Modelling of River Flow System: Issues for Water Resources Planning and Management

Otache, M. Y[1]., Jimoh, O. D[2]., Abubakar, A. S[3] and Animashaun, I. M[1]

[1]Department of Agricultural & Bioresources Engineering, Federal University of Technology, Minna

[2]Department of Civil Engineering, Federal University of Technology, Minna

[3]Department of Geography, Federal University of Technology, Minna

Abstract

Planning and management of water resources systems in recent time predictably entail modelling. This has assisted to a great extend in understanding some hydrological processes and river flow dynamics; however, there are still some flaws or bottlenecks arising from model error. Though, inherent limitation of models as representations of any real system is crystal clear, appropriateness of variables and their combination couple with their mathematical relationships could bring closeness of models to the real system. Thus, stochastic approach to river flow system is better incorporated with possibility for nonlinearity and heteroscedasticity of the entire dynamics for determining an appropriate model structure. More importantly, since the nonlinear mechanism, vis-a-vis conditional heteroscedasticity in hydrological processes has been very austere, modelling data with time varying conditional variance could be attempted in various ways using either the non-parametric or parametric approach in the context of integrated modelling and forecasting of hydrologic processes.

Introduction

Inherent in the principles of water resources management is the judicious utilisation and conservation of the available water resources. One of the ways to enhance this is the proper estimation of water demand both quantitatively and qualitatively. Within this overall management system, the hydrologist is often required to estimate the magnitude of the extreme events, whereas operations of some of the design works are dependent on reliable estimates of flow for an ensuing period of time. Since river is an essential component of the hydrologic cycle, its flow forecasting provides veritable and basic information on a wide range of problems related to the design and operation of the entire river system. A very common constraint encountered in the context of water resources planning is inadequacy of streamflow records. Thus, a system designed on the basis of the historical records only faces a chance of being inadequate for the unknown flow sequence that the system might experience. Hence, the reliability of a system has to be evaluated under different probable scenarios which might not be possible with the historical records alone.

Besides the issues of forecasting, one aspect that bothers hydrologists of late is the structure of hydrological processes. The tremendous spatial and temporal variability of the hydrological processes has been believed until recently to be due to the influence of a large number of variables (Otache *et al.*, 2011a). Consequently, the majority of previous investigations on modelling hydrological processes have essentially employed the concept of stochastic processes (Otache *et al.*, 2011b); but however, recent studies have indicated that even simple deterministic systems influenced by a few nonlinear interdependent variables might give rise to very complicated structures, i.e., deterministic chaos (Sivakumar, 2000). It is imperative therefore to note that since river flow dynamics is linked both to the climate, through precipitation, radiation, etc., and to the inflow-runoff transformation (Sivakumar, 2000), the existence of chaos cannot be excluded *a priori*. However, this does not connote that river flows are exclusively reducible to a nonlinear interactions of a few degrees of freedom purely based on the concept of fractals or nonlinear determinism.

Basically, in the light of the desire to understand the river flow dynamics, it suffices to realise that a physical variable that is not very useful for forecasting on its own can often be useful when used in conjunction with other variables. Given the number of physical variables that could be considered as potentially relevant, it is apparent that a very large number of different combinations of both variables and mathematical relationships that link them together are available when developing a forecasting model. But then determining an appropriate model structure by trial-and-error process is not always practical (Zealand *et al.*, 199); this is so because river flow is usually treated as a random process, purely stochastic (Otache *et al.*, 2011a); rather than incorporating the possibility for nonlinearity and at best, heteroscedasticity of the entire dynamics (Otache *et al.*, 2011a).

Materials and Methods

(a) Data

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 was used; a total of 26 years (1974–2000) water stage and discharge data were obtained and used. For purposes of this study, the daily flow data were 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, i.e., months of April to March for the streamflow process. In doing this, the whole data series was considered rather than being separated into parts conforming with the period from which the river was dammed upstream; the simple reason being that the distance from Makurdi to the dam site upstream (Lagdao dam in Cameroon) is relatively large and thus it is believed that whatever effect the dam operations may have down the river channel must have been significantly mitigated.

(b) Stochastic Modelling

(i) Thomas Fiering or Markov Model (i.e. lag one)

Thomas and Fiering describe a linear stochastic model for simulating synthetic flow data. On a monthly basis this model represents the means, the standard deviations, the serial correlations between successive flows, and the Skewness. This model uses a linear regression relationship to relate the flow Q_{t+1} in the (t+1)th month, (t being from the start of the generated sequence) to the flow Q_t in the t month. If $\bar{Q}_j$ and $\bar{Q}_{j+1}$ be the mean monthly discharges during months j and j+1, respectively, within a repetitive annual cycle of 12 months, b_j be the regression coefficient for estimating the flow in the (j+1)th month from the j month, and t be a normal deviate with zero mean and unit variance, the Thomas-Fiering Schema can be expressed according as in equation (1).

$$Q_{t+1} = \bar{Q}_{j+1} + b_j\left(Q_t - \bar{Q}_j\right) + t_t\sigma_{j+1}\left(1 - r_j^2\right)^{1/2} \quad 1$$

If an average first-order serial correlation coefficient r_1 is used to replace the 12 monthly r_j values, it can easily be shown using the relationship

$$y_t = \frac{Q_t - \bar{Q}_j}{\sigma_j} \quad 2$$

that the model (1) is the first-order case of the general non-seasonal autoregressive model

