

## **Modelling West African total precipitation depth: a statistical approach**

Sovoe, S.\*

\*Environmental Protection Agency, Ho, Volta Region, Ghana; [sovo35@yahoo.com](mailto:sovo35@yahoo.com);  
+233 246099870 (cell); +233 362026529 (office); +233 362025277 (home)

**Sponsor(s): International Foundation for Science, Grant No. W/4330-1**

### **Abstract**

Even though several reports over the past few decades indicate an increasing aridity over West Africa, attempts to establish the controlling factor(s) have not been successful. The traditional belief of the position of the Inter-tropical Convergence Zone (ITCZ) as the predominant factor over the region has been refuted by recent findings. Changes in major atmospheric circulations such as African Easterly Jet (AEJ) and Tropical Easterly Jet (TEJ) are being cited as major precipitation driving forces over the region. Thus, any attempt to predict long term precipitation events over the region using Global Circulation or Local Circulation Models could be flawed as the controlling factors are not fully elucidated yet. Successful prediction effort will require models which depend on past events as their inputs as in the case of time series models such as Autoregressive Integrated Moving Average (ARIMA) model. In this study, the researcher used historical precipitation data to build appropriate Seasonal Multiplicative Autoregressive Integrated Moving Average model, ARIMA (p,d,q)\*(P,D,Q) in an R programming language. The model was then used to predict long term precipitation events over the Ghanaian segment of the Volta Basin which could be used in planning and implementation of development policies.

**KEYWORDS: Modelling; West Africa; Total Precipitation Depth; Statistical Approach**

## **INTRODUCTION**

### **1.1 Background**

Climate change has resulted in extreme drought condition in some parts of the world and flooding in other parts (van de Giesien *et al.*, 2010). Environmental changes in Africa have mostly been directly related to rainfall (Nicholson, 2001; Zheng *et al.*, 1997). Precipitation over most regions of the continent traditionally has been associated with the seasonal excursion of the Inter Tropical Convergence Zone (ITCZ). Recent studies have however revealed that the seasonal development of the tropical rain belt over Africa is driven by several factors of the general atmospheric circulation which in turn control the location and characteristics of the ITCZ (Zheng *et al.*, 1997; Nicholson, 2001; Nicholson and Grist, 2003; Grist *et al.*, 2002). This atmospheric circulation is believed to generate and maintain wave disturbances that modulate the rainfall field. Nicholson *et al.* (2007) found the Tropical Easterly Jet (TEJ) as one of the most intense circulation features over Africa and concluded that the TEJ may be a critical factor in the development of the rainy season and the overall climate in West Africa as opposed to the traditional belief of seasonal movements of the inter-tropical convergence zone (ITZC). It is obvious that long term prediction of rainfall over West Africa using Global Circulation Model (GCMs) could be flawed as factors controlling precipitation over the region are still not well understood (van de Giesen *et al.*, 2010). Successful prediction effort will require models which depend on

past events as their inputs as in the case of Autoregressive Integrated Moving Average (ARIMA) model. The overarching goal of this research is to develop a model that can reliably forecast the catchment's long term rainfall event for effective and sustainable water resources management.

## MATERIALS AND METHODS

### 3.1 The study area

The study focused on the Ghanaian segment of the Volta River Basin. The area stretches from the mid section of the country (7°30'20.35"N, 2°59'47.22"W to 7°10'24.80" .09"E) to the northern section bordering three West African countries namely Togo in the east, Burkina Faso in the north and La Cote d'Ivoire in the west. Below the mid section it tapers towards south eastern and ended at 5°46'11.60"N, 0°43'02.74"E", where the Volta River empties into the Atlantic Ocean. The study area covers a total of 167, 692 square kilometres (70% of the country's total land area) constituting 40.18% of the total area of the Volta River Basin (417, 382 square kilometres) (Barry *et al.*, 2005; Leehuis *et al.*, 2009, Opoku-Duah *et al.*, 2008, Gordon and Amatekpor, 1999).

### 3.2 Selection of appropriate ARIMA model

Rainfall over West Africa could not be modelled adequately using Global Circulation Model (GCMs) since the factors controlling precipitation over the region are still not well understood (van de Giesen *et al.*, 2010; Philippon and Fontaine, 2002). Successful prediction effort will require statistical modelling approach which uses historical records to estimate the model parameters as in the case of Autoregressive Integrated Moving Average (ARIMA) model. The general forms of these models are (Sabry *et al.*, 2007):

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \epsilon_t + \mu - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$$

where:  $Y_t$  = response variable at time  $t$ ,  $Y_{t-1}, \dots, Y_{t-p}$  = response variables at time lags at  $t-1, \dots, t-p$  respectively,  $\phi_0, \phi_1, \dots, \phi_p$  = coefficients to be estimated for the AR(p) model,  $\theta_1, \dots, \theta_p$  coefficients to be estimated,  $\mu$  = constant mean of the process,  $\epsilon_t$  = error term at time  $t$ .

