Chinese Journal of Oceanology and Limnology
Vol. 26 No. 2, P. 142-151, 2008
DOI: 10.1007/s00343-008-0142-0

Analysis of stochastic characteristics of the Benue River flow process*

Martins Y. OTACHE**, Mohammad BAKIR, LI Zhijia (李致家)**
State Key Laboratory of Hydrology-Water Resources & Hydraulic Engineering, Hohai University, Nanjing 210098, China

Received Oct. 23, 2007; revision accepted Dec. 12, 2007

Abstract Stochastic characteristics of the Benue River streamflow process are examined under conditions of data austerity. The streamflow process is investigated for trend, non-stationarity and seasonality for a time period of 26 years. Results of trend analyses with Mann-Kendall test show that there is no trend in the annual mean discharges. Monthly flow series examined with seasonal Kendall test indicate the presence of positive change in the trend for some months, especially the months of August, January, and February. For the stationarity test, daily and monthly flow series appear to be stationary whereas at 1%, 5%, and 10% significant levels, the stationarity alternative hypothesis is rejected for the annual flow series. Though monthly flow appears to be stationary going by this test, because of high seasonality, it could be said to exhibit periodic stationarity based on the seasonality analysis. The following conclusions are drawn: (1) There is seasonality in both the mean and variance with unimodal distribution. (2) Days with high mean also have high variance. (3) Skewness coefficients for the months within the dry season period are greater than those of the wet season period, and seasonal autocorrelations for streamflow during dry season are generally larger than those of the wet season. Precisely, they are significantly different for most of the months. (4) The autocorrelation functions estimated "over time" are greater in the absolute value for data that have not been deseasonalised but were initially normalised by logarithmic transformation only, while autocorrelation functions for i = 1, 2, …, 365 estimated "over realisations" have their coefficients significantly different from other coefficients.

Keyword: trend; stationarity; seasonality; over time; over realisation; stochastic; skewness

1 INTRODUCTION

Although data used in the design, planning and operational studies of water resource schemes could be restricted to historical records; however, a serious drawback may then result, because a particular sequence of observations does not occur in an identical form over a future period; thus, the principal aim of time series analysis is to describe the history of movement in time of some variables such as the rate of flow in a river at a particular site. This becomes increasingly important considering that river flow and other hydrological sequences are essentially characterized by variability and oscillatory behavior. As such, the over riding objective of any time series study is to understand the mechanism that generates the data and also but not necessarily, to produce likely future sequences or forecast events over a short period of time. These are attempted by making inferences regarding the underlying laws of stochastic process from one or more sequences of recorded observation and then by postulating a model that fits the data, which are again used for estimation purpose. But to do this, it is necessary to identify and analyze the different components of a given time series.

Since the past two decades, extensive researches have been devoted to developing methods to analyze stochastic characteristics of hydrologic time series and to identifying, and testing the goodness of fit to stochastic models. Towards this end, studies have been carried out to examine the stochastic properties of streamflow process and to wit, detect changes in the streamflow processes globally (Zhang et al., 2001; Burn and Hag Elnur, 2002; Lins

* Supported by the Post Graduate Research Fund from Federal Government of Nigeria under the BEA Scholarship Program (No.CSC2005566002)
**Corresponding authors: martinso3@yahoo.com; zhijia-li@vip.sina.com

and Slack, 1999; Wang et al., 1999). In doing so, from the view of hydrological modeller, it is an important task to study the characteristics (components), say trend, stationarity, shifts/jumps, and dependence structures of streamflow processes. The objective is to determine if there is the existence of any trend or other forms of non-stationarity in the streamflow data and how to achieve stationarity when the data is otherwise non-stationary. It is imperative to detect the trend and stationarity, and analyse any form of dependence structures in a hydrological time series to understand possible links between hydrological processes and global environmental changes; for instance, the autocorrelation structure has special significance in linear dynamics modelling.

It is generally assumed that annual hydrologic time series have homogeneous mean, variance, skewness, and dependence structure; similarly, seasonal series generally have periodic mean, standard deviation, skewness, and dependence structure (Salas et al., 1982). Studies of the characteristics of river flows have been made by means of the theory of stochastic processes (Papoulis et al., 1965). Stochastic characterisation of the underlying processes is vital in constructing mathematical models that are generally used for stochastic simulations and forecasting of hydroclimatic processes. The stochastic characteristics of such processes, for example, streamflow series, depend largely on the type of data at hand; data may be available on a continuous time scale or at discrete points in time series defined on hourly, daily, weekly, monthly, and annual time intervals. Basically, the information contained in a discrete series is affected to a great extent by the choice of sampling interval which depends largely on the purpose to be served (Kottegoda et al., 1990).

