



Journal Homepage: - www.journalijar.com
**INTERNATIONAL JOURNAL OF
 ADVANCED RESEARCH (IJAR)**

Article DOI: 10.21474/IJAR01/1466
 DOI URL: <http://dx.doi.org/10.21474/IJAR01/1466>



RESEARCH ARTICLE

MODELLING THE TREND OF FLOWS WITH RESPECT TO RAINFALL VARIABILITY USING VECTOR AUTOREGRESSION.

Wahab A. Iddrisu*, Kaku S. Nokoe and Isaac Akoto.

Department of Mathematics and Statistics, University of Energy and Natural Resources, P.O. Box 214, Sunyani, Ghana.

Manuscript Info

Manuscript History

Received: 12 July 2016
 Final Accepted: 19 August 2016
 Published: 27 September 2016

Key words:-

Rainfall, Flows, Vector Autoregressive (VAR) models, Forecasting, Bui

Abstract

Vector Autoregressive (VAR) models for multivariate time series have been extensively developed and used in econometrics. However, use of VAR models outside of these areas is rather limited. This paper models the trend of flows at Bui with respect to rainfall variability and investigates whether there is any causality (Granger or instantaneous) between rainfall and flow in relation to forecasting. Monthly data on rainfall and flows from January 1954 to December 2005 was obtained from Bui Power Authority in Ghana. Both unrestricted and Bayesian VARs were estimated and compared in order to select the best VAR model for forecasting and structural analysis. The results showed that the unrestricted VAR model outperforms the Bayesian VAR model in terms of forecasting flows and rainfall at Bui. Results from the structural analysis also revealed a two-way causality from rainfall to flow and vice-versa. We conclude that modelling flows and rainfall together at Bui improves the forecasting of both of them.

Copy Right, IJAR, 2016., All rights reserved.

Introduction:-

Rainfall and stream flow are key components of the hydrologic cycle consisting of precipitation, evaporation, evapotranspiration, infiltration, precipitation and surface water flow. Air humidity and temperature are very influential at every stage of the cycle. When temperatures are relatively high there is likely to be evaporation of surface water and water vapour would be formed which in certain altitude would form a core of condensation to form clouds. The clouds would then drop to the ground in the form of rain, snow, dew, fog, etc due to the influence of low temperatures and relatively immense droplets (Raghunath, 2006). Climate variability is likely and expected to impact on some of these hydrological processes including evapotranspiration, water temperature, stream flow, soil moisture, timing and magnitude of runoff, and frequency and severity of floods, which may lead to changes in environmental variables (Nijssen et al., 2001; Zhang et al., 2005).

Despite the fact that stocks of water in natural and artificial reservoirs are helpful to increase the available water resources for human society, the main focus in water resources assessments should be on the flow of water (Oki & Kanae, 2006). Rainfall and other hydrological data in Ghana generally occur with quite some temporal variability. The Northern regions are more affected by prolonged droughts while the Southern regions including the Brong-Ahafo region experience abundant rainfall that at times cause erosion and floods. Giving the peculiar location of the

Corresponding Author:-Wahab A. Iddrisu.

Address:- Department of Mathematics and Statistics, University of Energy and Natural Resources, P.O. Box 214. Sunyani. Ghana.

Bui dam (at the border of the Northern Region (Bole) and Brong-Ahafo Region (Wenchi)), quantifying changes of water flows in time is fundamental in addressing issues of power generation and survival of ecosystem downstream.

A number of studies have been conducted the world over concerning rainfall and flow modelling. Examples include; Top-down and data-based mechanistic modelling of rainfall - flow dynamics at the catchment scale (Young, 2003); Simulating rainfall and river flow dynamics in Ghana (Ampadu, 2007); Statistical modelling of rainfall and river flow in Thailand (Boochabun et al., 2004); Comparative analysis of several conceptual rainfall-runoff models (Franchini & Pacciani, 1991) and Rainfall-runoff modelling and water resources assessment in north western Ivory Coast (Servat & Dezetter, 1993).

In Ghana, Awotwi et al. (2015) conducted a study on the prediction of hydrologic response to climate change in the White Volta catchment in West Africa. An ensemble of Regional Climate Model (REMO) was used to simulate and project the climate at local scale in order to investigate the hydrological impact of possible future climate change in White Volta Catchment (West Africa). They found that, with a small increase of 8% and 1.7% of the mean annual precipitation and temperature respectively, annual surface runoff, annual base flow and evapotranspiration recorded increment of 26%, 24% and 6% respectively. In a study conducted by the CSIR Water Research Institute (CSIR – WRI) in 2000 (CSIR & WRI, 2000), under the United Nations Framework Convention on Climate Change (UNFCCC) and coordinated by the Environmental Protection Agency (EPA), it was found that there is a decrease and highly variable rainfall pattern, and frequent and pronounced dry spells in Ghana. The study also estimated a general reduction in annual river flows in Ghana by 15-20 % for the year 2020 and 30-40 % for the year 2050.

VAR models have been extensively developed and used in econometrics. Park (1990) used a set of rigorous diagnostic techniques to evaluate the forecasting performance of some multivariate time-series models for the U.S. cattle sector, including Bayesian vector autoregression (BVAR), unrestricted vector autoregression (UVAR), restricted vector autoregression (RVAR), and vector autoregressive moving-average (VARMA) models. He found using the root-mean-squared-error criterion along with an evaluation of the rankings of forecast errors that the Bayesian vector autoregression (BVAR) and the unrestricted VAR (UVAR) models generate forecasts which are superior to both a restricted VAR (RVAR) and a vector autoregressive moving-average (VARMA) model. Bessler (1984) used VAR models to study agricultural prices, money supply and industrial prices in Brazil; Kaylen (1988) used VAR together with other models to forecast the U.S Hog market; McCarty & Schmidt (1997) used VAR models to study State-Government expenditure; Andersson (2007) compared the forecast performance of random walk (RW), autoregressive (AR) and VAR models to forecast Swedish real GDP growth in which VAR emerged as the best in terms of forecasting.

In relation to natural phenomena, especially rainfall forecasting, use of VAR models has been rather scarce and virtually nonexistent in hydrological modelling. Saputro et al. (2011) studied the correlation between rainfall in a region and rainfall in other nearby regions. Adenomon et al. (2013) examined the dynamic relationship between rainfall and temperature in Niger State, Nigeria using monthly data from January 1981 to December 2010 and found a bi-directional causation from rainfall to temperature and from temperature to rainfall. Nugroho et al. (2014) applied VAR models to forecast future rainfall in Semarang, Central Java, Indonesia using monthly data on rainfall, humidity and temperature from 2001 to 2013 collected from 5 measurement stations. The VAR method was found to perform better than ARIMA method regarding forecasting as having smaller Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) and was also quite accurate in forecasting rainfall in their study area.