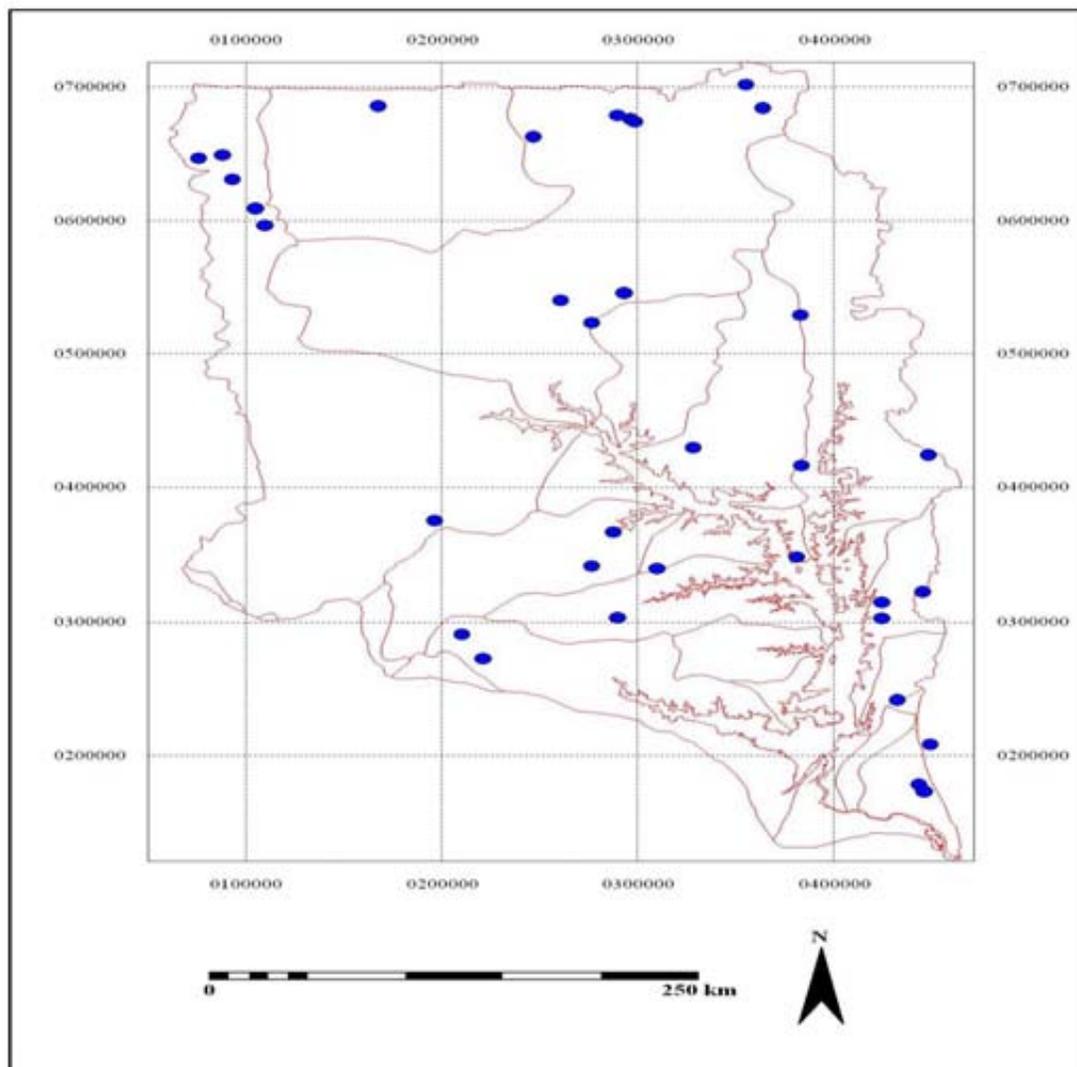
Rearranging the ARMA (p, q) model and introducing a differencing order gives the general Autoregressive Integrated Moving Average, ARIMA (p, d, q) model as shown below (Sabry *et al.*, 2007):

$$\begin{aligned} Y_t - \phi_0 - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} &= \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \\ (1 - \phi_0 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Y_t &= \mu + (1 - \theta_1 B - \dots - \theta_q B^q) \epsilon_t \\ \phi(B)(1 - B)^d Y_t &= \mu + \theta(B) \epsilon_t \\ Y_t &= \phi^{-1}(B)(1 - B)^{-d} (\mu + \theta(B) \epsilon_t) \end{aligned}$$

where  $\phi(B) = \phi_0 + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \epsilon_t$  and  $\theta(B) = \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$