$$y_t = \sum_{j=1}^{p} \phi_j y_{t-j} + a_t \quad 3$$

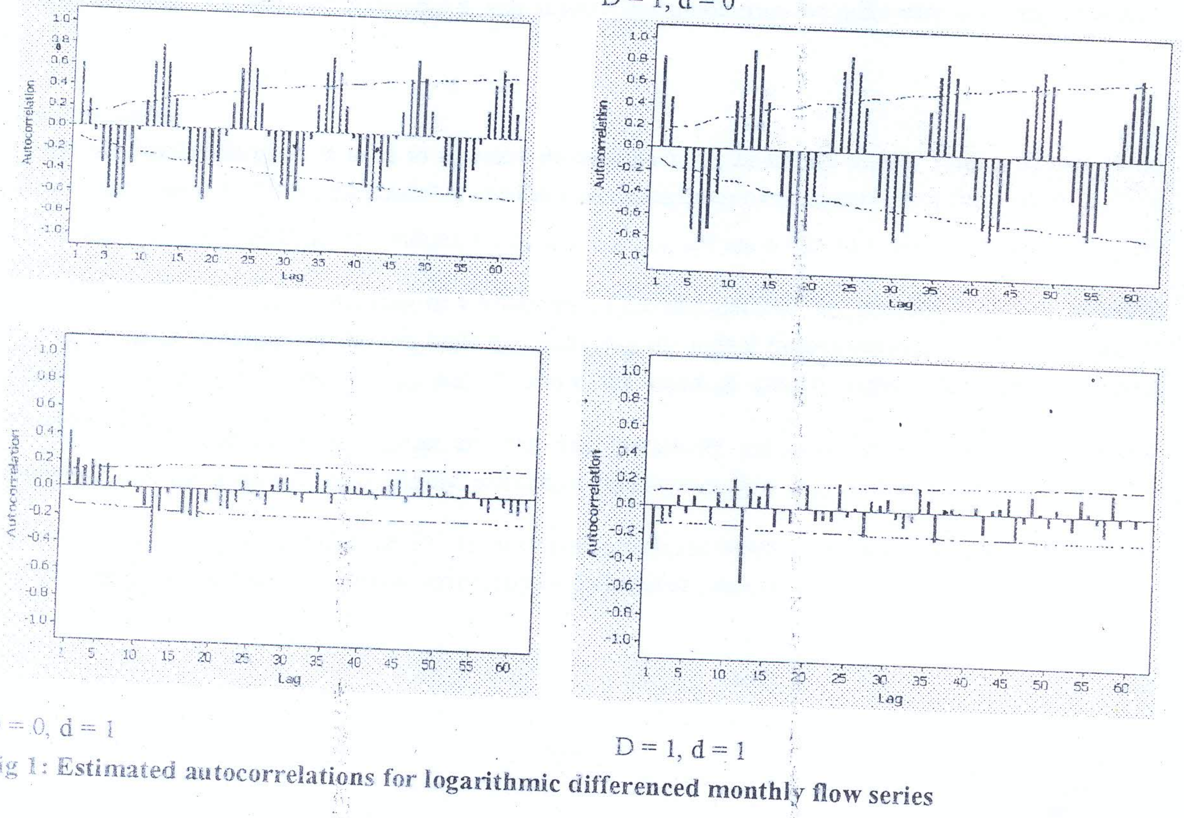
Application of the Thomas-Fiering model

In tropical countries like Nigeria, there can be practically no flow during the dry season in a stream. To generate synthetic discharge sequences for this case, the Thomas-Fiering model was used with the following modifications: Calculate mean and standard deviation of each monthly flow. Find out the correlation coefficients between all the successive months. If N is the number of years of data available, then let n_j be the number of years out of N for which the flow is available for the months. To illustrate, if 20 years of data are available then for June, say let there be 18 years for which flow is available. The value of $n_j = n_6 = 18$. Here, j is taken as 6 for June. Thus, calculate $p_j = n_j/N$ for all the months. Fit the Thomas-Fiering model to these successive pairs of months. The synthetic sequences of monthly flows are as follows. For month j, choose a pseudo-random number rectangularly distributed over (0, 1). The pseudo-random numbers generated are available in statistical handbooks. These numbers are continuously matched with the p_j values for all months. For the month of say July, the value of p_j is always 1; therefore flow is definitely to be generated for this month. For any month if the number is less than pj but greater than zero, then flow is to occur in the month j, otherwise no flow is to occur. If no flow is to occur in the month j, then generation of the flow for the month may not be carried out. If flow is to occur in the month j and flow also has already occurred in the month j-1, then use the regression equation of Thomas-Fiering model to obtain the flow for the month j. Sometimes negative values are generated while using Thomas-Fiering model. Where negative values are encountered, it is recommended that these values should be retained and used to derive the subsequent values in the sequence. Once the generated sequence is completed, all the negative values in the generated sequence should be replaced by zero. For the specific detail of the application here, the monthly streamflows were logarithmically transformed to overcome the occurrence of negative values in the generated sequence. In addition too, since there is flow all year round in the Benue River, article 6 was ignored.

(ii) ARIMA Analysis of the monthly flow data

D = 0, d = 0

D = 1, d = 0



D = 0, d = 1

D = 1, d = 1

Fig 1: Estimated autocorrelations for logarithmic differenced monthly flow series

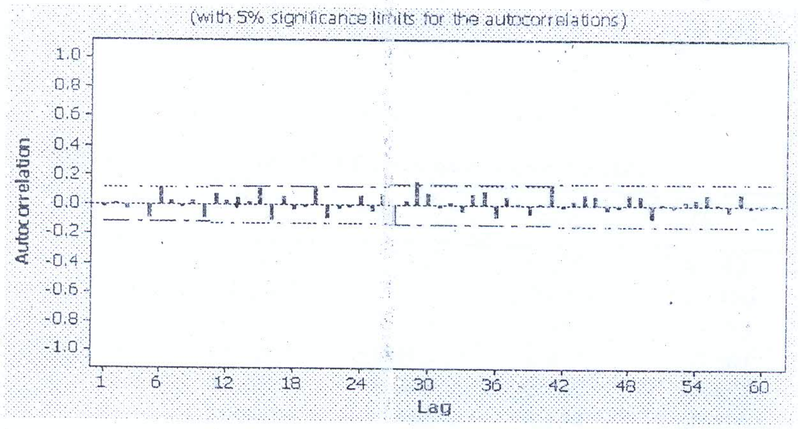


Fig 2: Residual autocorrelation function for ARIMA $(1,0,2)\times(1,1,1)_{12}$

To be able to identify the most suitable model to fit the flow series, serial correlations were calculated for possible differencing schemes d = 0, 1, 2 and D = 0, 1, 2, where d and D stand for non-seasonal and seasonal differencing, respectively. Fig. 1 shows the autocorrelation function plots for these differencing schemes. To account for runoff phenomenon in the streamflow data, the prospect of seasonal differencing seem more promising or more likeable since seasonality cannot really be accounted for by non-seasonal differencing, nor is an integrated moving average scheme expected to account for the non-seasonal autoregressive behaviour. Thus considering this factor, a multiplicative ARIMA model $(1,0,2)\times(1,1,1)_{12}$ was examined. This model has the form

$$(1-\phi_1 B)(1-\phi_{1,12}B^{12})z_t=(1-\Theta_{1,12}B^{12})(1-\theta_1 B-\theta_2 B^2)a_t \qquad 4$$

where, ϕ_i, Θ_i, and θ_i stand for non-seasonal autoregressive, seasonal autoregressive, and seasonal moving average parameters respectively; whereas z_t and a_t are logarithmic transformed series and model random shocks.