Studies relating to rainfall and flow forecasting in Ghana are mostly conducted using univariate time series models. Some of the recent ones include; Abdul-Aziz et al. (2013) who examined rainfall patterns over time, from 1974 to 2010, in the Ashanti region of Ghana using Seasonal Autoregressive Integrated Moving Average (SARIMA). Ampaw et al. (2013) developed a time series model for predicting rainfall in the New Juaben Municipality of the Eastern region of Ghana. Ampadu et al. (2013) presented a review of some of the approaches employed in rainfall-riverflow modelling highlighting on the rationale and structure of the modelling approaches, their strengths and weaknesses which may assist in making an informed choice of a modelling approach for hydrological studies. In this paper, vector autoregression (VAR) for multivariate time series is used to model the trend of water flow at Bui with respect to rainfall variability and investigates whether there is any causality (Granger or instantaneous) between rainfall and flow in relation to forecasting.

Materials and methods:-

Study area:-

The Bui Dam as shown in Figure 1 is located at the border of the Bole and Wenchi districts, approximately 150km upstream of Lake Volta. With portions of the Dam site falling within Bui National Park, the entire Dam lies within Ghana. The terrain of the area is gently undulating with some deeply incised stream valleys. Deep and wide valleys of the Black Volta River, inselbergs and ridges, and the Banda Hills are some of the major relief and geomorphologic features of the area. Eroded granites corresponding to the Gonja Plateau and greywackes of the Birrimian and Tarkwaian formations are the two distinct geological features underlying the area. Upland soils and alluvial soils are the two major soil types occurring in the area (Ministry of Energy/Bui Development Committee, 2007).

The area is characterized by double-peak wet season, maxima in May-June and October. Average yearly rainfall in the area is around 1140mm. While monthly temperatures range from around 26^o C in August to around 30^o C in March; the area has a mean annual relative humidity of 75% and a mean annual pan evaporation of 1781mm. The dam is drained into the Volta Lake by the Black Volta which has its headwaters in Burkina Faso, where it is called the Mohoun River (Ministry of Energy/Bui Development Committee, 2007).

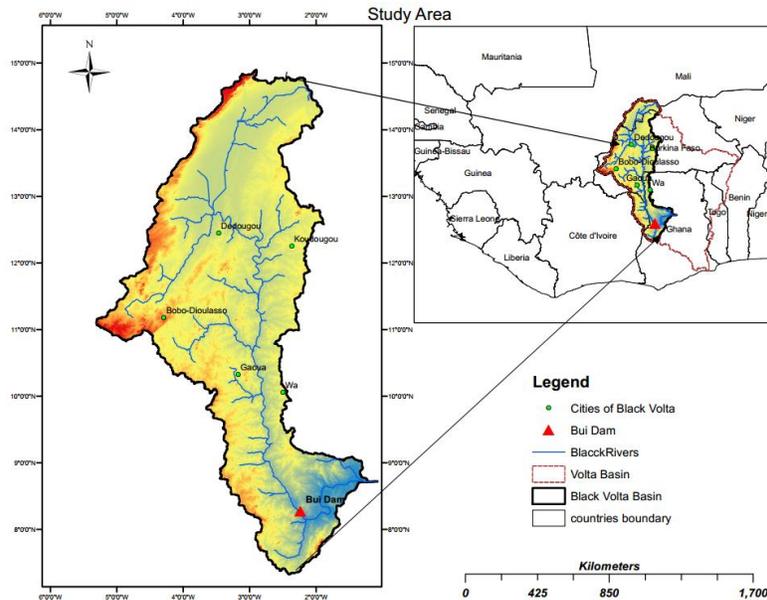


Figure 1:- Map of Study Area.

Data:-

Data was obtained from Bui Power Authority in the Brong-Ahafo Region of Ghana. The data consisted of Monthly Rainfall and Flow values from January 1954 to December 2005.

Vector autoregressive (VAR) model:-

Vector autoregressive (VAR) model as one of the most widely used and flexible models for analyzing multivariate time series is a natural extension of the univariate autoregressive model to dynamic multivariate time series and has superior forecast ability compared with those from univariate time series models and theory-based simultaneous equations models.

The basic vector autoregressive (VAR) model of order p , as suggested by Sims (1980) has the form

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + C D_t + u_t \quad (1)$$

Where $y_t = (y_{1t}, y_{2t}, \dots, y_{Kt})'$ is a vector of K observable endogenous variables, D_t contains all deterministic variables which may consist of a constant, a linear trend, seasonal dummy variables as well as user specified other dummy variables, u_t is a K -dimensional unobservable zero mean white noise process with positive definite covariance matrix $E(u_t u_t') = \Sigma_u$. A_i , and C are parameter matrices of suitable dimension on which you can impose

various restrictions. Example, by imposing zero restrictions, the right-hand side variables in equation (1) may be different in some of the equations. If we consider only one y variable (K=1), a univariate AR model is obtained.

The parameters in equation (1) are estimated by generalized least squares (GLS). This is done by first estimating the individual equations of the system by ordinary least squares (OLS). The residuals are then used to estimate the white noise covariance matrix Σ_u as $\widehat{\Sigma}_u = T^{-1} \sum_{t=1}^T \widehat{u}_t \widehat{u}_t'$ which is used to compute the GLS estimator.

Bayesian vector autoregressive (BVAR) model:-

VAR models generally uses equal lag length for all variables in the model which leads to the estimation of many parameters, some of which may not be significant. This has become a major drawback of the model resulting in possibly large out-of-sample forecasting errors.

Doan et al. (1984) proposed the BVAR model which involves the use of Bayesian prior information to impose certain restrictions on the coefficients of the VAR so as to overcome the problem of over parameterization associated with VAR models. The prior means and variances they suggested, also known as the Minnesota priors take the following form;

$$\beta_i \sim N(1, \sigma_{\beta_i}^2) \quad \text{and} \quad \beta_j \sim N(0, \sigma_{\beta_j}^2) \tag{2}$$

Where β_i represents the coefficients associated with the lagged dependent variable and β_j denotes any other coefficient in each equation of the VAR model. Whiles the prior means for lagged dependent variables are set to one, a prior mean of zero is assigned to all other coefficients in the equation. This is done in belief that lagged dependent variables are important whiles all other variables in each equation of the VAR model are viewed as less important. The prior variances indicate uncertainty about the prior means of one, for lagged dependent variables and prior means of zero, for all other variables.

Doan et al. (1984) proposed a formula to generate the standard deviations which is a function of a few hyperparameters: θ , φ , and a weighting matrix $w(i, j)$, to deal with the problem of over parameterization associated with VAR models. The standard deviation of the prior imposed on variable j at lag k in equation i is:

$$\sigma_{ijk} = \theta w(i, j) k^{-\varphi} \left(\frac{\widehat{\sigma}_{uj}}{\widehat{\sigma}_{ui}} \right) \tag{3}$$

Where $\widehat{\sigma}_{ui}$ and $\widehat{\sigma}_{uj}$ are the estimated standard errors from a univariate autoregression involving variables i and j respectively. θ is labelled as the ‘overall tightness’ to reflect the standard deviation of the prior on the first lag of the dependent variable. $k^{-\varphi}$ is a lag decay function with $0 \leq \varphi \leq 1$ which has the effect of imposing the prior means of zero more tightly as the lag length increases. The tightness of the prior for variable j in equation i relative to the tightness of the own-lags of variable i in equation i , is indicated by the function $w(i, j)$. The weighting matrix used in the standard Minnesota prior is given in equation (4) while the overall tightness and lag decay hyperparameters used are $\theta = 0.1$ and $\varphi = 1.0$ respectively.

$$W = \begin{bmatrix} 1 & 0.5 & \dots & 0.5 \\ 0.5 & 1 & & 0.5 \\ \vdots & & \ddots & \vdots \\ 0.5 & 0.5 & \dots & 1 \end{bmatrix} \tag{4}$$

BVAR models with the standard Minnesota prior are usually estimated using the mixed estimation method in Theil and Goldberger (Theil & Goldberger, 1961).