Though the relevance of the choice of sampling interval with regards to the information content of a time series cannot be over-emphasised, there is a dearth of substantial information available in existing literature as to the appropriate length of data series to be used for analysis. Considering this, the study is carried out under conditions of data austerity with the view to examining whether effective and meaningful conclusions could be drawn based on the available length of river flow data; this is important in view of the seeming fact that in a non-stationary time series, statistical properties differ from one segment to another and are therefore time dependent. Examples of these are daily river flow and other series in which seasonal changes occur.

For this study, time-based characteristics of the river flow data are examined. The stochastic analysis or examination deals with the diagnosis of trend, stationarity and periodicity/seasonality in autocorrelation structures. The paper is organised as follows: (i) Trend analysis, (ii) Stationarity, (iii) Periodicity/Seasonality in autocorrelation structures with conclusions on each subject headings accordingly presented.

2 HYDROLOGY OF THE STUDY RIVER (BENUE RIVER)

The Benue River is the major tributary of the Niger River. It is approximately 1 400 km long and almost navigable during the rainy season; precisely, July, August, September to October, the height of the rainy season. Hence, it is an important transportation route in the regions it flows through. Its headwaters rise in the Adamawa Plateau of the Northern Cameroon, then into Nigeria south of the Mandara Mountains through the east-central part of Nigeria before entering the Niger River at Lokoja (Fig.1).

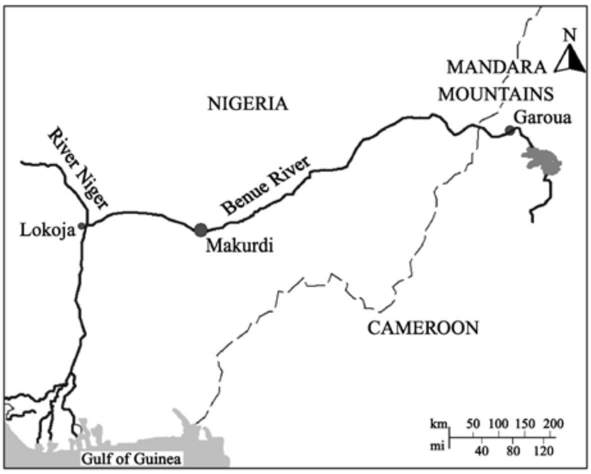


Fig.1 Map of the Benue River and its traverse

The wide flood plain is used for agriculture, with main crops being sugar cane and rice. There is only one high-water season because of its southerly location; this normally occurs from May to October while on the other hand, the low-water period is from December to June. There are definite wet and dry seasons which give rise to changes in river flow and salinity regimes. The flood of the Benue River (upper, middle, and downstream) lasts from July to October, and sometimes, early November.

In 1982, the upstream of the Benue River (i.e., in the Upper Benue River Basin) was impounded for hydropower generation, irrigation and fisheries (Lagdao Dam) around the Garoua (Fig.2); the surface area of the reservoir covers 700 km^2. This was not without its appurtenant damages. Based on routine monitoring of the Benue River (upper stream), it was found that some immediate deleterious effects are (Toro, 1997): (i) siltation of the riverbed and channel, (ii) frequent flooding events, and (iii) reduction in flow. The consequences of these developments include: (a) constrained irrigation, navigation and fishing activities which were formerly undertaken along the river, (b) siltation of water supply intake structures and irrigation abstraction facilities along the river. Correspondingly, these effects have a carry over effect downstream of the river.

3 DATA USED

In this study, historical time series for gauging stations at the base of the Benue River (i.e., Lower Benue River Basin) at Makurdi (centre at 7°44′ N, 8°32′ E) location is used; a total of 26 years (1974–2000) water stage and discharge data were collected and used.

To investigate the stochastic characteristics of the Benue River flow process, the daily flow data are aggregated to monthly and annual data series by taking the average of each month's flow and calendar year. Similarly, the annual maximum and minimum daily average discharges are obtained according to the water year for the streamflow process.

4. ANALYSIS OF STOCHASTIC CHARACTERISTICS

4.1 Trend analysis

Trend analysis is done at timescales of annual and monthly. A rank-based non-parametric method, the Mann-Kendall's test is adopted to study the trends in both the annual and monthly series. The choice of this method is based on the fact that it has the advantage of being less sensitive to outliers over the parametric method. The essence of this test at the annual timescale is to be able to obtain an overall view of the changes in the streamflow process and then at the seasonal/periodic scale for changes in seasonal patterns.