Flow forecasting

The ARIMA model forecast is used to forecast flows for one to 24-month ahead. With reference to an origin at time t (here, t = 288), the model is used to make minimum mean square error forecasts of z_{t+L} for $L\geq 1$, where L is the lead time; the values forecasted for z_{t+L} for an origin at t with lead time L will be written as $\hat{z}_t(L)$. Diagnostic checking or verification of the adequacy of the model is done by evaluating the autocorrelation function for the residuals by modifying the model to take account of any non-random features. Figure 2 and Tables 1, 2, and 3, show the residual autocorrelation function for model $(1,0,2)\times(1,1,1)_{12}$, parameter estimation and the final parameter values, as well as the corresponding diagnostic check for adequacy, respectively as fitted to the 26 years flow data. Values of the Q-test statistic compared with the value of χ^2 at the 5% level is not significant whereas the autocorrelation plot of the model residual reflects that the residual series may be considered random.

Table 1: Estimation of ARIMA model parameters

Iteration	SSE	ϕ_1	$\phi_{1,12}$**	θ_1	θ_2	$\Theta_{1,12}$**
0	64.5264	0.100	0.100	0.100	0.100	0.100
1	54.8937	0.012	0.055	-0.050	0.089	0.145
2	50.5089	-0.056	0.152	-0.151	0.071	0.295
3	45.4580	-0.162	0.171	-0.301	0.037	0.386
4	42.0178	-0.275	0.149	-0.451	-0.007	0.424
5	39.6586	-0.394	0.124	-0.601	-0.055	0.447
6	36.2933	-0.244	0.075	-0.507	-0.071	0.479
7	36.1324	-0.094	0.072	-0.359	-0.035	0.480
8	35.9804	0.056	0.069	-0.212	0.003	0.481
9	35.8173	0.206	0.066	-0.064	0.041	0.482
10	35.6239	0.356	0.063	-0.083	0.081	0.484
11	35.3723	0.506	0.059	0.231	0.120	0.486
12	34.9998	0.656	0.054	0.377	0.161	0.488
13	34.2944	0.806	0.043	0.521	0.204	0.494
14	32.1154	0.956	-0.002	0.642	0.253	0.519
15	28.0190	0.943	-0.102	0.544	0.269	0.699
16	27.2837	0.946	-0.181	0.486	0.266	0.700
17	27.2450	0.954	-0.179	0.483	0.270	0.719
18	27.2367	0.957	-0.173	0.484	0.274	0.730
19	27.2331	0.958	-0.168	0.485	0.276	0.738
20	27.2317	0.959	-0.166	0.486	0.277	0.742
21	27.2317	0.960	-0.166	0.486	0.278	0.743
22	27.2312	0.960	-0.166	0.486	0.278	0.745

** $\phi_{1,12}$ is seasonal autoregressive parameter; $\Theta_{1,12}$ is seasonal moving average parameter

Table 2: Final model parameter estimates

Parameters			Statistics	
Type	Coef	SE Coef	T	P
AR 1	0.9600	0.0284	33.83	0.000
SAR 12**	-0.1657	0.0748	-2.22	0.028
MA1	0.4859	0.0667	7.29	0.000
MA 2	0.2782	0.0646	4.31	0.000
SMA 12*	0.7447	0.0544	13.69	0.000
Constant	0.000318	0.001155	0.28	0.783

** Seasonal autoregressive parameter, * seasonal moving average parameter

Table 3: Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	10.2	29.0	46.3	60.4
Critical value	12.6	28.9	43.8	58.1
DF	6	18	30	42
p-Value	0.115	0.044	0.011	0.021

In terms of the forecasting function, the general ARIMA model can be written in three alternative forms: as a difference equation; as an infinite sum of the current and weighted previous values of shocks at; and as an infinite sum of weighted previous observations plus the current value of at. Conditional expectation of any of these forms supplies a forecasting function; in this regard, the difference equation is used here.

By recalling that $z_t(L)=[z_{t+L}]$, using square brackets to signify conditional expectation, noting that

$$[z_{t-j}]=z_{t-j} \qquad j=0,1,2,\cdots$$
$$[z_{t+1}]=\hat{z}_t(j) \qquad j=1,2,\cdots$$
$$[a_{t-j}]=a_{t-1}=z_{t-j}-\hat{z}_{t-j-1}(1) \qquad j=0,1,2,\cdots$$
$$[a_{t+j}]=0 \qquad j=1,2,\cdots \qquad 5$$

and taking expectation of the model which has the general form $(1-\phi_1 B)(1-\phi_{1,12}B^{12})z_t=(1-\Theta_{1,12}B^{12})(1-\theta_1 B-\theta_2 B^2)a_t$, the forecasting function can be obtained thus

$$z_{t+L}=\phi_1 z_{t+L-1}+\phi_{1,12}z_{t+L-12}-\phi_1\phi_{1,12}z_{t+L-13}+a_{t+L}-\theta_1 a_{t+L-1}-\theta_2 a_{t+L-2}-\Theta_{1,12}a_{t+L-12}$$
$$+\theta_1\Theta_{1,12}a_{t+L-13}+\theta_2\Theta_{1,12}a_{t+L-14} \qquad 6$$

This can be expanded for the respective lead time (L) to make the forecasts with z_{t+L} as the dependent forecast variable as a function of L.