Optimal lag length selection criteria:-

For a range of lag orders n the individual equations of the system are estimated by OLS. The optimal lag order is selected by minimizing one of the following information criteria:

$$\text{Akaike Information Criterion, } AIC(n) = \log \det(\widehat{\Sigma}_u(n)) + \frac{2}{T} nK^2 \tag{5}$$

$$\text{Hannan-Quinn criterion, } HQ(n) = \log \det(\widehat{\Sigma}_u(n)) + \frac{2 \log \log T}{T} nK^2 \tag{6}$$

$$\text{Schwarz Criterion, } SC(n) = \log \det(\widehat{\Sigma}_u(n)) + \frac{\log T}{T} nK^2 \quad (7)$$

$$\text{Final Prediction Error criterion, } FPE(n) = \left(\frac{T+n^*}{T-n^*}\right)^K \det(\widehat{\Sigma}_u(n)) \quad (8)$$

Where $\widehat{\Sigma}_u(n)$ is estimated by $T^{-1} \sum_{t=1}^T \widehat{u}_t \widehat{u}_t'$, n^* is the total number of parameters in each equation of the model when n is the lag order of the endogenous variables.

Structural analysis:-

Despite the fact that VAR coefficients capture the anticipated impact of a variable, there are often a lot of coefficients to interpret. It is usually more common to examine the model's residuals which represent unforeseen contemporaneous events. The next subsections provide relatively non-technical explanations of some of the common techniques used for structural analysis of VAR models.

Causality analysis:-

Both the Granger-causality and instantaneous causality were investigated. For both tests, the vector of endogenous variables is divided in two sub vectors, y_{1t} and y_{2t} , with dimensions K_1 and K_2 , respectively, so that $K = K_1 + K_2$. The sub vector y_{1t} is said to be Granger-causal for y_{2t} if the past of y_{1t} significantly help predicting the future of y_{2t} via the past of y_{1t} alone (Granger, 1969). For testing this property, a model of the form

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \sum_{i=1}^p \begin{bmatrix} \alpha_{11,i} & \alpha_{12,i} \\ \alpha_{21,i} & \alpha_{22,i} \end{bmatrix} \begin{bmatrix} y_{1,t-i} \\ y_{2,t-i} \end{bmatrix} + CD_t + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \quad (9)$$

is considered. In this model setup, y_{1t} is not Granger-causal for y_{2t} if and only if $\alpha_{21,i} = 0$, $i = 1, 2, \dots, p$.

Therefore this null hypothesis is tested against the alternative that at least one of the $\alpha_{21,i}$ is nonzero. An F-test statistic which is distributed as $F(pK_1K_2, KT - n^*)$ is used for testing the restrictions. Here n^* is the total number of parameters in the system including the parameters of the deterministic term (Lütkepohl, 1991). The role of y_{1t} and y_{2t} can be reversed to test Granger-causality from y_{2t} to y_{1t} .

Instantaneous causality is characterized by nonzero correlation of u_{1t} and u_{2t} . Thus the null hypothesis

$$H_0: E(u_{1t}u_{2t}') = 0$$

is tested against the alternative of nonzero covariance between the two error vectors in testing for instantaneous causality. A Wald test statistic is used to test this hypothesis.

Impulse response analysis:-

In impulse response analysis the exogenous and deterministic variables are treated as fixed and may therefore be dropped from the system. The adjusted endogenous variables are now denoted by y_t . If the process y_t is stationary (I(0)), it has a Wold moving average (MA) representation

$$y_t = \Phi_0 u_t + \Phi_1 u_{t-1} + \Phi_2 u_{t-2} + \dots, \quad (10)$$

Where $\Phi_0 = I_K$ and the Φ_s can be computed recursively as

$$\Phi_s = \sum_{j=1}^s \Phi_{s-j} A_j, \quad s = 1, 2, \dots,$$

With $\Phi_0 = I_K$ and $A_j = 0$ for $j > p$. The coefficients of this representation may be interpreted as reflecting the responses to impulses hitting the system. The (i, j) th elements of the matrices Φ_s , regarded as a function of s , trace out the expected response of $y_{i,t+s}$ to a unit change in y_{jt} holding constant all past values of y_t .

Forecast error variance decomposition:-

Denoting the (i, j) th element of the orthogonalized impulse response coefficient matrix Ψ_n by $\psi_{ij,n}$, the variance of the forecast error $(y_{k,T+h} - y_{k,T+h|T})$ is

$$\sigma_k^2(h) = \sum_{n=0}^{h-1} (\psi_{k1,n}^2 + \dots + \psi_{kK,n}^2) = \sum_{j=1}^K (\psi_{kj,0}^2 + \dots + \psi_{kj,h-1}^2).$$

The term $(\psi_{kj,0}^2 + \dots + \psi_{kj,h-1}^2)$ is interpreted as the contribution of variable j to the h -step forecast error variance of variable k . Dividing the above terms by $\sigma_k^2(h)$ gives the percentage contribution of variable j to the h -step forecast error variance of variable k ,

$$w_{kj}(h) = (\psi_{kj,0}^2 + \dots + \psi_{kj,h-1}^2) / \sigma_k^2(h)$$

Stability analysis:-

Even though there are several options available for checking model stability including recursive residuals and parameter estimates, Chow tests and Cumulative sum (CUSUM) tests, this study used the CUSUM test to check model stability.

The cumulative sum of recursive residuals up to period τ is

$$CUSUM_{\tau} = \sum_{t=M+1}^{\tau} \hat{u}_t^{(r)} / \hat{\sigma}_u \tag{11}$$

Where $\hat{\sigma}_u = \sqrt{(T - M)^{-1} \sum_{t=1}^T \hat{u}_t^2}$ is the usual residual standard deviation estimator based on the full sample.

It is plotted for $\tau = M + 1, \dots, T$ together with the lines $\pm c_{\gamma} [\sqrt{T - M} + 2(\tau - M) / \sqrt{T - M}]$, where c_{γ} depends on the desired significance level of the resulting test. If the CUSUMs wander beyond these lines, then there is evidence against structural stability of the underlying model.

Results and Discussion:-

Descriptive statistics:-

At an initial step, descriptive statistics for the data on rainfall and flow are produced in Table 1. The results contains among others, the minimum, maximum, and average values of rainfall and flow for the period considered. Whiles the minimum, maximum, and average values of rainfall are 0.00mm, 407.50mm, and 95.09mm respectively, the minimum, maximum, and average values of flow are 0.01cms, 2417.30cms, and 206.42cms respectively.

Table 1:- Descriptive Statistics.