4.1.1 Mann-Kendall test for annual series

Kendall (1938) proposed a measure τ to determine the strength of the monotonic relationship between variables; in furtherance of this approach, Mann (1945) suggested a test for the significance of Kendall's τ, where one of the variables is time as a test for trend. This test is popularly known as the Mann-Kendall's (MK) test. In using this method, it is advisable to pre-whiten the time series by removing the impact of serial correlation from the series through $m_i = x_i - \phi x_{i-1}$ (Wang et al., 2005), where m_i is the pre-whiten series value, x_i is the original series value, and ϕ is the estimated lag 1 serial correlation. The pre-whitening is necessary because it has been found that positive serial correlation inflates the variance of the MK statistic S thereby increasing the possibility of rejecting the null hypothesis of no trend (Von, 1995).

In order to test for the presence or non-existence of trend in the flow sequence, the mean daily flow values were aggregated to annual mean values. The null hypothesis H_0 for this test is that a flow series $\{x_1,..........,x_N\}$ come from a population where the random variables are independent and identically distributed.

The Mann-Kendall (MK) test statistic S is expressed as

$$S = \sum_{i=1}^{N-1} \sum_{k=i+1}^{N} \operatorname{sgn}(x_k - x_i) \quad (1)$$

where

$$\operatorname{sgn}(x) = \begin{cases} +1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

Mann-Kendall test statistic tau, τ is computed as

$$\tau = 2S/(N(N-1)) \quad (2)$$

and

$$\sigma_s^2 = \frac{1}{18}\left[\begin{array}{l} N(N-1)(2N+5) - \\ \sum_{i=1}^{m} p_i(p_i-1)(2p_i+5) \end{array}\right]$$

where m is the number of tied groups in the data set and p_i, the number of data points in the ith tied group.

Similarly too, under the null hypothesis, the quantity Z is taken to be standard normally distributed.

Based on this,

$$Z = \begin{cases} (S-1)/\sigma_s & S>0 \\ 0 & S=0 \\ (S+1)/\sigma_s & S<0 \end{cases} \tag{3}$$

4.1.2 Mann-Kendall test results for annual series

The Mann-Kendall test results as shown in Table 1, indicate that at 5% level of significance, i.e., ±1.96, the computed value of the test statistic Z, (i.e., two-tailed) for each cases is within the range of ±1.96, so there is no reason to suspect the presence of trend in the annual flow series. This result is in conformity with that of the heuristic by visual examination as shown in Fig.2. The ten-year segmented data series for three periods do not indicate a discernible presence of trend; it shows a unimodal distribution and in between the periodic segments, the peak flow pattern is random (Fig.2). As shown in Table 1, the values of the p statistic are not statistically significant (Table1) which agrees with the heuristic of no probable reason to suspect the presence of trend.

4.1.3 Mann-Kendall test for monthly series

The monthly flow processes for the entire 26 water-years were examined in detail for possible changes in trend with the Mann-Kendall test that allows for serial dependence. In line with Hirsch et al. (1982), the Kendall test that allows for seasonality in observations collected over a time period by computing the Mann-Kendall test on each seasons is used.

Table 1 Mann-Kendall tests on annual Series

Flow series	Statistics		
	τ	Z	p-value
Annual means	−0.13	−0.88	0.3789
Annual maxima	−0.19	−1.37	0.1707
Annual minima	−0.19	−1.37	0.1707

Let the monthly flow series be represented by the matrix

$$\mathbf{X} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix} \tag{4}$$

Here, p is the number of seasons for n years under consideration; similarly, let the matrix

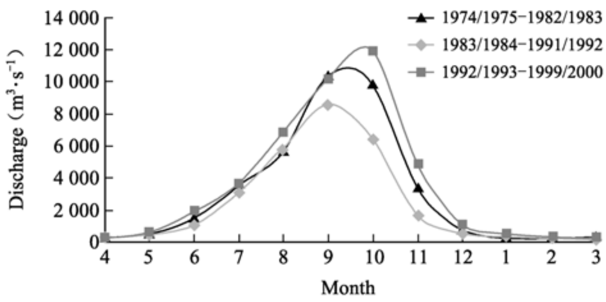


Fig.2 Seasonal flow patterns

$$\mathbf{R} = \begin{pmatrix} R_{11} & R_{12}\cdots & R_{1p} \\ R_{21} & R_{22}\cdots & R_{2p} \\ \vdots & \vdots & \vdots \\ R_{n1} & R_{n2}\cdots & R_{np} \end{pmatrix} \tag{5}$$

denote the ranks corresponding to the observations in x where the n observations for each season are ranked among themselves. Thus each column of R is a permutation of (1, 2, . . . , n). Specifically, the rank matrix R_{ij} is computed as

$$R_{ij} = \frac{1}{2}\left[n+1+\sum_{k=1}^{n} \operatorname{sgn}\left(x_{ij}-x_{kj}\right) \right] \tag{6}$$

The Mann-Kendall test statistic for each season is

$$S_i = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}\left(x_{ji}-x_{ki}\right) \tag{7}$$

where n = 26 (water years); i = number of seasons (12) and a season is defined as one calendar month; S_i is the S-statistic in the MK test for season i (i = 1, 2, …, 12).

