Stochastic Characteristics of Rainfall and River Flow of Kainji Reservoir System

Abstract

The analysis of time series is essential to detect intrinsic stochastic characteristics of hydrologic variables, which are cardinal in understanding the reliability and uncertainty associated with water supply yields derived from surface water reservoirs which are central for planning purposes. To this end analysis of stochastic characteristics like moments and dependence structures of the data series were done to be able to evaluate randomness and trend pattern of the time series. Analytical tools employed were moment equations; Autocorrelation for serial dependency, trend, and Hurst Phenomenon. Analysis for serial dependency was done employing Durbin Watson standard test on the rainfall of the basin and inflow. The lag 1 autocorrelation value obtained was 0.19. The DW value computed was 1.42 for seasonal rainfall which indicates that there is no serial autocorrelation in the seasonal rainfall time series. Also the lag 1 autocorrelation value (0.667) obtained and DW value (0.67) for the seasonal inflow depicted positive autocorrelation. Analysis of long term dependency on the inflows employing Modified Rescaled Range Statistic (R/S) using Lo’s (1991) q opt indicates long term dependency of the inflow time series with computed QN, q value of 3161.6. Mann-Kendall and step trend test were employed for the trend analysis. The trend test analysis for the rainfall, draw-down and inflow time series showed that the rainfall time series exhibit positive trend with Z value of 1.98 at 0.05 level of significance and Sen’s slope Q of 3.236. Also the z value was greater than 1.96 i.e. the chosen significant level of 5%, with Z 0.025= ±1.96, indicating the null hypothesis that trend does not exist is void. The analysis of the step or jump trend of the rainfall draw-down showed that the critical value 1.05 (18) = 1.734, 1.446 so that the hypothesis H1 is accepted, as 1 (μ1≠ μ2), meaning that the step or jump trend is significant and positive at 5% probability. From the annual Mann-Kendall trend analysis of the flow, the Z statistic was determined to be -0.23, the Sen’s slope Q was determined to be-0.012 indicating the alternative hypothesis that trend does exist is void.

Introduction

Climate change is a long-term change in the statistical distribution of weather patterns over periods of time that range from decades to millions of years. Climate change may be limited to a specific region of weather patterns over periods of time that range from decades to millions of years. Climate change constitutes another factor for flow regime alteration and will interact with other anthropogenic flow modifications. Climate change "is now evident from observations of increases in global average air and ocean temperatures, widespread melting of snow and ice and rising global average sea level" [2]. Higher temperatures could potentially increase evaporation rates at surfaces and transpiration by plants, which lead to a reduction in runoff [3]. Additionally, in snow or glacier-affected river basins, runoff is altered by a decline in melt water [4]. Regionally and seasonally differing developments are simulated for precipitation amounts and patterns which will cause, depending on the location and season, higher or lower runoff values in the future [1,5]. Moreover, it is expected that climate change will accelerate the hydrological cycle with an increasing intensity of rainfalls and frequency of extreme weather events [6]. All these implications will interact in different ways at different climatic locations inducing substantial alterations in the river flow regimes with large geographical disparities in directions and causes. The consequences of these alterations are manifold. In 2000, approximately 2.4 billion people lived in water stressed river basins and this number is supposed to rise in the future [7]. An increase in water stress can be caused by a reduction in total flow amounts or just by changes in flow seasonality.

In the last century, natural flow regimes have been heavily modified by different anthropogenic impacts [8]. Worldwide around 50 000 large dams and an estimated number of 800 000 smaller dams exist, usually generating a less variable flow with elevated low flows and dampened flood peaks [9]. Water demands of an exponential growing world population leads to reduced river discharge due to withdrawals for irrigation, electricity production, manufacturing, domestic purposes and others. In addition, population growth and society’s increasing demands on resources have caused immense land-use changes. Urbanization and deforestation result in large sealed areas that alter flow magnitudes and timing through lower evapo-transpiration rates and faster runoff [10]. Many rivers have also been artificially modified by construction works such as channelization, embanking, straightening, widening or deepening with further impacts on flow and flow velocity.