Rainfall		Flow	
Minimum	0.00	Minimum	0.10
1st Quartile	17.48	1st Quartile	14.72
Median	83.45	Median	64.10
Mean	95.09	Mean	206.42
3rd Quartile	146.15	3rd Quartile	267.27
Maximum	407.50	Maximum	2417.30

Figure 2 shows the time series plot for both rainfall and flow. The figure does not show any level of trend which is an indication of stationarity even though the Augmented Dickey Fuller (ADF) test will still be used formally to test for stationarity.

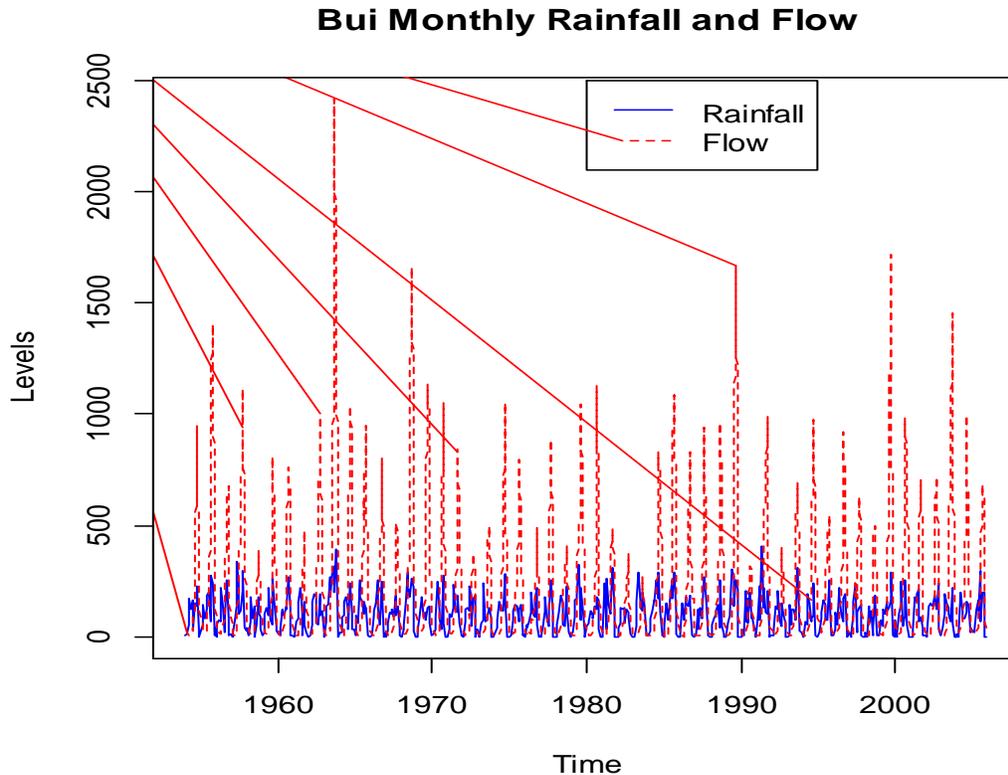


Figure 2:- Monthly Rainfall and Flow Levels at Bui.

Stationary test and optimal lag selection:-

To examine the stationarity (a necessary condition for VAR modeling) of the two time series data, the Augmented Dickey Fuller (ADF) unit root test was employed. The ADF test results for rainfall and flow are shown in Table 2. The results reveal stationarity for both series confirming the results from the time series plot. This indicates that VAR is suitable for modelling the time series data.

Table 2:- Augmented Dickey-Fuller Unit Root Test.

Rainfall			Flow		
Coefficients:	Estimate [Std. Error]	t value [Pr(> t)]	Coefficients:	Estimate [Std. Error]	t value [Pr(> t)]
(Intercept)	90.76351 [6.80386]	13.34 [< 2e-16]	(Intercept)	123.497115 [11.833747]	10.436 [< 2e-16]
z.lag.1	-0.95384 [0.06424]	-14.847 [< 2e-16]	z.lag.1	-0.5954 [0.042936]	-13.867 [< 2e-16]
z.diff.lag1	0.30701 [0.05669]	5.416 [8.76e-08]	z.diff.lag1	0.634798 [0.041373]	15.343 [< 2e-16]
z.diff.lag2	0.30663 [0.05143]	5.962 [4.22e-09]	z.diff.lag2	-0.008697 [0.046262]	-0.188 [0.8509]
z.diff.lag3	0.28955 [0.04605]	6.287 [6.14e-10]	z.diff.lag3	0.216862 [0.039406]	5.503 [5.48e-08]
z.diff.lag4	0.20482 [0.03967]	5.163 [3.30e-07]	z.diff.lag4	0.100111 [0.040355]	2.481 [0.0134]

According to Lütkepohl & Saikkonen (1997), when the fitted VAR model order (p) is assumed to vary as the cube root of the size of the time series, then $VAR(p + 1)$ is fitted to the data such that p goes to infinity with sample size. Hence for our dataset, we considered VAR models from lag 1 to lag 10 and computed AIC, HQ, SC, and FPE values for each. By choosing the fitted candidate model corresponding to the minimum values of these measures, one is attempting to select the candidate model which is rendered most plausible by the data at hand. Therefore, the smaller the AIC, HQ, SC, and FPE value, the better the model. We observe from Table 3 that, while the AIC, the HQ and the FPE criterion favour VAR(10), the SC information criterion rather favoured VAR(4). However for purposes of forecasting, economists' wisdom suggest it is better to use the model favoured by the Schwartz (SC) information criterion (Galvao, 2012). Therefore, based on SC values we selected VAR (4) for model fitting.

Table 3:- Optimal Lag Length Selection.

Lags	Criteria			
	AIC(n)	HQ(n)	SC(n)	FPE(n)
1	1.9239E+01	1.9261E+01	1.9296E+01	2.2657E+08
2	1.8796E+01	1.8829E+01	1.8882E+01	1.4553E+08
3	1.8607E+01	1.8651E+01	1.8722E+01	1.2042E+08
4	1.8488E+01	1.8544E+01	1.8632E+01	1.0699E+08
5	1.8461E+01	1.8528E+01	1.8634E+01	1.0410E+08
6	1.8454E+01	1.8532E+01	1.8655E+01	1.0336E+08
7	1.8420E+01	1.8510E+01	1.8650E+01	9.9929E+07
8	1.8383E+01	1.8484E+01	1.8642E+01	9.6306E+07
9	1.8374E+01	1.8486E+01	1.8662E+01	9.5468E+07
10	1.8357E+01	1.8480E+01	1.8674E+01	9.3846E+07

VAR estimation results:-

Table 4 shows the VAR(4) results for rainfall and flow. The parameter estimates together with their respective standard errors, t-values and p-values are presented. Without exceptions, all parameter estimates are significant at the 0.05 level of significance. It is however imperative to note that interpreting the impulse response function for the model is much more informative than interpreting the individual parameter estimates.

Table 4:- VAR Estimation results.