(iii) Forecasts from the Thomas-Fiering and ARIMA models

Both models were used to make forecasts of the monthly flow series, and subsequently, the forecasts from the models are compared with the flows that actually occurred; in doing this, the final 2 years data are used for the comparison. Because the last 2 years flow data is used for the comparison, the parameters are re-estimated for both models for the entire flow series shortened by 2 years (i.e., the model fit is done with 26 years of flow data). The flow forecasts are considered from the aspect of choosing a particular time origin and taking cognisance of the behaviour of the forecast function as the lead time L increases; that is, the long-term behaviour of the forecast function should be a useful theoretical check on the fit of a model. Taking the origin t = 288, forecasts for the logarithms of the flow were made using both models. Figure 3 shows the behaviour of the ARIMA model forecast function; the forecasts are quite close to the monthly means. Baring data quality problems, stationarity issues, and model over-fitting, intuitively, for an ideal forecast function this behaviour is to be expected; forecasts in the distant future for a trend-free series should be the unconditional estimates of the means. Considering Fig.3, the forecasts are all within the bounds with respect to the actual flows; based on this, and considering the data size for model fitting, the ARIMA model has reproduced the monthly means well.

Figure 4 illustrates the standard errors of the forecasts. This figure compares the monthly standard deviations of the logarithms of the monthly flow with the standard errors for forecasts of the two models under discourse, respectively, for $1 \le L \le 24$. As L becomes large (say, greater than 4), the standard error of a forecast for Thomas-Fiering model, tends closely to that of the historic flow, whereas the ARIMA model deviates away significantly. The behaviour of the Thomas-Fiering model in this regard is further explained by Figure 5, where it is used to simulate the flow regime for 26 years; it was able to reproduce the flow dynamics clearly well, this reinforces its suitability to be used for long-term flow forecasting of the Benue River. The failure of the ARIMA model to account for the seasonal pattern in the standard deviations is a major limitation of the model. In particular, it leads to problems in transforming forecasted flow logarithms into forecasted flows

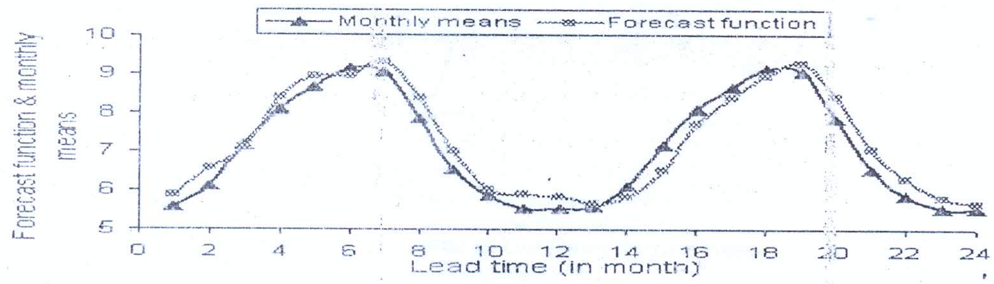


Fig 3: Forecasting for flow logarithms of $ARIMA(1,0,2)\times(1,1,1)_{12}$ model

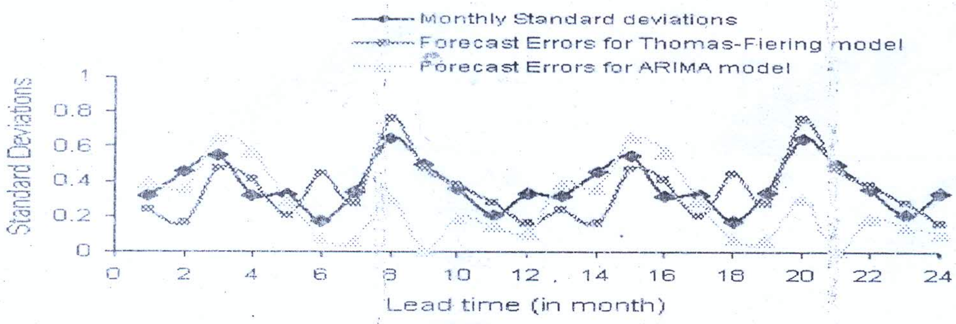


Fig 4: Monthly standard deviation of logarithms and forecast errors for Thomas-Fiering and $ARIMA(1,0,2)\times(1,1,1)_{12}$ models.

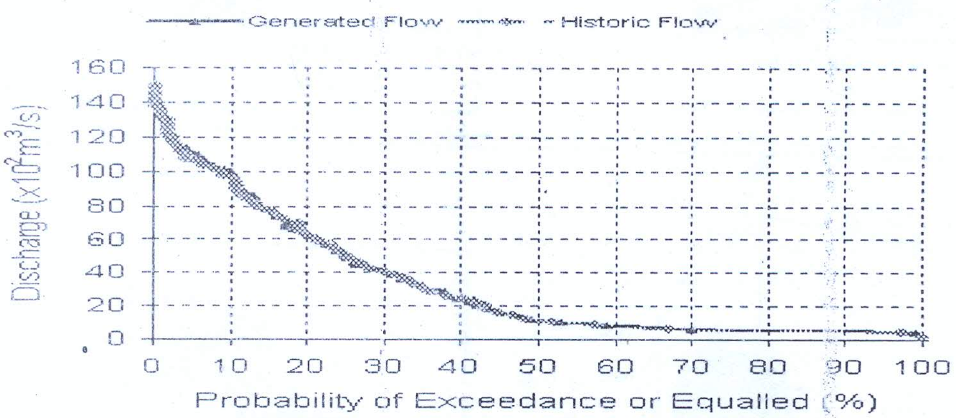


Fig 5: Long – term flow duration curve based on simulations (i.e., Thomas-Fiering model)

(c) Nonlinear Determinism

The first step in the search for a deterministic behaviour is to attempt to reconstruct the dynamics in phase space. The delay time method as proposed by Takens (1981) and Packard *et al.* (1980) was employed; this was done by analysing the autocorrelation structure as in Fig. 6 while Fig. 7 shows the attractor diagram or phase-space constructs.