Climate is changing in term of rising global mean temperature, rise level of sea and variability in precipitation [11]. Under climate conditions and

Keywords: Trend analysis; Rainfall; Reservoir system; Climate change
change conditions, many water systems are projected to be less reliable and more vulnerable in meeting user demands, exacerbating existing competition for water resources. In general, negative effects of climate change on water resource systems will complicate changing economic activity, water quality, increasing population, land use change and urbanization [12,13]. A reduction of available water resources is expected in regions where runoff is projected to decrease; conversely, where rainfall increases are expected, increased water supply is projected [14]. However, the benefit of this might be reduced by negative effects of higher variability of precipitation and seasonal runoff in water supply, food risks, and water quality. Identification of temporal changes in hydrological regimes of river basins is an important topic in contemporary hydrology because of the potential impacts of climate change on river basin. For this purpose, generally parametric and nonparametric techniques have been employed; the latter have been widely used mainly because of a fewer number of assumptions involved in their implementation [15]. In any sign investigation study, mostly it is vital to estimate hydrological trends at local scales, i.e. point estimates [16]. Precipitation is one of the most important meteorological variables which can impact the occurrence of drought or floods [17]. Analysis of precipitation and drought data offers important information which can be applied to improve water management strategies, to protect the environment, to plan agricultural production or in general to impact on economic development of a certain region. In recent years, a plethora of scientists have compared and analyzed the precipitation trends in the worldwide [18,19].

The analysis of time series is essential for building mathematical models to generate synthetic hydrologic records, to forecast hydrologic events, to detect intrinsic hydrologic characteristics of hydrologic variables as well to fill missing and extend records [20]. The assessment of the dynamics and regime of a particular hydrologic phenomenon is imperative; especially the time-based characteristics. Time-based characteristics of hydrological data are of great significance in the planning, designing and operation of water systems. This significance is informed more largely due to the variability and oscillatory behavior of hydrological sequences. Against this backdrop therefore, as noted by the lack of complete understanding of the physical processes involved and the consequent uncertainties in the magnitudes and frequencies of future events highlight the importance of time series analysis [21]. Though generally, hydrologic processes such as precipitation and runoff evolve on a continuous time scale and their estimation correspondingly unduly difficult, in particular, rainfall modeling and its quantitative estimation or forecasting are important considering the fact that it is a critical weather parameter in the estimation of crop water requirement, and development of long lead time flood and flash-flood warning systems [22].

Material and Methods

The study area

Geographically, Kainji hydroelectric dam is located in New Bussa town now headquarters of Borgu local government area of Niger State, Nigeria. The lake created behind the dam span between latitude 9° 8’ to 10° 2” and between longitude 4° 5’ to 4° 7’ E with reference point 9.54 N and 4.38 E northwest of the Federal Capital Territory (FCT, Abuja) [23].

Hydrology of the niger river system

The average rainfall at the headwaters of Niarand and Milo rivers at the source of the Niger at the Fouta Djallon Mountains in Guinea and its exit to the sea in Nigeria is 2200 mm. The river flow regime is characterized by two distinct flood periods occurring annually namely the White and Black floods. The black flood derives its flow from the tributaries of the Niger outside Nigeria (flow lag October to May) and arrives at Kainji reservoir (Nigeria) in November and lasts until March at Jebba after attaining a peak rate of about 2,000 m$^3$/sec in February [24]. The White flood is a consequent of flows from local tributaries especially the Sokoto-Rima and Malendo river systems. The White flood is heavily laden with silts and other suspended particles (flow lag June to September) and arrives Kainji in August in the pre-Kainji Dam River Niger having attained a peak rate of 4,000 to 6,000 m$^3$/sec in September-October in Jebba [25]. The critical low flow period into the Kainji reservoir is March and July each year.