Rainfall			Flow		
Coefficients:	Estimate [Std. Error]	t value [Pr(> t)]	Coefficients:	Estimate [Std. Error]	t value [Pr(> t)]
Rainfall.L1	0.10212 [0.04170]	2.449 [0.0146*]	Rainfall.L1	-0.316032 [0.112927]	-2.799 [0.00530*]
Flow.L1	0.12212 [0.01541]	7.923 [1.11e-14*]	Flow.L1	0.900972 [0.041747]	21.582 [< 2e-16*]
Rainfall.L2	0.09650 [0.04189]	2.304 [0.0216*]	Rainfall.L2	0.365812 [0.113442]	3.225 [0.00133*]
Flow.L2	-0.16560 [0.01998]	-8.290 [7.26e-16*]	Flow.L2	-0.576499 [0.054102]	-10.656 [< 2e-16*]
Rainfall.L3	0.06204 [0.04101]	2.513 [0.01309*]	Rainfall.L3	0.938559 [0.111081]	8.449 [< 2e-16*]
Flow.L3	0.01321 [0.02083]	1.634 [0.05261*]	Flow.L3	0.027008 [0.056405]	1.479 [0.063223*]
Rainfall.L4	-0.08373 [0.04286]	-1.954 [0.0512*]	Rainfall.L4	0.646973 [0.116078]	5.574 [3.75e-08*]
Flow.L4	-0.06893 [0.01579]	-4.364 [1.50e-05*]	Flow.L4	-0.205122 [0.042777]	-4.795 [2.05e-06*]
Const	104.93703 [7.87593]	13.324 [< 2e-16*]	const	22.384929 [21.330364]	2.049 [0.029439*]

* Significant at 0.05 level

Figure 3 is the diagram of fit and residuals for the rainfall model. The standardized residuals plot in the figure shows the residuals are distributed randomly around zero which is an indication that the residuals from the model have a constant variance. The Auto-Correlation Function (ACF) and the Partial Auto-Correlation Function (PACF) residual plots in the figure suggest that there is no autocorrelation in the residuals of the model which indicate no lost of information by the rainfall model.

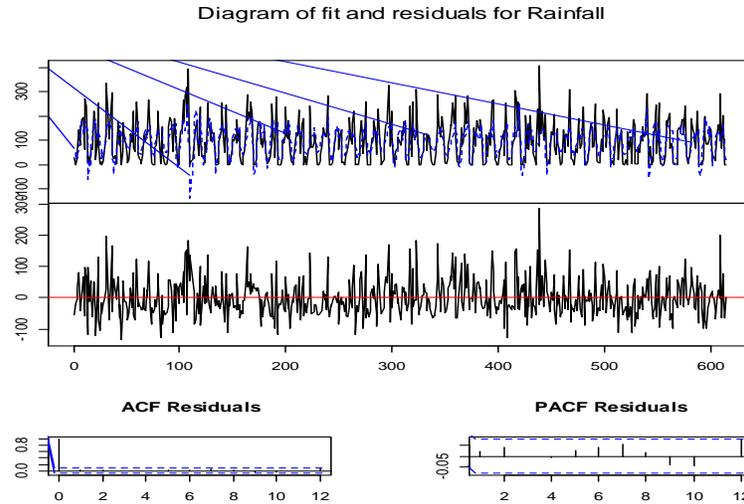


Figure 3:- Diagram of Fit for Rainfall

Figure 4 is the diagram of fit and residuals for the flow model. The standardized residuals plot in the figure shows the residuals are distributed randomly around zero which is an indication that the residuals from the model have a constant variance. Also, the Auto-Correlation Function (ACF) and the Partial Auto-Correlation Function (PACF) residual plots in the figure suggest that there is no autocorrelation in the residuals of the model which indicate no lost of information by the flow model.

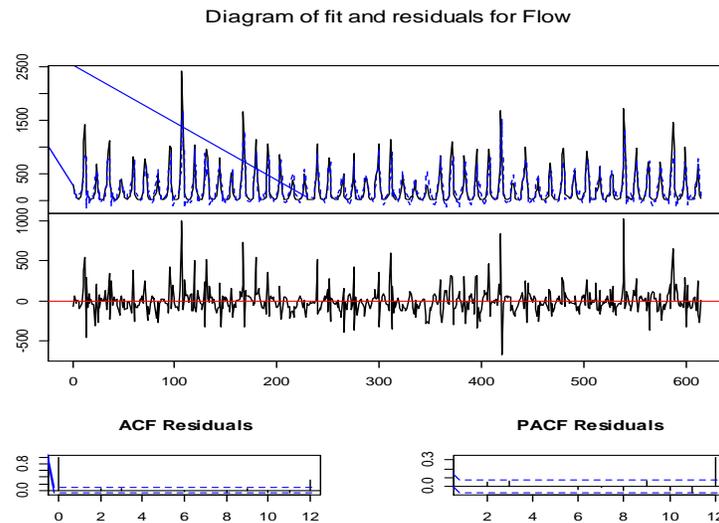


Figure 4:- Diagram of Fit for Flow.

BVAR estimation results:-

The BVAR estimation results for both rainfall and flow are presented in Table 5. For the rainfall model, only flow at lag 1 is insignificant among the lags for flow whiles the only significant lag among the lags for rainfall is lag 1. On the other hand, rainfall at lags 1 and 2 are significant in the flow model whiles flow at lag 3 is insignificant.

Table 5:- BVAR Estimation results.

Rainfall				Flow			
Coefficients:	Estimate	t value	[Pr(> t)]	Coefficients:	Estimate	t value	[Pr(> t)]
Rainfall.L1	0.4165	11.3237	0.0000*	Rainfall.L1	-0.0221	-0.2528	0.8005
Rainfall.L2	0.0097	0.3204	0.7488	Rainfall.L2	0.1303	2.3092	0.0213*
Rainfall.L3	0.0031	0.1265	0.8993	Rainfall.L3	0.1524	3.7776	0.0002*
Rainfall.L4	-0.0298	-1.4898	0.1368	Rainfall.L4	0.0536	1.7120	0.0874
Flow.L1	0.0164	1.6612	0.0972	Flow.L1	0.8811	26.7440	0.0000*
Flow.L2	-0.0417	-5.6611	0.0000*	Flow.L2	-0.3144	-9.7304	0.0000*
Flow.L3	-0.0192	-3.5936	0.0004*	Flow.L3	-0.0435	-1.7175	0.0864
Flow.L4	-0.0108	-2.5995	0.0096*	Flow.L4	-0.0459	-2.3077	0.0213*
Const	68.5370	13.0182	0.0000*	Const	78.3977	5.7378	0.0000*

* Significant at 0.05 level

Model selection:-

Table 6 presents the mean absolute errors (MAEs) and coefficients of determination for both flow and rainfall under the VAR and BVAR models. The VAR model accounts for a little over seventy percent of the variability in flow and a little over forty percent of the variability in rainfall whereas the BVAR model only accounts for about twenty-nine and sixty-one percent of the variability in rainfall and flow respectively. Also, by inspecting the MAE values which is a measure of prediction accuracy, we observe that MAE values for VAR models are smaller than those for the BVAR models. Hence the VAR model is selected for structural analysis and forecasting.

Table 6:- Model Selection.

	VAR		BVAR	
	Rainfall	Flow	Rainfall	Flow
Multiple R-Squared	0.4182	0.7055	0.2926	0.6173
Adjusted R-squared	0.4096	0.7012	0.2833	0.6123
MAE	10.1103	6.5224	14.3451	12.6013

Results of the structural analysis:-

As indicated earlier, it is often more informative and common to conduct structural analysis after fitting a VAR model by examining the model's residuals. The next subsections provide results of the structural analysis of the fitted VAR model.