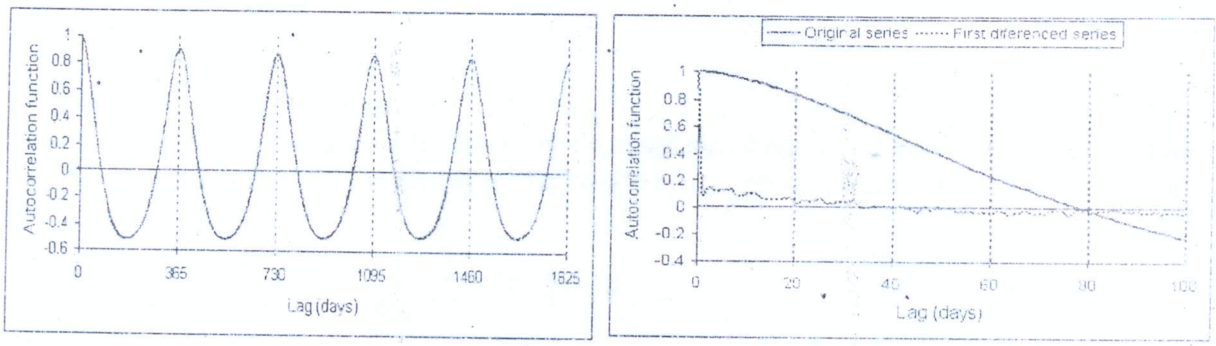


Fig. 6: Autocorrelation Functions of the Daily Time Series

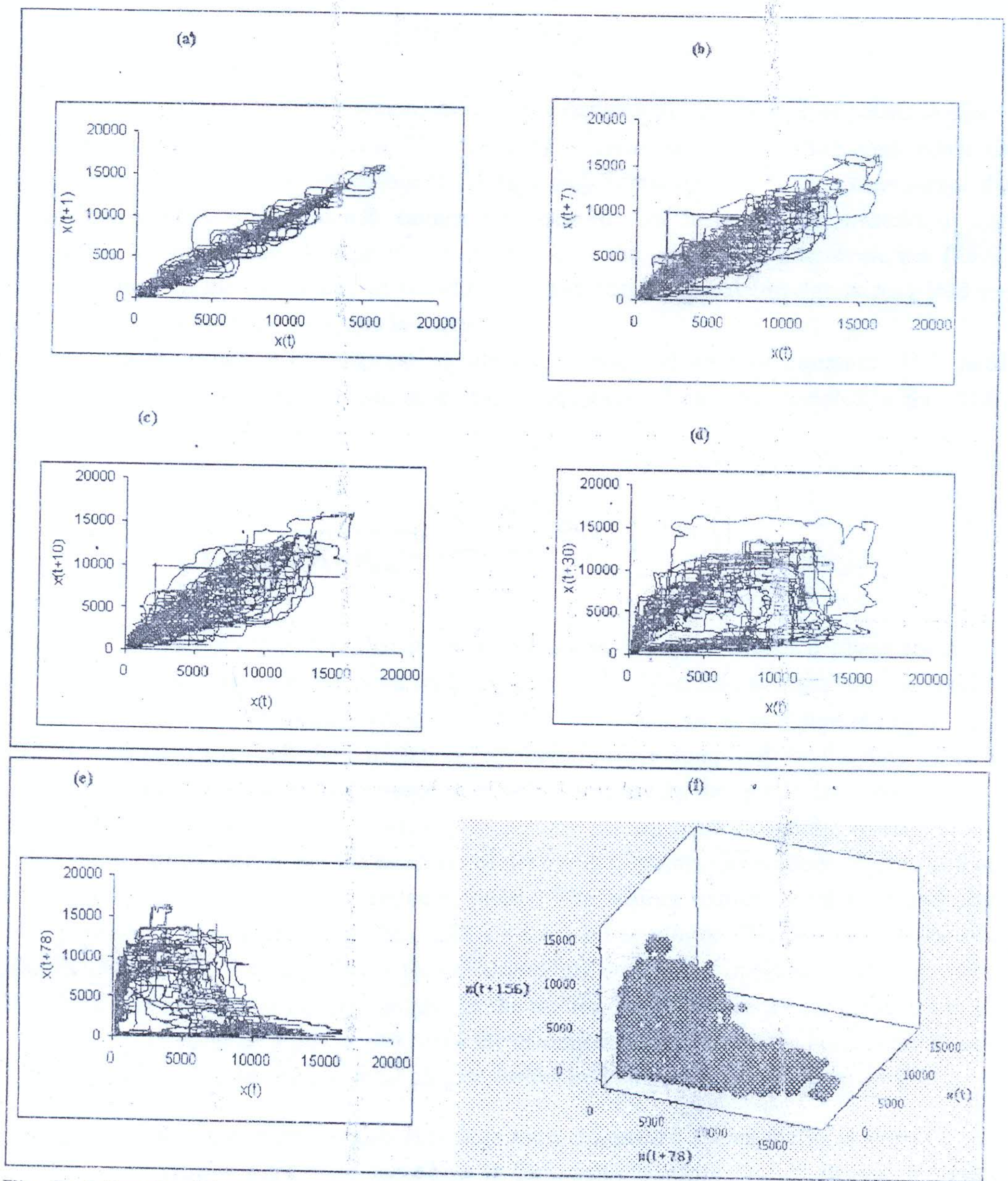


Fig. 7: Attractor Diagrams

(i) **Estimation of Correlation Dimension**

The most commonly used algorithm for computing the correlation dimension is the Grassberger-Procaccia algorithm (Grassberger and Procaccia, 1983a). For m-dimensional phase-space, the basic formula is given by

$$C(r)=\frac{2}{N(N-1)}\sum_{i=1}^{N}\sum_{j=i+1}^{N}\Theta\left(r-\left\|x_i-x_j\right\|\right) \qquad 7$$

subject to the constraints:

$$\begin{cases}\Theta(x)=0, & \text{if } x\leq 0\\ \Theta(x)=1, & \text{if } x>0\end{cases}$$

where, $\Theta(x)$ is the Heaviside step function, and r is the radius between the pair of points in phase space. The straightforward estimator, Equation (7), is biased towards too small dimensions when the pairs entering the sum are statistically independent (Holger and Schreiber, 1997). For time series data with nonzero autocorrelations, independence cannot be assumed; for instance, the embedding vectors at successive times are often also close in phase space due to the continuous time evolution [ibid], called temporal correlation. In the calculation of correlation dimension, temporal correlation may lead to serious underestimation unless the necessary care is taken.

