vulnerable to recurrent dryness conditions. The study also demonstrates a high association of number of rainy days, rain onset, cessation of rain and length of the rainy season. The changing length of the rainy season and rainy days is highly attributed to the increasing trends in dryness severity. Increased dryness is an indication of higher climatic variability for the future. While the incidences of severe to extreme drought are rare, the study was able to show a general increase in dryness severity for the period of record. Simulations undertaken using NARX multistep ahead artificial neural network showed a near normal situation most of the time and no major changes in the next five years. However, vulnerability to more severe drought is not completely ruled out because of the high variability in rainfall and because of the location of the Delta in a semi-arid zone.

Acknowledgements

This study has been funded with assistance from Mobility for Engineering Graduates in Africa (METEGA). The data used was provided by Department of Meteorological services (DMS) of Botswana. The authors gratefully appreciate the comments from the reviewers that enriched the manuscript.

References

- Abramowitz M. and Stegun I.A. (1964). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*: Courier Corporation.
- Acharya G. (2000), Approaches to valuing the hidden hydrological services of wetland ecosystems, *Ecological economics*, **35**(1), 63-74.
- Acreman M., Blake J., Booker D., Harding R., Reynard N., Mountford J. and Stratford C. (2009), A simple framework for evaluating regional wetland ecohydrological response to climate change with case studies from Great Britain, *Ecohydrology*, **2**(1), 1-17.
- Araya A. and Stroosnijder L. (2011), Assessing drought risk and irrigation need in northern Ethiopia, *Agricultural and Forest meteorology*, **151**(4), 425-436.
- Ardalani-Farsa M. and Zolfaghari S. (2010), Chaotic time series prediction with residual analysis method using hybrid Elman–NARX neural networks, *Neurocomputing*, **73**(13), 2540-2553.
- Batisani N. (2011), The spatio-temporal-severity dynamics of drought in Botswana, *Journal of Environmental Protection*, **2**(06), 803.
- Batisani N. (2012), Climate variability, yield instability and global recession: the multi-stressor to food security in Botswana, *Climate and Development*, **4**(2), 129-140.

- Batisani N. and Yarnal B. (2010), Rainfall variability and trends in semi-arid Botswana: implications for climate change adaptation policy, *Applied Geography*, **30**(4), 483-489. doi: 10.1016/j.apgeog.2009.10.007
- Bauer P., Gumbricht T. and Kinzelbach W. (2006), A regional coupled surface water/groundwater model of the Okavango Delta, Botswana, *Water Resources Research*, **42**(4).
- Beguería S., Vicente-Serrano S.M., Reig F. and Latorre B. (2014), Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International journal of climatology*, **34**(10), 3001-3023.
- Box G.E. and Jenkins G.M. (1976), *Time series analysis: forecasting and control, revised ed*: Holden-Day.
- Bunn S.E. and Arthington A.H. (2002), Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity, *Environmental management*, **30**(4), 492-507.
- Burkett V. and Kusler J. (2000). Climate change: Potential impacts and interactions in wetlands of the United States: Wiley Online Library.
- Dastorani M. and Afkhami H. (2011), Application of artificial neural networks on drought prediction in Yazd (Central Iran), *Desert*, **16**(1), 39-48.
- De Silva R., Dayawansa N. and Ratnasiri M. (2007), A comparison of methods used in estimating missing rainfall data, *Journal of Agricultural Science*, **3**, 101-108.
- Demuth H., Beale M. and Hagan M. (2009). Neural network toolbox user's guide, The MathWorks. Inc., Natick, USA.
- Diaconescu E. (2008), The use of NARX neural networks to predict chaotic time series, *WSEAS Transactions on Computer Research*, **3**(3), 182-191.
- Droogers P. and Allen R.G. (2002), Estimating reference evapotranspiration under inaccurate data conditions, *Irrigation and drainage systems*, **16**(1), 33-45.
- Farquharson F., Gash J. and Roberts J. (1990), Report on a review of evaporation estimation for the Okavango Delta.
- Gao Y. and Meng Joo E. (2005), NARMAX time series model prediction: feedforward and recurrent fuzzy neural network approaches, *Fuzzy sets and systems*, **150**(2), 331-350.
- 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.
- Haktanira T. and Bozduman A. (1995), A study on sensitivity of the probability-weighted moments method on the choice of the plotting position formula, *Journal of Hydrology*, **168**(1), 265-281.
- Hecht-Nielsen R. (1987). *Kolmogorov's mapping neural network existence theorem*. Paper presented at the Proceedings of the international conference on Neural Networks.
- Hosking J.R.M. and Wallis J.R. (2005), *Regional frequency analysis: an approach based on L-moments*: Cambridge University Press.
- Hosking J.R.M., Wallis J.R. and Wood E.F. (1985), Estimation of the generalized extreme-value distribution by the method of probability-weighted moments, *Technometrics*, **27**(3), 251-261.
- Illeperuma G. and Sonnadara U. (2009), Forecasting Droughts using Artificial Neural Networks. *Promoting Knowledge Transfer to Strengthen Disaster Risk Reduction & Climate Change Adaptation*, 100.
- Kysely J. (2002), Comparison of extremes in GCM-simulated, downscaled and observed central-European temperature series, *Climate Research*, **20**(3), 211-222.
- 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.
- Machiwal D. and Madan Kumar J. (2012). *Hydrologic time series analysis: Theory and practice*: Springer Science & Business Media.
- Maier H.R. and Dandy G.C. (2000), Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications, *Environmental modelling & software*, **15**(1), 101-124.