Causality results:-

Results of both the Granger and instantaneous causality tests in Table 7 indicate a two-way causality from Rainfall to Flow and vice versa. Granger causality is an F-test while Instantaneous causality is a Wald Chi-squared test. Testing the hypothesis that Rainfall Granger-causes Flow, the results in Table 7 reveals an F-value of 15.283 with p-value < 2.2e-16. This indicates the null hypothesis that, Rainfall do not Granger-cause Flow is rejected even at 1% level of significance. Furthermore, results of the instantaneous causality between Rainfall and Flow reveals a Chi-squared value of 74.743 with p-value < 2.2e-16 thereby rejecting the null hypothesis of no instantaneous causality between Rainfall and Flow at 1% level of significance.

Table 7:- Causality Tests.

Granger and Instantaneous Causality Tests	
\$Granger	Granger causality H0: Rainfall do not Granger-cause Flow F-Test = 15.283, df1 = 10, df2 = 1184, p-value < 2.2e-16
	Granger causality H0: Flow do not Granger-cause Rainfall F-Test = 24.715, df1 = 10, df2 = 1184, p-value < 2.2e-16
\$Instant	H0: No instantaneous causality between: Rainfall and Flow Chi-squared = 74.743, df = 1, p-value < 2.2e-16

Impulse response analysis:-

Impulse response analysis was utilized to analyze the dynamic interactions between rainfall and flow of the VAR (4) process. The orthogonal impulse response of rainfall to flow is presented in Figure 5. The response of rainfall has an obvious fluctuation; there is a highest positive effect of flow on rainfall in the fifth month (May) and lowest negative effect of flow on rainfall in the ninth month (September). In Figure 6, the orthogonal impulse response of flow to rainfall is presented, which shows another obvious fluctuation. While the highest positive effect of rainfall on flow was recorded in the second month (February), the lowest negative effect of rainfall on flow was recorded in the fifth month (May).

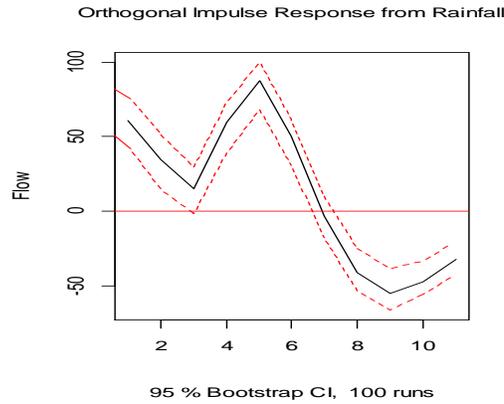


Figure 5:- Impulse Response from Rainfall

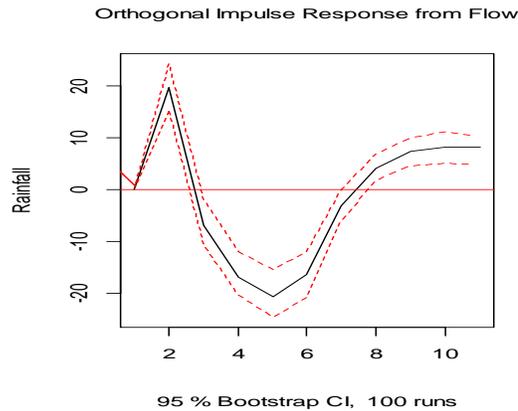


Figure 6:- Impulse Response from Flow.

Forecast error variance decomposition:-

Forecast error variance decompositions (FEVDs) are also popular in interpreting VAR models. Results for the FEVDs for both rainfall and flow are presented in Figure 7. The results reveal that on the average, while about 81% of the variability in the trend of rainfall has been explained by past innovations in rainfall figures at Bui, a significant proportion (about 19%) of the variability in the trend of rainfall have been explained by past innovations in flow values at Bui. The results further reveal that on the average, while about 78% of the variability in the trend of flow has been accounted for by past innovations in flow values at Bui, about 22% (which is very significant) of the variability in the trend of flow has been explained by past innovations in rainfall figures at Bui. This confirms the results of the two-way relationship between rainfall and flow obtained from the granger-causality test, and indicate that modelling rainfall and flow together will further improve the forecast of each of them.

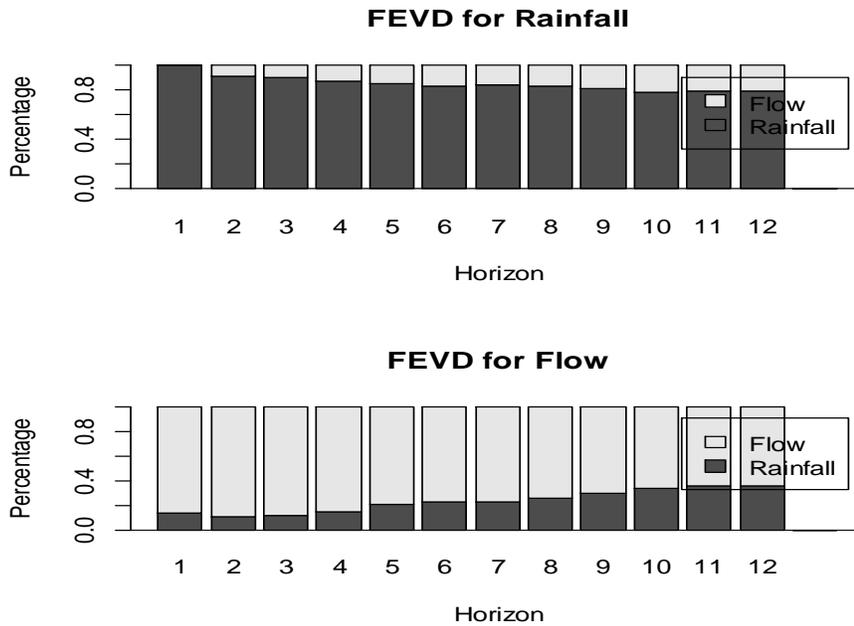


Figure 7:- Forecast Error Variance Decompositions

Stability results:-

Parameter stability throughout the study period is an important assumption in VAR modelling. If the parameters of the model are different during the forecast period than they were during the sample period, then the estimated model will not be very useful irrespective of how well it was estimated. Furthermore, if the parameters of the model were unstable over the sample period, then the model was not even a good representation of how the series evolved over the sample period. Visual examinations of the graphs of the recursive residuals are useful in evaluating the stability of the model. The ordinary least squares cumulative sum (OLS-CUSUM) graphs for the fitted models in Figure 8 indicate structural stability of the underlying models since the cumulative sums wander within the control lines for both models.