Data collection: The reservoir daily inflow data for twenty five years (1990-2014), draw-down for nineteen years (1995-2014) and monthly average rainfall spanning fifty years (1965-2014) were collected from the hydrological unit of Kainji Hydropower station in Nigeria [24].

(a) Analysis of variability in rainfall: Rainfall analysis of the study area was done to show inter- annual decadal variation in the rain fall series. To this end average monthly data for 50 years were grouped into inter- annual- decadal time sequence. The time series were then examined to establish whether it does exhibit intermittency or otherwise as well as seasonal characteristics like trend and moments. To achieve this, the equations of measure of central tendency, coefficient of variation and standard deviation were employed as stated below:

$$
\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i
$$

(1)

Where $\bar{x}$ is the measure of central tendency $n$ = the size of the sample, $x$ is the observed parameter, $\Sigma$ is the summation symbol and $N$ is the number of observations.

$$
C_v = \frac{\sigma_p}{\mu}
$$

(2)

Where $C_v$ = coefficient of variation, $n\sigma_p$ is the standard deviation and $\mu$ is the mean.

$$
\sigma_p = \frac{\sqrt{\sum (x-\mu)^2}}{n-1}
$$

(3)

$\sigma_p$ = standard deviation, $x$ = is the observed parameter, $\mu$ = mean, $n$ = the size of the sample

(b) Trend analysis: MAKESENS was use in estimating trends in the time series of the average annual and monthly values of the rain fall, the reservoir level or draw-down and the river flows [25]. The data series was broken into wet and dry season and twelve calendar months. The procedure is based on the nonparametric Mann-Kendall test for the trend and the nonparametric Sen’s method for the
magnitude of the trend. MAKESENS was used to perform two types of statistical analyses. First the presence of a monotonic increasing or decreasing trend was tested with the nonparametric Mann-Kendall test and secondly the slope of a linear trend was estimated with the nonparametric Sen’s method. The null hypothesis of no trend Ho (the data is independent and randomly ordered) was tested, i.e. the observations (xi) are randomly ordered in time, against the alternative hypothesis, H1, where there is an increasing or decreasing monotonic trend. Based on the data series collected, the Z statistic was computed. The absolute value of Z was compared to the standard normal cumulative distribution to define if there is a trend or not at the selected 5% a level of significance. A positive (negative) value of Z indicates an upward (downward) trend. The equations employed in the computation of test statistic (S) and the statistics (Z) are as below:

\[ S = \sum_{i=1}^{n} \sum_{j=i+1}^{n} sgn(X_j - X_i) \]  

(4)

Where \( X_i \) and \( X_j \) are annual value in years \( j, k \), respectively, and:

\[ sgn(X_j - X_i) = \begin{cases} 1 & \text{if } X_j - X_i > 0 \\ 0 & \text{if } X_j - X_i = 0 \\ -1 & \text{if } X_j - X_i < 0 \end{cases} \]  

(5)

\[ Var(S) = \left[ n(n-1)(2n+5) - \sum_{j=1}^{n} (t_j - 1)(2t_j + 5) \right]/18 \]  

(6)

Where \( q \) is the number of tied groups and \( t_j \) is the number of data values in the \( p \)th group. The values of \( S \) and \( \text{VAR}(S) \) were employed to compute the test statistic \( Z \) as below:

\[ Z = \frac{S - n(n-1)/2}{\sqrt{\text{VAR}(S)}} \]

\[ = \begin{cases} Z > \text{0.5} & \text{if } S > 0 \\ Z < \text{0.5} & \text{if } S < 0 \end{cases} \]  

(7)

To determine the trend in the draw-down time series, the step or jump trend was considered, it can be expressed as:

\[ h_t = \mu_t + (\mu_{t+1} - \mu_t)_{j \rightarrow k} + \varepsilon \]

(8)

Where \( \mu_t \) and \( \mu_{t+1} \) are the mean process values for the first and second periods of the time series respectively, \( \varepsilon \) is an independent random variable with mean zero and variance \( \sigma^2 \), and

\[ (\mu_{t+1} - \mu_t)_{j \rightarrow k} = 0 \text{ for } \leq n_i \]