$$S' = \sum_{i=1}^{p} S_i \text{ , } p = \text{seasons; } \sigma^2_{S'} = \sum_{i=1}^{p} Var\left(S_i\right).$$

In the presence of serial correlation, as in the monthly flow processes, the variance of S' is defined (Hirsch and Slack, 1984) as

$$\sigma^2_{S'} = \sum_{i=1}^{p} Var\left(S_i\right) + \sum_{g=1}^{p-1} \sum_{h=g+1}^{p} \sigma_{gh} \tag{8}$$

where the covariance matrix σ_{gh} is expressed as

$$\hat{\sigma}_{gh} = \frac{1}{3}\left[K_{gh} + 4\sum_{i=1}^{n} R_{ig}R_{ih} - n\left(n+1\right)^2 \right] \tag{9}$$

$$K_{gh} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sgn}\left[\left(x_{jg} - x_{ig}\right)\left(x_{jh} - x_{ih}\right)\right] \quad (10)$$

This is for a no-missing data situation, and *g* and *h* are different seasons respectively with the test statistic *z*' which is standard normally distributed, is evaluated as:

$$z' = \begin{cases} \left(S' - 1\right) / \sigma_{s'} & S' > 0 \\ 0 & S' = 0 \\ \left(S' + 1\right) / \sigma_{s'} & S' < 0 \end{cases} \quad (11)$$

4.1.4 Mann-Kendall test results for monthly series

Based on the trend analysis as shown in Tables 2, 3, even though the aggregate yearly and monthly *p*-values, by the measured statistics indicate to be statistically insignificant connoting the absence of any discernible presence of trend, month by month values are to the contrary. Presence of change in the trend situation seems to be evident for the months of January, February and August; this can be explained against the background of high fluctuations usually encountered in these months due to high seasonality effect often occasioned by protracted rainfall variability. The prevailing situation could be attributable to the flow regimes of the Benue River upstream which is being hampered by the sudden releases and occasional impoundments of water resulting from the impact of the Lagdao dam built across it (Toro, 1997).

Despite this though, detection or assessment of trend requires that data be collected at a given location by using consistent collections and measurement techniques on a regular schedule and over a substantial number of years. Thus, seasonal variation in trend as indicated in the preceding section is far from being significant for this interval of time considering the short length of data series used for analysis. This assertion is informed based on the presence of strong seasonality that results in the existence of different distributions for different times of the year and more so that one of the problems in detecting and evaluating trends in hydrologic data is the confounding effect of serial dependence.

Table 2 Seasonal Kendall tests on monthly flow Series

Streamflow	*S*	*τ*	*Z*-Statistic	*p*-value
Makurdi Station	677	0.166 6	0.01	0.992 0

4.2 Stationarity test and analysis

4.2.1 ADF stationarity test

The stationarity test is carried out by using the augmented Dickey-Fuller (ADF) unit root test method that was first proposed by Dickey and Fuller (1979) and modified by Said and Dickey (1984). Because the ADF test is based on linear regression that assumes a normal distribution, log-transformation of the flow data can convert exponential trend that might possibly be present in the data series (Gimeno et al., 1999; Wang et al., 2005). As such, for this test, before applying the ADF test, this pre-processing was carried out; in addition, the flow series was also deseasonalised to test for the effect of seasonality in the series.

One important practical aspect of the ADF test is the specification of the truncation lag values. To solve this problem, the choice of the number of lag length is suggested as $p = \text{int}[x\ (N/100)^{1/4}]$ (Wang et al., 2005); the values of x chosen for this study though subjectively, are 3, 4, and 12 for the daily and monthly flows whereas it was pegged at 1 for annual flow in order to exclude the effect of serial correlation as suggested in Wang et al (2005) (Wang et al., 2005) and implemented in unit root DLL (Beta) (Kurt Annen).

4.2.2 Stationarity test result

From the test analysis, the null hypothesis of a unit root is rejected in favour of the stationary alternative at a 5% level of significance for both the daily and monthly flow series, whereas it is accepted for the annual flow series even at 1% and 10% significant levels. This conclusion is informed by the values of the t-statistic; despite this, for monthly flow series, the extent of stationarity could not be determined using this framework whether it is level stationary or trend stationary; but considering the fact that monthly flow series are known to exhibit strong seasonality, it may be hypothesized that at best, it is periodic stationary. As shown in Table 4, deseasonalising the flow series has no discernible impact on stationarity.