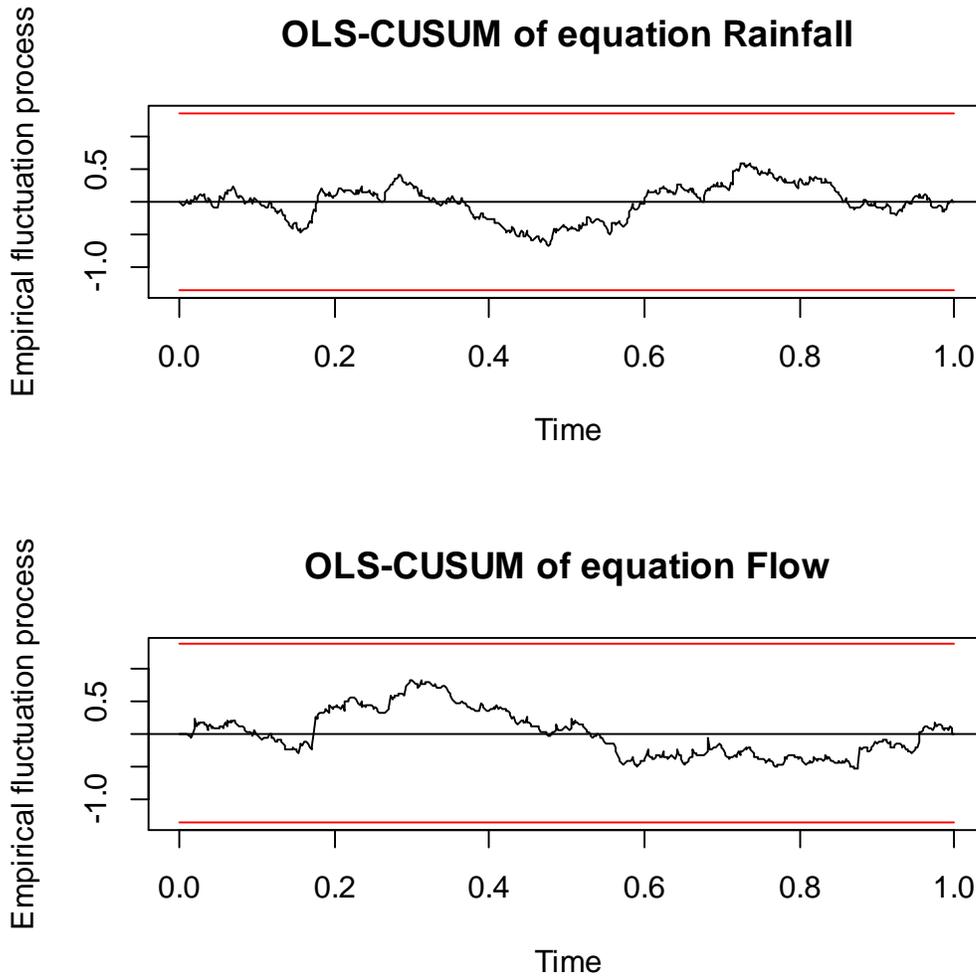


Figure 8:- Diagram of Stability for Rainfall and Flow

Flow and rainfall forecast:-

The VAR(4) model developed is a predictive model for making forecasts of future flow and rainfall figures at Bui. The predicted rainfall and flow values for the first half of 2006 are presented in Figures 9 and 10 respectively. In each case, the black line represents the observed values for the first half of 2006 while the red line represent the predicted values for the same period. The blue lines are the 95% confidence interval for the predicted values. We observe from Figure 9 that, the trends of both the observed and predicted lines for rainfall under the model are similar at least for the first 3 months. Also, the observed and predicted values themselves are close. This suggests that the VAR(4) model developed is very good in forecasting rainfall figures at Bui. The model even performs better at predicting flow as observed from Figure 10. The trends of both the observed and predicted lines for flow under the model are similar for the entire 6 months and the values are also close.

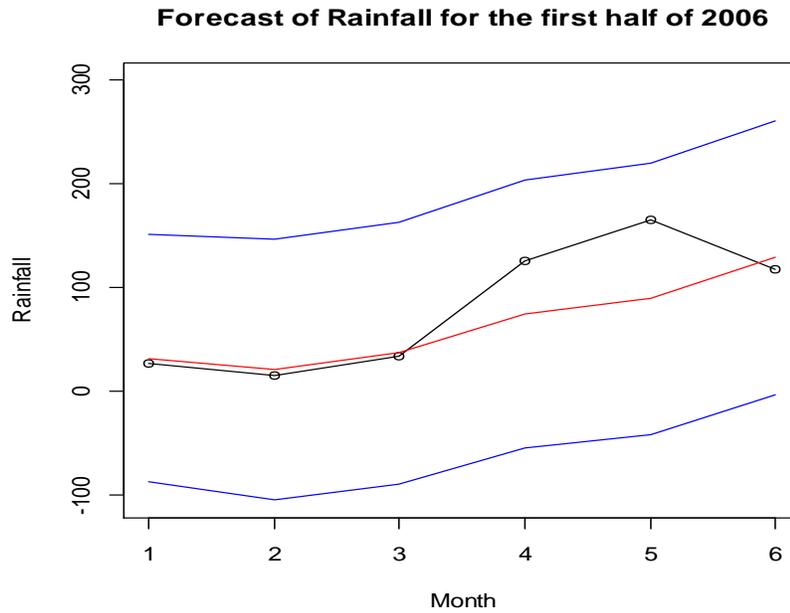


Figure 9:- Predicted Rainfall Figures for the First Half of 2006 under the VAR (4) Model.

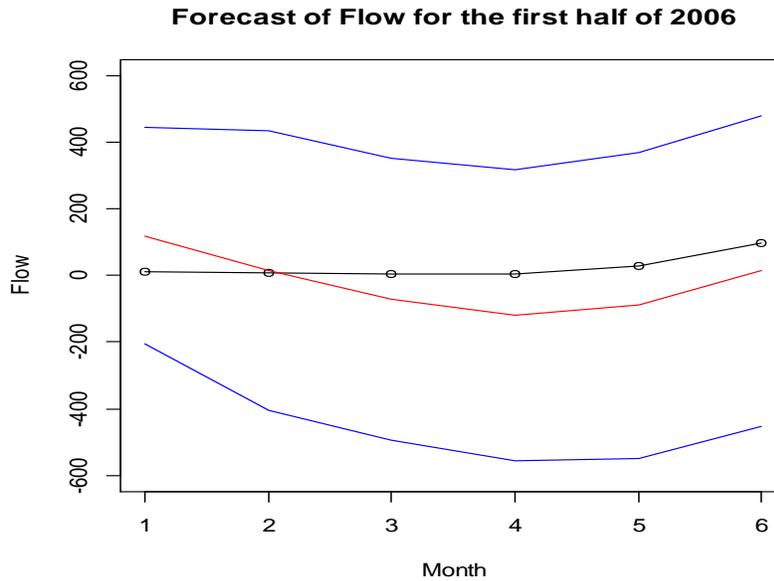


Figure 10:- Predicted Values of Flow for the First Half of 2006 under the VAR (4) Model

Conclusion:-

The implications from this study are quite clear. First of all, we investigated to determine between unrestricted and Bayesian VAR models the one that is best suited for forecasting flows with respect to rainfall variability at Bui. The unrestricted VAR model proved to have superior forecasting abilities than the Bayesian VAR model based on MAE values. We also investigated to determine whether past rainfall figures at Bui significantly helps in predicting future trends of flows via past flow values alone. The Granger causality test conducted revealed a two-way causality from rainfall to flow and from flow to rainfall which was confirmed by results of the forecast error variance decomposition. Finally, we modelled the trend of flow with respect to rainfall variability at Bui. Results from the fitted VAR(4) model revealed that modelling rainfall and flow together improves the forecast of each of them at Bui,

which is an improvement in our understanding and knowledge of hydrological (Flows) variability in the context of climate (Rainfall) variability.