- Maier H.R. and Dandy G.C. (2001), Neural network based modelling of environmental variables: a systematic approach, *Mathematical and Computer Modelling*, **33**(6), 669-682.
- Manatsa D., Chingombe W. and Matarira C. (2008), The impact of the positive Indian Ocean dipole on Zimbabwe droughts, *International Journal of Climatology*, **28**(15), 2011-2029.
- McCarthy T. (2006), Groundwater in the wetlands of the Okavango Delta, Botswana, and its contribution to the structure and function of the ecosystem, *Journal of Hydrology*, **320**(3), 264-282.
- McEvoy D.J., Huntington J.L., Abatzoglou J.T. and Edwards L.M. (2012), An evaluation of multiscale drought indices in Nevada and Eastern California, *Earth Interactions*, **16**(18), 1-18.
- 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.
- Menezes J.M.P. and Barreto G.A. (2008), Long-term time series prediction with the NARX network: an empirical evaluation, *Neurocomputing*, **71**(16), 3335-3343.
- Ministry of Environment Wildlife and Tourism of the Republic of Botswana. (2012). *Second National Communication to the United Nations Framework Convention on Climate Change (UNFCCC)*. Gaborone, Botswana: Republic of Botswana.
- Mishra A.K. and Singh V.P. (2011), Drought modeling—A review, *Journal of hydrology*, **403**(1), 157-175.
- Mishra K.A. and Desai V. (2006), Drought forecasting using feed-forward recursive neural network, *Ecological Modelling*, **198**(1), 127-138.
- Mkhabela M., Bullock P., Gervais M., Finlay G. and Sapirstein H. (2010), Assessing indicators of agricultural drought impacts on spring wheat yield and quality on the Canadian prairies, *Agricultural and Forest meteorology*, **150**(3), 399-410.
- Moalafhi D.B., Parida B. and Kenabatho P. (2014), A hybrid stochastic-ANN approach for flow partitioning in the Okavango Delta of Botswana, *Global Nest Journal*, **16**(1), 68-79.
- Moalafhi D.B., Tshoko R., Athlhopeng J.R., Odirile P.T. and Masike S. (2012), Implications of climate change on water resources of Botswana., *Advance Journal of Physical Sciences*, **1**(2), 4-13
- Modarres R. and de Paulo Rodrigues da Silva V. (2007), Rainfall trends in arid and semi-arid regions of Iran, *Journal of Arid Environments*, **70**(2), 344-355.
- Moeletsi M. and Walker S. (2012), Rainy season characteristics of the Free State Province of South Africa with reference to rain-fed maize production, *Water SA*, **38**(5), 775-782.
- Nasr M. and Zahran H.F. (2014), Using of pH as a tool to predict salinity of groundwater for irrigation purpose using artificial neural network, *The Egyptian Journal of Aquatic Research*, **40**(2), 111-115.
- Ngetich K., Mucheru-Muna M., Mugwe J., Shisanya C., Diels J. and Mugendi D. (2014), Length of growing season, rainfall temporal distribution, onset and cessation dates in the Kenyan highlands, *Agricultural and Forest meteorology*, **188**, 24-32.
- Nicholson S.E., Leposo D. and Grist J. (2001), The relationship between El Niño and drought over Botswana, *Journal of Climate*, **14**(3), 323-335.
- Nsubuga F.W., Botai O., Olwoch J.M., Rautenbach C.D., Bevis Y. and Adetunji A.O. (2014), The nature of rainfall in the main drainage sub-basins of Uganda, *Hydrological Sciences Journal*, **59**(2), 278-299. doi: 10.1080/02626667.2013.804188
- Pachauri R. and Reisinger A. (2007), IPCC fourth assessment report. *IPCC, Geneva*.
- Palmer W.C. (1965), *Meteorological drought*: US Department of Commerce, Weather Bureau Washington, DC, USA.
- 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.
- Potop V. and Možný M. (2011), The application a new drought index—Standardized precipitation evapotranspiration index in the Czech Republic, *Mikroklima a mezoklima krajinných struktur a antropogenních prostředí*, **2**.

- Romero R., Guijarro J., Ramis C. and Alonso S. (1998), A 30-year (1964-1993) daily rainfall data base for the Spanish Mediterranean regions: First exploratory study, *International Journal of climatology*, **18**(5), 541-560.
- Singh C. (1999), Principal components of monsoon rainfall in normal, flood and drought years over India, *International Journal of climatology*, **19**(6), 639-652.
- Sönmez F.K., Koemuescue A.U., Erkan A. and Turgu E. (2005), An analysis of spatial and temporal dimension of drought vulnerability in Turkey using the standardized precipitation index, *Natural Hazards*, **35**(2), 243-264.
- Stathakis D. (2009), How many hidden layers and nodes?, *International Journal of Remote Sensing*, **30**(8), 2133-2147.
- Stokes S., Thomas D.S. and Washington R. (1997), Multiple episodes of aridity in southern Africa since the last interglacial period, *Nature*, **388**(6638), 154-158.
- Tabari H., Somee B.S. and Zadeh M.R. (2011a), Testing for long-term trends in climatic variables in Iran, *Atmospheric Research*, **100**(1), 132-140.
- Tsheko R. (2004), Rainfall reliability, drought and flood vulnerability in Botswana, *Water SA*, **29**(4), 389-392.
- Tuerkes M. (1996), Spatial and temporal analysis of annual rainfall variations in Turkey, *International Journal of climatology*, **16**(9), 1057-1076. doi: 10.1002/(SICI)1097-0088
- Vicente-Serrano S.M., Beguería S. and López-Moreno J.I. (2010), A multiscale drought index sensitive to global warming: the standardized precipitation evapotranspiration index, *Journal of Climate*, **23**(7), 1696-1718.
- 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.
- 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.