(9)

\[ (\mu_{t+1} - \mu_t)_{j \rightarrow k} = \mu_{t+1} - \mu_t > n_i \]

(10)

\( h_t \) and \( \mu_t \) are the number of observations in the first and second periods of the time series respectively. \( \mu_t, h_t, \) and \( \sigma^2 \) was estimated with their sample estimators, respectively as:

\[ h_t = \frac{1}{n_i} \sum_{i=1}^{n_i} h_i \]

(11)

\[ h_t = \frac{1}{n_k} \sum_{k=1}^{n_k} h_k \]

(12)

\[ S^2 = \frac{1}{n_i + n_k - 2} \sum_{i=1}^{n_i} (h_i - h_i)^2 + \sum_{k=1}^{n_k} (h_k + h_k)^2 \]

(13)

The significance of the step or jump trend was detected by performing a hypothesis as:

\[ H_0: \mu_t = \mu_t \]

or

\[ H_1: \mu_t \neq \mu_t \]

A t statistic was calculated as:

\[ t = \frac{|h_i - h_j|}{\sqrt{2a(n-2)}} \]

(16)

Which follows a student t distribution with a degree of freedom on \( n-2 \) with a significance level of \( a \) (5 %), the critical value \( t_{(n-2)} \) was determined from the student table. If \( t > t_{(n-2)} \), \( H_1 \) hypothesis is accepted, and the trend is significant.

(c) Determination of autocorrelation for serial dependency:
Autocorrelation for serial Dependency was done to establish serial correlation in the inflow series; in this regard the Durbin- Watson (DW or d) standard test was considered. The test employs null hypothesis \( (H_0) \) of no existence of serial correlation or alternative hypothesis \( (H_1) \) of existence of serial correlation. The DW statistic boundary condition lie in the 0-4 range, with a value near two indicating no first-order serial correlation. Positive serial correlation is associated with DW values below 2 and negative serial correlation with DW values above 2. The DW test was evaluated employing the equations below:

\[ DW = \frac{\sum_{i=1}^{n} (\varepsilon_i - \varepsilon_{i-1})^2 - \sum_{i=1}^{n} (\varepsilon_i - \bar{\varepsilon})^2}{\sum_{i=1}^{n} (\varepsilon_i - \bar{\varepsilon})^2} \]

(17)

\[ d = 2(1 - \rho) \]

(19)

where, \( \bar{\varepsilon} \) and \( \varepsilon_i \) are time series at time \( t_i \) is \( \bar{\rho} \) the value of first order autocorrelation coefficient of the time series and \( N \) is the length of data.

(d) Establishment of hurst phenomenon in the inflow time series: The Hurst phenomenon or long-term memory dependency was employed in the inflow time series to establish persistence in the trend. Hypothesis testing was conducted employing rescaled Range Statistic (R/S), the null hypothesis \( (H_0) \) of no long-rang dependence is rejected at 5% significance level if \( Q_{(n_i)} \) is not within the range of 0.809 and 1.862, which is the 95% (asymptotic) level, and the

\[ H_0: \mu_t = \mu_t \]

or

\[ H_1: \mu_t \neq \mu_t \]

Table 1: Monthly in flows, outflows, storage and demand regime.
alternative hypothesis ($H_1$) of existence of long range dependence if $Q \left( \frac{1}{N} \right)$ is within the range of 0.809 and 1.862. Lo’s modified (R/S) method was employed in this research, as the works of [26]. The relationships of the modified (R/S) are as below:

$$S(N,q) = \left( \frac{1}{N} \sum_{j=1}^{N} (x_j - \bar{x})^2 + 2 \sum_{j=1}^{N-1} \alpha_j(q) \sum_{j=1}^{N-1-j} (x_{j+j} - x_j)^2 \right)^{\frac{1}{2}}$$ \hspace{2cm} (20)

where $\bar{x}_N$ denotes the sample mean and the $\alpha_j(q)$ are given by:

$$\alpha_j(q) = 1 - j/(q+1), qN$$ \hspace{2cm} (21)