4.3 Seasonality analysis

4.3.1 Seasonality in autocorrelation structure

Assume that the daily flow time series, without any gaps covering, say, an N year observation period is compiled as follows:

Table 3 Mann-Kendall Test for seasonal Streamflows at Makurdi

Variable	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mar
S	60	47	27	27	96	31	57	72	59	94	83	24
τ	0.18	0.14	0.08	0.08	0.30	0.10	0.18	0.22	0.18	0.29	0.26	0.07
Z	1.30	1.01	0.57	0.57	2.09	0.66	1.23	1.56	1.28	2.05	1.81	0.51
p-value	0.194	0.313	0.569	0.569	0.037	0.509	0.219	0.119	0.201	0.040	0.07	0.610

Table 4 ADF Unit root test (constant with no trend model)

Series	Lag	Information criterion					
		Akaike		Schwarz		Hannan-Quinn	
		t-s*	p-value	t-s*	p-value	t-s*	p-value
Log-daily	38	−9.673	0.00	−7.033	3.3E-14	−8.379	7.0E-32
	13	−5.588	5.52E-07	−5.588	5.52E-07	−5.588	5.52E-07
	10	−5.046	1.47E-05	−4.791	5.25E-05	−5.046	1.47E-05
Log-deseasonalised daily	38	−8.912	0.00	−9.681	0.00	−9.264	0.00
	13	−9.539	0.00	−9.681	0.00	−9.539	0.00
	10	−10.062	0.00	−10.752	0.00	−10.062	0.00
Log-monthly	16	−10.402	0.00	−1.152	0.00	−10.402	0.00
	6	−10.967	0.00	−10.967	0.00	−10.967	0.00
	4	−9.329	1.08E-30	−9.329	1.08E-30	−9.329	1.08E-30
Log-deseasonalised monthly	16	−10.895	0.00	−10.895	0.00	−10.895	0.00
	6	−10.949	0.00	−10.949	0.00	−10.949	0.00
	4	−9.363	3.38E-31	−.363	3.38E-31	−9.363	3.38E-31
Log-annual	1	−2.530	0.121	−2.530	0.121	−2.530	0.121
Log-deseasonalised annual	1	−2.530	0.121	−2.530	0.121	−2530	0.121

* t-s: t-statistic value

$$\begin{pmatrix} x_{1,1} & \cdots & x_{1,365} \\ \vdots & \ddots & \vdots \\ x_{N,1} & \cdots & x_{N,365} \end{pmatrix} \tag{12}$$

In compiling matrix (12), the flow on the last day of a leap year is omitted. It is assumed that this might not introduce any significant error but rather allows symmetry of notation (Mitosek, 2000); Wang et al., 2005). It is assumed that the seasonal effects repeat in the same time periods of any year as a result of the time-invariant annual circulation of the Earth around the Sun and the daily rotation of the Earth. The flow time series is log transformed or normalised and then deseasonalised (i.e., standardised). After normalisation, for each *i*th column (i.e., day), test statistics (mean, standard deviation and coefficient of variation) are computed; in this case, successive columns contain flows of the same period in an annual cycle.

Mean: $$\bar{X}_i = \frac{1}{N}\sum_{j=1}^{N} X_{j,i} \tag{13}$$

Standard deviation:

$$s_i = \left(\frac{1}{N}\sum_{j=1}^{N}(X_{j,i} - \bar{X}_i)^2\right)^{1/2} \tag{14}$$

Coefficient of variation:

$$CV_i = \frac{s_i}{\bar{X}_i} \tag{15}$$

Standardisation

$$m_{j,i} = \frac{x_{j,i} - \bar{X}_i}{s_i} \tag{16}$$

Matrix (12) is rewritten as one realisation according to the chronology of events:

$$\left\{x_{1,1}, x_{1,2}, \ldots, x_{1,i}, \ldots, x_{1,365}, x_{2,1}, \ldots, x_{j,i}, \ldots, x_{N,365}\right\}$$

This after renumbering becomes $\{x_1, x_2, ..., x_n\}$, where $n = 365N$; it serves as a basis for "over time" estimation of the mean value

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_i, \quad (17)$$

and the autocorrelation function $r(k) = c_k / c_0$, for $k = 0, 1, 2\ldots$ where

$$c_k = \frac{1}{n-k}\sum_{i=1}^{n-k}(x_i - \bar{x})(x_{i+k} - \bar{x}) \quad (18)$$

Taking into consideration of annual cyclicity in autocorrelation function, it is necessary to examine the daily flow process in the form of matrix (12) and to estimate values of the autocorrelation function "over realisation" between components X_i and X_{i+k} of the 365-dimensional random variable (2), where $i = 1, 2, \ldots, 365$ and $k = 0, 1, 2, \ldots, k_{\max}$, ($k_{\max} \le 365$), described by (Mitosek, 2000); Wang et al., 2005).