In order to avoid the problem of temporal correlation, a modified form of Equation (7) is used in the computation of the correlation dimension, as given by Equation (8) and implemented in the TISEAN 3.0 Software package.

$$C(r)=\frac{2}{\left(N-n_{\min}\right)\left(N-n_{\min}-1\right)}\sum_{i=1}^{N}\sum_{j=1+n_{\min}}^{N}\Theta\left(r-\left\|x_i-x_j\right\|\right) \qquad 8$$

where, n_{min} is a threshold value such that pairs of vectors in the m-dimensional phase space which are closer in time than it are discarded to avoid temporal correlations that may contaminate the result. For the implementation of Equation (8), n_{min} is set to 182 for the daily flow series as suggested by Wang (2006). The correlation integral C(r) and the correlation exponent D were computed for the data using different values of the delay time in order to be disposed to examine objectively the appropriate delay time.

For a finite data set, there is a radius below which there are no pairs of points, whereas at the other extreme, when the radius approaches the diameter of the cloud of points, the number of pairs will no longer increase as the radius increases (i.e., saturation point). The scaling region would be found somewhere between depopulation and saturation. When lnC(r) versus lnr is plotted (Fig. 8) for a given embedding dimension m whereas Fig. 9 clearly shows the corresponding correlation integrals. The range of lnr where the slope of the curve is approximately constant is the scaling region where fractal geometry is indicated. In this region, C(r) increase as a power of r, with the scaling exponent being the correlation dimension D. To vividly display the scaling region, local slopes should be computed. Local slopes of lnC(r) versus lnr can indeed show clearly the scaling region if it does exist. Figure 10 shows the local slopes Δ_i computed for this case; computations were done according to Equation (9) while Fig. 11 depicts the relationship between correlation exponent and embedding dimension for both the original and processed flow data.

$$\Delta_i=\frac{\log\left[C(r)\right]_{i+1}-\log\left[C(r)\right]_{i-1}}{\log(r)_{i+1}-\log(r)_{i-1}} \qquad 9$$

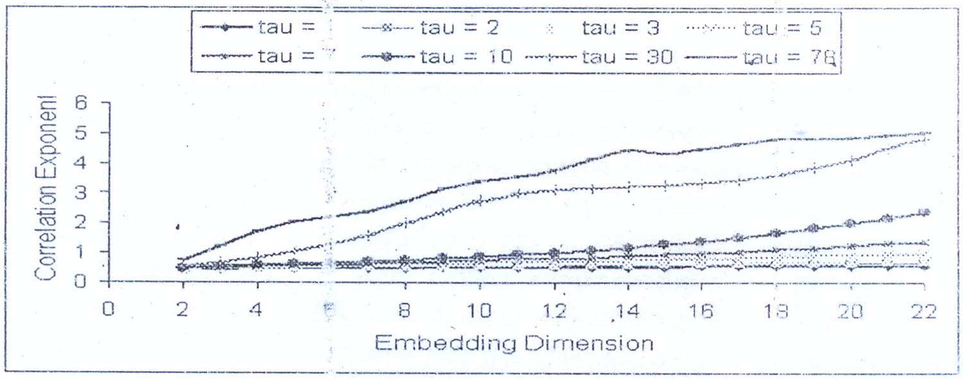


Fig 8: Correlation Exponent versus Embedding Dimension for various delay time (tau) values