Acknowledgements:-

This study was supported by the Akosombo Kpong Dams Reoperation and Reoptimization Study hosted by the Water Resources Commission (WRC) of Ghana and funded by the African Development Bank. We sincerely thank WRC and ADB for the support.

References:-

1. Abdul-Aziz, A. M., Kwame, A., Munyakazi, L., & Nsowa-Nuamah, N. N. N. (2013). Modelling and Forecasting Rainfall Pattern in Ghana as a Seasonal Arima Process: The Case of Ashanti Region. *International Journal of Humanities and Social Science*, 3(3), 224–233.
2. Adenomom, M. O., Ojehomon, V. E. T., & Oyejola, B. A. (2013). Modelling the dynamic relationship between rainfall and temperature time series data in Niger State, Nigeria. *Mathematical Theory and Modelling*, 3(4), 53–70.
3. Ampadu, B. (2007). *Simulating rainfall and riverflow dynamics in Ghana*. Lancaster University. Retrieved from <http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.507277>
4. Ampadu, B., Chappell, N. A., & Kasei, R. A. (2013). Rainfall-riverflow modelling approaches: making a choice of data-based mechanistic modelling approach for data limited catchments: a review. *Canadian Journal of Pure and Applied Sciences*, 2571.
5. Ampaw, E. M., Akuffo, B., Larbi, S. O., & Lartey, S. (2013). Time Series Modelling of Rainfall in New Juaben Municipality of the Eastern Region of Ghana. *International Journal of Business and Social Science*, 4(8). Retrieved from <http://search.proquest.com/openview/f5a5a9712e1054d5489e4b4b381e978d/1?pq-origsite=gscholar>
6. Andersson, J. (2007). Forecasting Swedish GDP Growth. Retrieved from <http://lup.lub.lu.se/student-papers/record/1336739>
7. Awotwi, A., Kumi, M., Jansson, P. E., Yeboah, F., & Nti, I. K. (2015). Predicting Hydrological Response to Climate Change in the White Volta Catchment, West Africa. *Journal of Earth Science & Climatic Change*, 2015. Retrieved from <http://www.omicsonline.org/open-access/predicting-hydrological-response-to-climate-change-in-the-white-volta-catchment-west-africa-2157-7617.1000249.php?aid=36768>
8. Bessler, D. A. (1984). Relative prices and money: a vector autoregression on Brazilian data. *American Journal of Agricultural Economics*, 66(1), 25–30.
9. Boochabun, K., Tych, W., Chappell, N. A., Carling, P. A., Lorsirirat, K., & Pa-Obsaeng, S. (2004). Statistical modelling of rainfall and river flow in Thailand. *JOURNAL-GEOLOGICAL SOCIETY OF INDIA*, 64(4), 503–516.
10. CSIR, & WRI. (2000). Climate change vulnerability adaptation assessment on water resources of Ghana. CSIR, Accra.
11. Doan, T., Litterman, R., & Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, 3(1), 1–100.
12. Franchini, M., & Pacciani, M. (1991). Comparative analysis of several conceptual rainfall-runoff models. *Journal of Hydrology*, 122(1), 161–219.
13. Galvao, A. B. (2012). Vector autoregressions II, Empirical macroeconomics-Lect2. Queen Mary University of London.
14. Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424–438.
15. Kaylen, M. S. (1988). Vector autoregression forecasting models: recent developments applied to the US hog market. *American Journal of Agricultural Economics*, 70(3), 701–712.
16. Lütkepohl, H. (1991). Introduction to multiple time series analysis. *Springer Verlag, Berlin*.
17. Lütkepohl, H., & Saikkonen, P. (1997). Impulse response analysis in infinite order cointegrated vector autoregressive processes. *Journal of Econometrics*, 81(1), 127–157.
18. McCarty, T. A., & Schmidt, S. J. (1997). A vector-autoregression analysis of state-government expenditure. *The American Economic Review*, 87(2), 278–282.
19. Ministry of Energy/Bui Development Committee. (2007). Environmental and Social Impact Assessment Study of the Bui Hydroelectric Power Project, Final Report. Environmental Resources Management, in association with SGS Environment. Retrieved from www.erm.com

20. Nijssen, B., O'Donnell, G. M., Hamlet, A. F., & Lettenmaier, D. P. (2001). Hydrologic sensitivity of global rivers to climate change. *Climatic Change*, 50(1–2), 143–175.
21. Nugroho, A., Hartati, S., Mustofa, K., & others. (2014). Vector Autoregression (Var) Model for Rainfall Forecast and Isohyet Mapping in Semarang–Central Java–Indonesia. *International Journal of Advanced Computer Science & Applications*, 1(5), 44–49.
22. Oki, T., & Kanae, S. (2006). Global hydrological cycles and world water resources. *Science*, 313(5790), 1068–1072.
23. Park, T. (1990). Forecast evaluation for multivariate time-series models: The US cattle market. *Western Journal of Agricultural Economics*, 133–143.
24. Raghunath, H. M. (2006). *Hydrology: principles, analysis and design*. New Age International. Retrieved from [https://books.google.co.uk/books?hl=en&lr=&id=9mdkJ0T2P30C&oi=fnd&pg=PA3&dq=\(Raghunath,+2006\)&ots=ej6oHRZYgV&sig=ALCtyBsgpfu9Qsg4U_yaNxO_cOE](https://books.google.co.uk/books?hl=en&lr=&id=9mdkJ0T2P30C&oi=fnd&pg=PA3&dq=(Raghunath,+2006)&ots=ej6oHRZYgV&sig=ALCtyBsgpfu9Qsg4U_yaNxO_cOE)
25. Saputro, D. R. S., Wigena, A. H., & Djuraidah, A. (2011). MODEL VEKTOR AUTOREGRESSIVE UNTUK PERAMALAN CURAH HUJAN DI INDRAMAYU (Vector Autoregressive Model for Forecast Rainfall In Indramayu). In *Forum Statistika dan Komputasi* (Vol. 16). Retrieved from <http://journal.ipb.ac.id/index.php/statistika/article/view/4916>
26. Servat, E., & Dezetter, A. (1993). Rainfall-runoff modelling and water resources assessment in northwestern Ivory Coast. Tentative extension to ungauged catchments. *Journal of Hydrology*, 148(1), 231–248.
27. Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 1–48.
28. Theil, H., & Goldberger, A. S. (1961). On pure and mixed statistical estimation in economics. *International Economic Review*, 2(1), 65–78.
29. Young, P. (2003). Top-down and data-based mechanistic modelling of rainfall–flow dynamics at the catchment scale. *Hydrological Processes*, 17(11), 2195–2217.
30. Zhang, G.-H., Nearing, M. A., & Liu, B.-Y. (2005). Potential effects of climate change on rainfall erosivity in the Yellow River basin of China. *Transactions of the ASAE*, 48(2), 511–517.