$$r(X_i, X_{i+k}) \equiv r_i(k):$$

$$r_{i(k)} = \begin{cases} \dfrac{\frac{1}{N}\sum_{j=1}^{N}(x_{j,i} - \bar{x}_i)(x_{j,i+k} - \bar{x}_{i+k})}{s_i s_{i+k}}, & \text{for } i + k \le 365 \\ \dfrac{\frac{1}{N-1}\sum_{j=1}^{N-1}(x_{j,i} - \bar{x}_i)(x_{j+1,i+k-365} - \bar{x}_{i+k-365})}{s_i s_{i+k-365}}, & \text{for } i + k > 365 \end{cases} \quad (19)$$

where

$$\bar{x}_i = \frac{1}{N}\sum_{j=1}^{N} x_{j,i};$$

$$\bar{x}_{i+k-365} = \frac{1}{N-1}\sum_{j=1}^{N} x_{j+1,i+k-365}$$

and

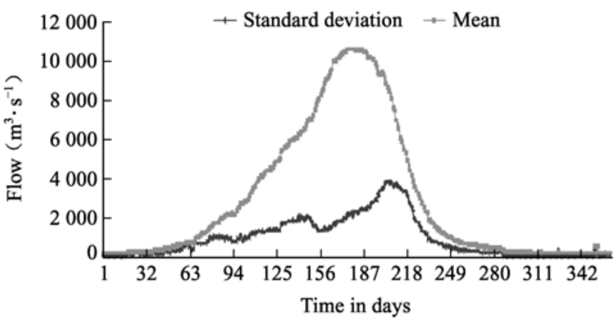
$$s_i = \left(\frac{1}{N}\sum_{j=1}^{N}(x_{j,i} - \bar{x}_i)^2\right)^{1/2}$$

$$s_{i+k-365} = \left(\frac{1}{N-1}\sum_{j=1}^{N-1}(x_{j+1,i+k-365} - \bar{x}_{i+k-365})^2\right)^{1/2}$$

Eq.(19) is a function of the lag k and the variate i, herein the successive days. The equation can be re-arranged and applied to monthly series.

4.3.2 Results of seasonality analysis

Based on the analysis, as seen in Fig.3, there is obvious seasonality in both the daily mean and standard deviation as well as in the coefficient of variation; while the mean values are greater than the standard deviations, the coefficient of variations on the other hand, cannot be considered to be constant even in the slightest approximation. The variations in the mean and standard deviations are characterised by unimodal distribution with the extremum in September as days with high mean values also have high standard deviations. From Figs. 4 and 5, one can appreciate the implications of evaluating the autocorrelation functions as averages on all months; it is succinctly clear from Figure 4 that the autocorrelation functions for $i = 1, 2 \ldots, 365$



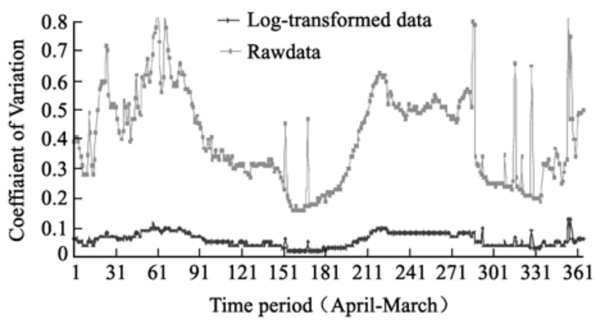


Fig.3 Variation in daily mean, standard deviation (left**) and seasonal coefficient of variation (**right**) over an annual cycle (raw and log-transformed flow data)**

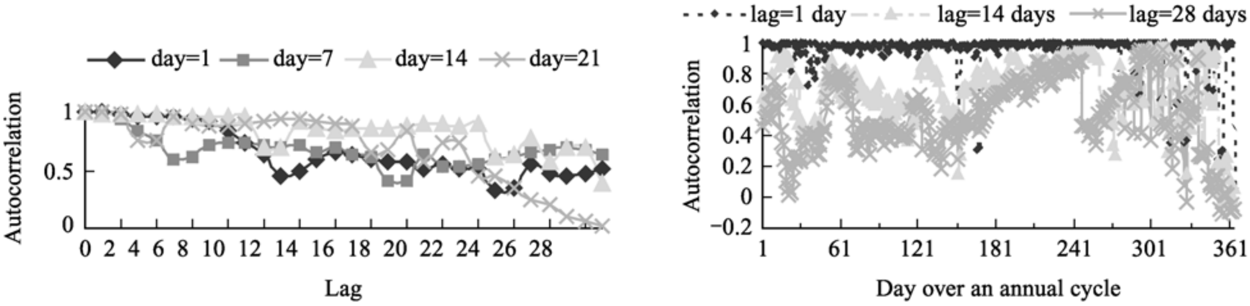


Fig.4 Daily autocorrelation function "over realisations at different lags for days (left**) and lags at day-to-day (**right**)**