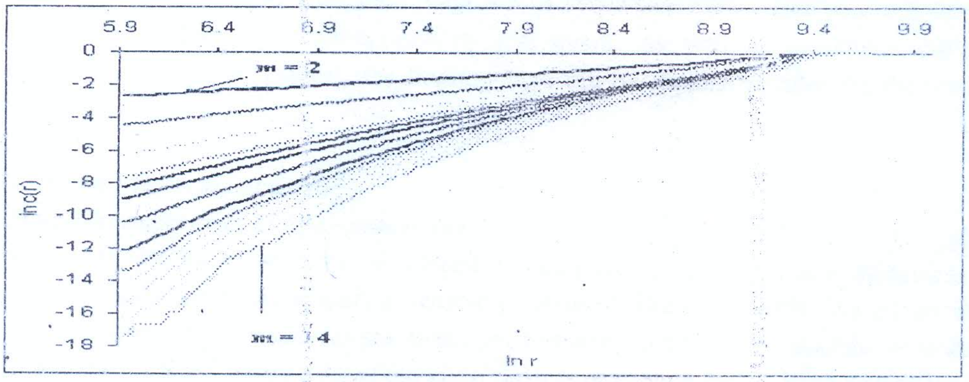


Fig 9: Correlation integrals for Average daily discharge

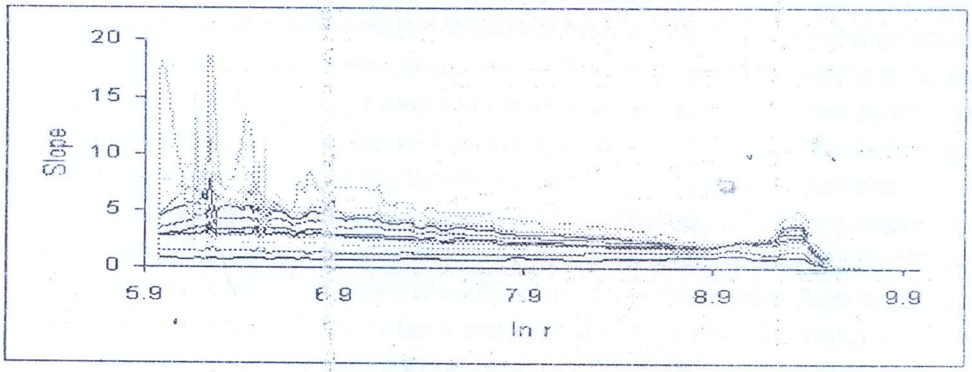


Fig 10: Correlation integral slopes for m between 2 and 14

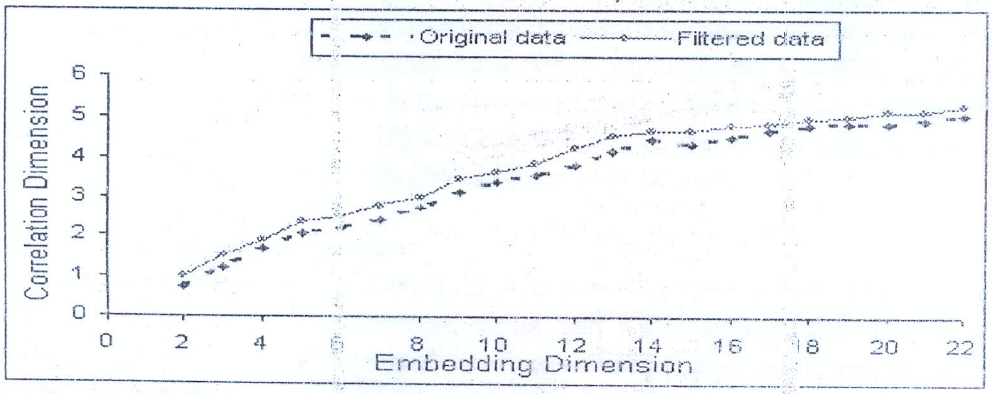


Fig 11: Relationship between correlation exponent and embedding dimension

Findings and Implications for Water Resource Planning and Management

There are quite a host of components to water resources planning and management in the overall with their respective appurtenant issues. In this light, the issue of systemic interactions cannot be ignored or downplayed at all. This is so because, one alteration to the natural functional dynamics of a particular component produces different corresponding effects on one or more other aspects of the riparian system. Hence to be able to understand the nature of these interactions majorly, modelling and simulation of some parts of the overall system is usually carried out; the objective basically is to enhance our understanding of the workings of the system writ-large for purposes of proactive actions and management of the system in the overall in a futuristic manner. To this end, usually both physical and process based data are extensively mobilised; but more often than not, these data are of secondary nature and thus not devoid of basic hydrologic problems like randomness, quantity and quality as well as continuity and homogeneity problems, etc. Hence, in this context, the findings here can be appraised under the themes as presented below.

Stochasticity and Heteroscedasticity

(a) Data Characteristics and pre-processing

In hydrologic modelling of systems of interest for purposes of planning and management of water resources, historical records are usually extensively collected. But statistically, the historical record is a sample out of a population of natural processes; for instance, streamflow. Thus, the generated sequence are representative of likely time series of the given process or phenomenon. Using streamflow as a case of analogy in this context, streamflow being a natural phenomenon has a random component, though not fully random since it has been observed that it exhibits heteroscedastic behavioural pattern (Otache *et al.*, 2012). Considering this, data length, quality as well as consistency issues become paramount; for instance, in the face of acute dearth of long and continuous data availability, can realistic generalisations be made from forecasting the dynamics of a particular river flow system. In the same twist, considering the complex nature of river flow and the significant variability it exhibits in both time and space, what is the appropriateness of employing stochastic approaches for modelling it? These portend serious issues for effective planning and management of water resources. Intertwined in all this, for effective stochastic modelling, usually, stationarity conditions are adhered to; doing this, entails some levels or forms of data pre-processing but not without its associated problems for water resource managers. Linear stochastic models like SARIMA, deseasonalised ARMA, and periodic models like Markov models are the most commonly used in modelling hydrological processes. But considering model formulation in terms of constitutive equations, though the seasonal variation in the variance present in the original time series can be dealt with basically by the deseasonalisation approach, neither of the SARIMA and ARMA models take into account the seasonal variation in the variance of the residual series since the innovation series is assumed to be independent and identically distributed. Generally, the pre-processing strategy irrespective of the type to a large extent distorts the original spectrum of the time series.

(b) Non-determinate nature of models

Following the line of thought in (a), forecasting river flow in general or after heavy rainfall events is important for public safety, environmental issues and water management. For these purposes, mathematical models have been developed based either on physical considerations or on statistical analysis. But the non-determinate nature of models poses critical issues for water resource managers; the non-determinate nature of model structure, for instance, streamflow models, can be really be appreciated in a wider context considering the fact that streamflow is usually treated as a random process, purely stochastic

(Otache *et al.*, 2011b). The justification is that both streamflow or river flow is a function of precipitation and other processes which, at present level of knowledge, seem to evolve randomly in time and space.

(c) Conditional heteroscedasticity

When modelling hydrologic time series, the focus usually is on modelling and predicting the mean behaviour, or the first order moments, and rarely, is it concerned with the conditional variance, or their second order moments; although unconditional season-dependent variances are usually considered. The increased importance played by risk and uncertainty considerations in water resources management and flood control practice, as well as in modern hydrology theory, however, has called for the development of new time series techniques that allow for the modelling of time varying variances. It is not hard to find evidence to argue that a time series with random appearance might be nonlinear dynamic; but, the difficulty is in telling what kind of nonlinear dynamics; especially, issues of long memory, non-normality and heteroscedasticity or volatility. This aspect is paramount as it takes into account excess kurtosis (i.e., fat tail behaviour), which is very common in hydrologic processes, and volatility clustering. Volatility clustering depicts a phenomenon in which large changes tend to follow large changes, and small changes tend to follow small changes. In either of these cases, the changes from one period to the next are typically of unpredictable sign. To address this time varying phenomenon is not out rightly easy and thus calls for the development of Generalised Autoregressive Conditional Heteroscedasticity type-models for real time modelling and simulation.