Lo’s modified R/S statistic is given as:

$$Q_{\alpha} = \frac{1}{\max(N,q)} \left\{ \max_{0 \leq i \leq N} \sum_{j=1}^{N} (x_{j+i} - \bar{x}) - \min_{0 \leq i \leq N} \sum_{j=1}^{N} (x_{j+i} - \bar{x}) \right\}$$ \hspace{2cm} (22)

The data driven formula for choosing $q$ is as:

$$q_{\omega} = \frac{\sqrt{2\rho} \sigma}{\bar{x}}$$ \hspace{2cm} (23)

Results and Discussion

Trend analysis

Analysis of stochastic characteristics like moments and dependence structure of the data series was done to be able to evaluate randomness and trend pattern. In this regards, the time series was examined to establish whether it exhibits intermittency or otherwise as well as seasonal characteristics like trend and moments. Analytical tools employed were moment equations; Autocorrelation for serial dependency, trend, and Hurst Phenomenon. The objective here is to evaluate the seasonality in the moments.

Figures 1,2 depicts the relative variation and dispersion in the means of the rainfall respectively; it shows the presence of seasonality in the moments, meaning that monthly statistics for dry season are significantly different from those of the wet season period. Unlike intermittent stream flow processes, the seasonal means have higher values than the seasonal deviations throughout the year. This can be visualised from Figure 3. The coefficient of variation varies from 0.35494 in the month of June to 2.3847 in the month of January (i.e., period of incipient rains, moderate-peak to late rains). The variance is maximum during the period of late rains and incipient dry season; more or less the interfacing period. This indicates atmospheric instability during this watershed period; i.e., the fringes of the raining season going to full harmattan period.

Figure 4 shows the inter-annual decadal variation in the rainfall series; long term pattern is seemingly evident. However, there is large variability among the monthly values of rainfall of different years, with the period 1985-1994 showing slight increases in the storm evident during peak seasons. The period of 1975-1984 shows the lowest rainfall decadal period. The storm increases from the incipient raining period of April (hydrological period) of the year to the peak period of September. The months of November to March shows dry periods.

Figure 5 shows the annual trend of the rainfall of the study area; it depicts variation in the annual rainfall with 2010 having the highest value with a peak of 1323.51 mm to the lowest of 639.68 mm in 1983. The high storm in 2010 correlate with high reservoir annual inflow of 2010 with the peak of 1467.08333 m$^3$/sec. From the trend analysis, the Z statistic was determined to be 1.98 and has 0.05 level of significance the Sen’s slope $Q$ was determined to be 3.235. The positive value of
the z statistic indicates upward or positive trend this implies that the variable have tendency to increase in future. Also the z value was greater than 1.96 i.e. the chosen significant level of 5%, with $Z_{0.025} = \pm 1.96$, indicating the null hypothesis that trend does not exist is void. Hydrologic processes such as precipitation and runoff evolves on a continuous time scale. Figure 4 time series plot exhibits typical characteristic movement with seasonality, cyclical or sinusoidal and random components. It shows clearly a discernable seasonal or periodic pattern; a random pattern is also evident. In this regard, it suffices to note that even though, monthly and annual rainfalls are usually non-intermittent, in semiarid and arid regions, monthly and annual precipitation may be intermittent. Seasonal autocorrelations for monthly precipitation are generally not significant from zero. This translates that the rainfall time series are uncorrelated, depicting strong homogeneity.