estimated "over realisations" have their coefficients significantly different from other coefficients (at least half), whereas the autocorrelation functions estimated "over time" are greater in the absolute value for data which have not been standardized but were initially normalised by logarithmic transformation only (Fig.5). The seeming implication of this is that the autocorrelation functions estimated "over realisations" should not be replaced with autocorrelation function estimated "over time" since they are not identical considering the fact that, especially for daily river flow process, it is not ergodic (Mitosek, 2000). Fig.5(upper) shows an exceedingly obvious periodicity (sine wave) indicating a strong serial persistence in the un-processed (raw) daily flow series; thus the daily flow series may not be second order stationary but rather, periodic stationary. As a result of the strong serial correlation, in modelling the daily flow series, working with the residuals may eliminate or reduce the persistence in the data (Fig.5(lower)).

For the monthly flow series analysis (Table 5), values of the skewness coefficients for the months within the dry season period are generally greater than those of the wet season indicating that data in the dry season depart more from normality than those for the wet season. Similarly, seasonal correlations for streamflow during the dry season are generally larger than those for the wet season, and are significantly different for most of the months except for some few cases that could be attributable to variability in the data (Table 6). Like as in the case for daily flow series, the raw monthly flow series exhibits high seasonality pattern in its autocorrelation structure as shown in Fig.6; this figure does illustrates further the relevance of working with the flow residuals rather the flow data in its original form, especially for the type of analysis or any similar sort as in this paper. Doing so without proper pre-processing may lead to biased results.

On the other hand, how much the length of the data as well as its variability may have impacted negatively on the results for this analysis cannot be ascertained directly; but one thing is certain, the dearth of continuous data on a large scale can create indeterminacy problem in a stochastic analysis of this form.

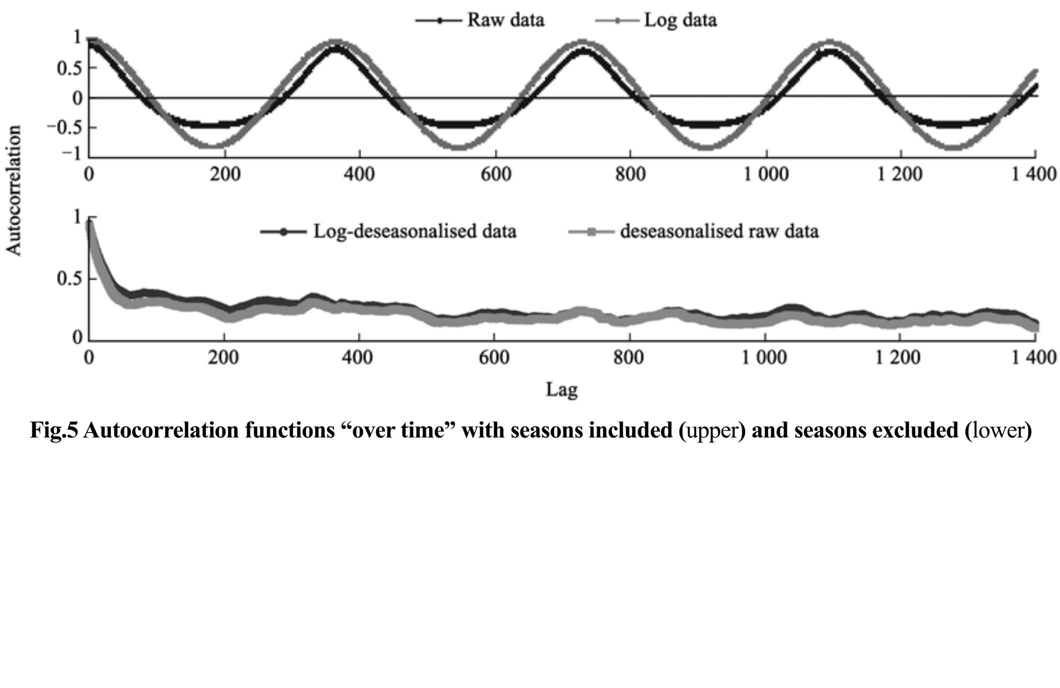


Fig.5 Autocorrelation functions "over time" with seasons included (upper**) and seasons excluded (**lower**)**

Table 5 Skewness coefficients of monthly flow

Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mar
0.390	0.423	0.516	0.270	0.264	0.160	0.312	0.529	0.481	0.365	0.212	0.423

Table 6 Seasonal autocorrelation coefficients over realisations of monthly flow (*r* (*i*, 1))*

Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mar
0.27	0.79	0.62	0.46	0.38	0.64	0.14	0.84	0.43	0.67	0.28	0.30

* r stands for autocorrelation function), i as used in the paper stands for the "day index", i.e., different days and 1 stands for lag 1. Hence, $r(i,1)$ connotes the autocorrelation function values over realisations; stand for different days as indexed by i