Nonlinearity and Nonlinear Determinism

(a) Nonlinearity

The physical mechanisms governing river flow or streamflow dynamics are many, and act on a huge range of temporal and spatial scales and thus pose a lot of concerns for water resource planners and managers in real time. Although not all components are complex in themselves, the vastness of the space-time domain, the number of the processes involved, and the fact that almost all of them present some degree of nonlinearity, make the problem of river flow or streamflow formation highly non-trivial (Amilcare and Ridolfi, 1997). To model such a complex system, one hopes that only few of the various mechanisms become prevalent in the different phases of the process, so that the system dynamics undergoes a simplification due to a reduction in the number of the effective degrees of freedom (Amilcare and Ridolfi, 2003). Therefore, it is important to attempt qualitative determination of the degree of linearity of particular catchments, because they permit an assessment of the range of approximation likely to be obtained from the application of either linear or low-order nonlinear methods for the postulation of inflow-outflow relationships (Otache *et al.*, 2011a). Generally, a reliable detection of nonlinearity can be essential information in addressing modelling efforts. Though the evolving paradigm points to the acceptance of the fact that many natural systems are nonlinear with feedbacks over many space and timescales, however, certain aspects of these systems may be less nonlinear than others and the nature of nonlinearity may not always be that clear; this is an aspect that water resource planners and managers who are interested in real time simulation of the river flow or streamflow system should keenly address. To be able to draw any objective conclusions on analysis aimed at detecting nonlinearity and nonlinear determinism, it is relevant to trace its course and causes, especially why and where deterministic components or nonlinear dynamics are to be expected in river-flow time series. The first source of determinism is related to climate dynamics that produces the input of the rainfall-runoff transformation and also determines in many ways the state of the basin with a sort of "parametric forcing" on vegetation cover, soil saturation, etc. Considering these facts though, it is imperative to note that even if the actual dynamics of atmosphere at a meteorological time scale is more likely of the kind of Spatio-temporal chaos of a high dimension, the coupling with the basin with its various feedbacks could also give rise to recurrent low-dimensional components, as reported by Amilcare and Ridolfi (2003). In view of the complex nature and dimension of the subject of nonlinearity and the corresponding issues associated with it, to explore all its different kinds will be a treacherous task.

(b) Nonlinear Determinism

One aspect which hydrologists have been extensively working on is the structure of hydrological processes, such as rainfall and runoff (Sivakumar, 2000). Even though, during the past decades, a number of mathematical models have been proposed for modelling hydrological processes, there is, however, no unified mathematical approach. In part, this difficulty stems from the fact that hydrological processes exhibit considerable and temporal variability; by extension, another part of this difficulty is due to the limitation in the availability of 'appropriate' mathematical tools to exploit the structure underlying the hydrological processes (Sivakumar, 2000). The tremendous spatial and temporal variability of the hydrological processes has been believed, until recently, to be due to the influence of a large number of variables. Consequently, the majority of the previous investigations on modelling hydrological processes have essentially employed the concept of a stochastic process. However, recent studies have indicated that even simple deterministic systems, influenced by a few nonlinear interdependent variables, might give rise to very complicated structures (i.e., deterministic chaos (Sivakumar, 2000). Therefore, it is now believed that the dynamic structures of the seemingly hydrological processes, such as rainfall and runoff, might be better understood using nonlinear deterministic chaotic models than the stochastic ones. Despite this posturing though, it is highly improbable that complex natural phenomena may be controlled only by the presence of a chaotic dynamics; rather, if it is present, it most likely coexists with other types (Amilcare and Ridolfi, 1997). It is therefore a weaker form of determinism which may be hidden in complex natural systems (Otache *et al.*, 2011a).

Conclusions

For purposes of effective water resource planning and management, an integrated forecasting framework is essential. But it suffices to note that the appropriateness of the stochastic process for every flow series may be debated in the context of nonlinear determinism and chaos; this holds true considering the fact that seemingly complex and irregular behaviours could be the outcome of simple deterministic systems with only a few nonlinear interdependent variables with sensitive dependence on initial conditions. In the overall, nonlinear deterministic methods could be viable complement to the linear stochastic ones for analysing the dynamics of river flow system. Despite this though, it is imperative to note that the study of the dynamics of river flow, conducted in the light of chaos theory may have conflicting results; some are more of a speculative character whereas others may have practical potentials. Since it is obvious from available literature that discussion on the nonlinear mechanism, vis-a-vis conditional heteroscedasticity in hydrological processes has been very austere, modelling data with time varying conditional variance could be attempted in various ways using either the non-parametric or parametric approach in the context of integrated modelling and forecasting of hydrologic processes.

References

Amilcare, P., and Ridolfi. (1997). "Nonlinear Analysis of River Flow Time Series", *Water Resources Research*, Vol. 33(6), pp. 1353-1367

Amilcare, P., and Ridolfi, L. (2003). "Detecting determinism and nonlinearity in river-flow time series", *Journal of Hydrological Sciences*, 48(5), pp: 763-780

Grassberger, P., and Procaccia, I. (1983a). "Measuring the strangeness of strange attractors", *Physica D, 9,* pp: 189-208

Holger, K., and Schreiber, T. (1997). "Nonlinear Time Series Analysis", *Cambridge University Press*, England; pp: 42-86

Otache, Y. M., Mohammed, A S., and Ahaneku, I. E. (2011a). Nonlinear Deterministic Chaos in Benue River Flow Daily Time Series, *Journal of Water Resource and Protection*, Vol. 3(10), pp: 747-757; http://www.SciRP org/journal/jwarp. doi; 10.4236/jwarp, 310085

Otache, Y. M., Mohammed, A.S., and Ahaneku, I. E. (2011b). Parametric Linear Stochastic Modelling of Benue River Flow Process, Open Journal of Marine Science, Vol. 3, pp: 1-9; http://www.SciRP.org/journal/ojms, doi; 10.4236/ojmh.2011. 13008.

Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S. (1980). "Geometry from a Time Series", *Physical Review Letters*, Vol. 45(9), pp: 712-716

Sivakumar, B. (2000). "Chaos Theory in Hydrology: Important Issues and Interpretation", *Journal of Hydrology*, Vol. 227(1-4), pp: 1-20

Takens, F. (1981). "Detecting Strange Attractors in Turbulence", In: Rand, D. A., and Young, L. S. Eds., *Lecture Notes in Mathematics*, Vol. 898, pp: 366-381

Wang, W. (2006). "Stochasticity, nonlinearity and forecasting of streamflow processes". Deft University Press, Amsterdam – Netherlands, ISBN 1-58603-621-1; pp: 1-17

Zealand, C. M., Burn, H. D., and Simonovic, S. P. (1999). "Short Term Forecasting Using Artificial Neural Networks", *Journal of Hydrology*, Vol. 214(1-4), pp: 32-48