**Trend pattern in the draw-down of the lake at the dam site**

Figure 6,7 show the quinquinal and decadal mean annual draw-down of the reservoir, with month of July having the lowest reservoir level of 133.134 m and December having the peak reservoir level of 141.17 m during the time period under consideration. The low reservoir level in July could be attributed to incipient period of rainfall and the peak at the month of December could be attributed to peak period of rainfall in August and September. Figure 7 shows the inter-annual decadal variation in the reservoir level; long term pattern is seemingly evident. Also there is large variability among the monthly values of reservoir level of different years, with the period 1995-2004 showing slight increases in the draw-down with peak of 141.56 m, evident during peak seasons. The water level of the reservoir is at the lowest in the month of July with reservoir level at 132.66 m in 2005-2014 time period. Figure 6 shows obviously seasonal or periodic pattern; it is a periodic-stochastic series. The analysis of the step or jump trend of the reservoir draw-down shows that the critical value $t_{0.05}(18) = 1.734, t = 4.46$ so that the hypothesis $H_0$ is accepted, as $t (\mu_1 \neq \mu_2)$, meaning that the step or jump trend is significant at 5% probability.

The monthly reservoir flows i.e., inflow and outflow, is shown in Figure 8. The highest inflow was in the month of September with a peak of 2516.39 m$^3$/sec during the period of peak storm. Similarly the outflow was at the peak in the month of October with a recorded value of 1456.98 m$^3$/sec. The low flows recorded were in the periods of incipient storm.

The Mann-Kendall trend test was explored in establishing the presence of trend in the annual flows. From the annual trend analysis of the flow, the Z statistic was determined to be -0.23, the Sen’s slope $Q$ was determined to be -0.012. The value of the z statistic 0.23 is less than 1.96 which indicates the trend in the annual time series is not significant considering the chosen significant level of 5%, with $Z_{0.025} = \pm 1.96$, indicating the alternative hypothesis that trend does
exist is void. This shows that annually there is no significant change in the pattern of the reservoir flow. But in the seasonal Man-Kendall test the value of Z statistic (1.25) (obtained shows the alternative hypothesis $H_a$ that trend those not exist is void. It is worth noting that hydrological processes such as runoff evolve on continuous time scale as depicted in Figure 9, the time series plot exhibit typical characteristic movement with cyclic or sinusoidal behavior. From the trend analysis it shows significant seasonal or periodic pattern in the flow i.e., seasons have significant effect on the flows. Figure 8 shows this as flows from July to November are significantly high.

Analysis of dependence structure

Autocorrelation for serial dependency:
The result obtained in the analysis for serial dependence employing Durbin-Watson (DW) standard Test for monthly rainfall indicates that the null hypothesis ($H_0$) of no serial correlation is accepted. Employing DW boundary condition 0-4 range, with a value near two indicating no first-order serial correlation. Positive serial correlation is associated with DW values below 2 and negative serial correlation with DW values above 2. The lag 1 autocorrelation value obtained was (0.19) i.e., not significantly different from zero. The DW value computed was (1.62) for seasonal rainfall indicates that there is no serial autocorrelation in the seasonal rainfall time series. Also the lag 1 autocorrelation value (0.667) obtained and DW value (0.67) for the seasonal inflow indicates that the null hypothesis ($H_0$) that no serial correlation for the inflow is rejected. The value of the lag 1 serial autocorrelation for the seasonal inflow depicted positive autocorrelation.

Establishment of hurst phenomenon in the inflow time series:
From the modified rescaled range statistic using Lo’s (1991) qopt, the null hypothesis of no long-range dependence is rejected at 5% significance level i.e., interval of $(0.809-1.862)$ considering the computed $QN$, $q$ value of 3161.6.

Conclusion

The assessment of stochastic characteristics of the time series like moments and dependence structure were done to be able to understand the dynamics or randomness and trend pattern of the annual and monthly series. The analysis showed that there is no serial autocorrelation in the seasonal rainfall time series. The seasonal inflow depicted positive autocorrelation while the analysis of long term dependency on the inflows indicates long term dependency of the inflow time series. The trend test for the rainfall, draw-down and inflow time series showed that the rainfall time series exhibit positive trend. The analysis of the step or jump trend of the reservoir draw-down shows that the step or jump trend is significant and positive.

References