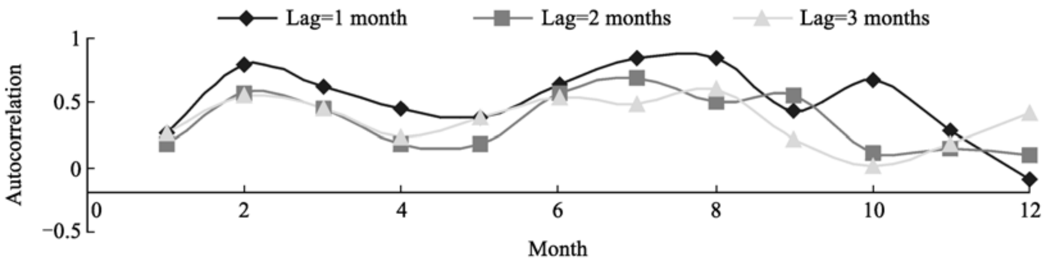


Fig.6 Seasonal autocorrelation functions for raw monthly flow series (hydrological year: April-March**)**

5. CONCLUSION

The stochastic characteristics (trend, stationarity, and seasonality) of the streamflow series of the Benue River are examined in this study with a view to providing a general insight of the dynamics of its regime. Resulting from the analyses, the following main conclusions can be drawn.

In general, there is no discernible reason to suspect the presence of probable trend in the annual flow series; though the aggregate yearly and monthly values of the p statistic are statistically insignificant, there is marginal presence of positive change in trend in some of the months. On the other hand, while the stationarity test shows that both the daily and monthly flow series are stationary, for annual series, it is rejected. Further analysis for seasonality indicates the presence of obvious seasonality in the streamflow series. It is also noted that in the evaluation of seasonality in the autocorrelation structures that autocorrelation functions estimated "over time" should not replace those estimated "over realisations" since they are not identical. Considering the fact that limited data is used for analysis in this study, the results obtained here are inconclusive from practical point of view rather than academic and thus subject to further analysis.

References

Burn, D. H. and M. A. Hag Elnur, 2002, Detection of hydrologic trends and variability. *Hydrology* **255**: 107-122.

Dickey, D. A. and W. A. Fuller, 1979. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **74**: 423-431.

Gimeno, R., B. Manchado and R. Minguez, 1999. Stationarity tests for financial time Series. *Physica A* **269**(1): 72-78.

Hirsch, R. M., J. R. Slack and R. A. Smith, 1982. Techniques of trend analysis for monthly water quality data. *Water Resources Research* **18**(1): 107-121.

Hirsch, R. M. and J. R. Slack, 1984. A non-parametric trend test for seasonal data with serial dependence. *Water Resources Research* **20**(6): 727-732.

Kendall, M. G, 1939. A new measure of rank correlation. *Biometrika*, **30**: 81-93.

Kottegoda, N. T., 1980. Stochastic Water Resources Technology. The Macmillan Press Ltd, London and Basingstoke.

Kurt Annen, *Unit root DLL (BETA)*; www.web-reg.de

Lins, H. F. and J. R. Slack, 1999. Streamflow trends in Sweden 1807-2002. *Hydrol. Sci. J.* **4**9(1): 69-83.

Mann, H. B, 1945. Non-parametric test against trend. *Econometrica*, **13**: 245-259.

Mitosek, H. T, 2000. On stochastic properties of daily river flow processes. *J. Hydrology* **228**: 188-205.

Papoulis, A., 1965. Probability, Random Variables, and Stochastic Processes. McGraw-Hill. New York.

Salas, J. D., D. C. Boes and R. A. Smith, 1982. Estimation of

ARMA Models with Seasonal Parameters. *Water Resources Research* **18**(4): 1 006-1 010.

Said, S. E. and D. Dickey, 1984. Testing for unit roots in autoregressive moving-average models with unknown order. *Biometrika*, **71**: 599-607.

Toro, S. M, 1997. Post construction effects of the Cameroonian Lagdao Dam on the River Benue. *J. Chartered Institute, Water Environmental Management,* **11**(2): 109-113.

Von Storch, H., 1995. Misuses of statistical analysis in climate research: *In*: H. V. Storch and A. Navarra (eds). Analysis of climate: Applications of Statistical Techniques. Springer-Verlag, New York. pp. 11-26.

Wang, W., P. H. A. J. M. van Gelder and J. K. Vrijing, 2005. Detection of changes in Streamflow series in Western Europe over 1901-2000. *Water Sci. and Technology*, *Water Supply* **5**(6): 289-299.

Zhang, X., K. D. Harvey and R. Uzyk, 2001. Trends in Canadian streamflow. *Water Resources Research* **37**(4): 987-998.