### 3.3 Modelling of spatial component of the rainfall total

The ARIMA model considers the temporal variations only since the observed data are collected at some locations within the study area, for the spatial dimension to be accounted for in the time series analysis, a dedicated spatial analysis package is required to model the spatial component in the observations. In this study, the researcher used Geographic Information System (GIS) to model historical spatial variation of the precipitation over nearly forty year (40) record. 1967 – 2006 monthly precipitation data of the study area spanning roughly forty years were obtained from Ghana Meteorological Agency. A total of 33 rainfall stations data were analyzed. Though the available number (stations) of rainfall data was few it covered the entire study area thus giving a good spatial representation (figure 1).



**Fig. 1: Rainfall stations of the study area**

The monthly rainfall for each year was interpolated over a grid using moving average interpolation method. Data gaps in the monthly recordings were filled using predicted values from the interpolation. The complete monthly data as obtained above were converted into a vector layer. The vector data was rasterized and a distance between each point was calculated. A Thiessen polygon map depicting area coverage of a particular rainfall station was created for each month. The Thiessen polygon map was crossed with sub-basins map of the study area. Monthly rainfall total for each sub-catchment was generated from the cross operation, the sub-catchments monthly rainfall were summed to monthly total precipitation within the study area. The monthly total rainfall data were then imported into a statistical programming environment R. Using R time series analysis function, the data was decomposed into seasonal, trend and irregular components.

Box-Jenkins approach was used to select the appropriate Seasonal Multiplicative Autoregressive Integrated Moving Average ARIMA (p, d, q) \* (P, D, Q) model which fits the time series data. The dataset was divided into two, comprising 1967 – 2000 (33 years) record for model building and 2000 – 2006 (6 years) for model validation. The researcher used graphical methods to identify normality, trend, cyclical and seasonal variations in the data. A second order differencing was then used to render the data stationary. Six different options of the ARIMA models were fitted and for each model the researcher computed the Box-Jenkins test and also compared histogram of the residuals, residuals quantile versus theoretical quantile plot, autocorrelation function and the p-values of the Ljung-Box statistics. The model with the least Akaike Information Criterion (AIC) was parsimoniously selected among the rest. The fitted model was used to predict a six-year (2000 – 2006) time step ahead. The predicted values were plotted with 95% confidence interval with the observed data (2000 – 2006 validation record) superimposed on the plot of the predicted data. The model prediction was very good. The validated model was therefore used to predict forty-year time step ahead of precipitation events over the study area.

## RESULTS

### 4.1 Rainfall pattern

Different model parameter estimates were obtained during the search stage (table 1). The Aike Information Criterion (AIC) for each model was compared. Based on the AIC and the principle of parsimony in model fitting, second order seasonal autoregressive integrated moving average model ARIMA (2, 1, 1) (1, 1, 1) was selected.

**Table 1: Results of parameter estimates for average monthly rainfall depth using ARIMA model**

ARIMA (p, d, q) (P, D, Q)	Parameters	1 <sup>st</sup> Order	2 <sup>nd</sup> Order	AIC
(1, 1, 1) (1, 1, 1) <sub>12</sub>	Non-seasonal AR	0.149	-	3464.51
	Non-seasonal MA	-0.976	-	
	Seasonal AR	-0.0555	-	
	Seasonal MA	-0.9360	-	
(2, 1, 1) (1, 1, 1) <sub>12</sub>	Non-seasonal AR	0.1413	0.0860	3463.728
	Non-seasonal MA	-0.9813	-	
	Seasonal AR	-0.0468	-	
	Seasonal MA	-0.9356	-	
(2, 1, 2) (1, 1, 1) <sub>12</sub>	Non-seasonal AR	0.2722	0.0667	3465.65
	Non-seasonal MA	-1.1133	0.1290	
	Seasonal AR	-0.0461	-	
	Seasonal MA	-0.9354	-	

Substituting the estimated parameters of the selected seasonal ARIMA (2, 1, 1) \* (1, 1, 1)<sub>12</sub> model that predicts the average monthly rainfall depth  $Y_t - Y_{t-1} = \Delta Y_t$ , in the general formula gave the Seasonal Autoregressive Integrated Moving Average Model as below:

$$\Phi(B^h)\phi(B)\Delta_n^D\Delta^d Y_t = \alpha + \Theta(B^h)\theta(B)\epsilon_t$$

$$(1 - \Phi_1 B^h - \dots - \Phi_P B^{Ph})(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B^h)^D(1 - B)^d \Delta Y_t = \alpha + (1 - \Theta_1 B^h - \dots - \Theta_Q B^{Qh})(1 - \theta_1 B - \dots - \theta_q B^q)\epsilon_t$$

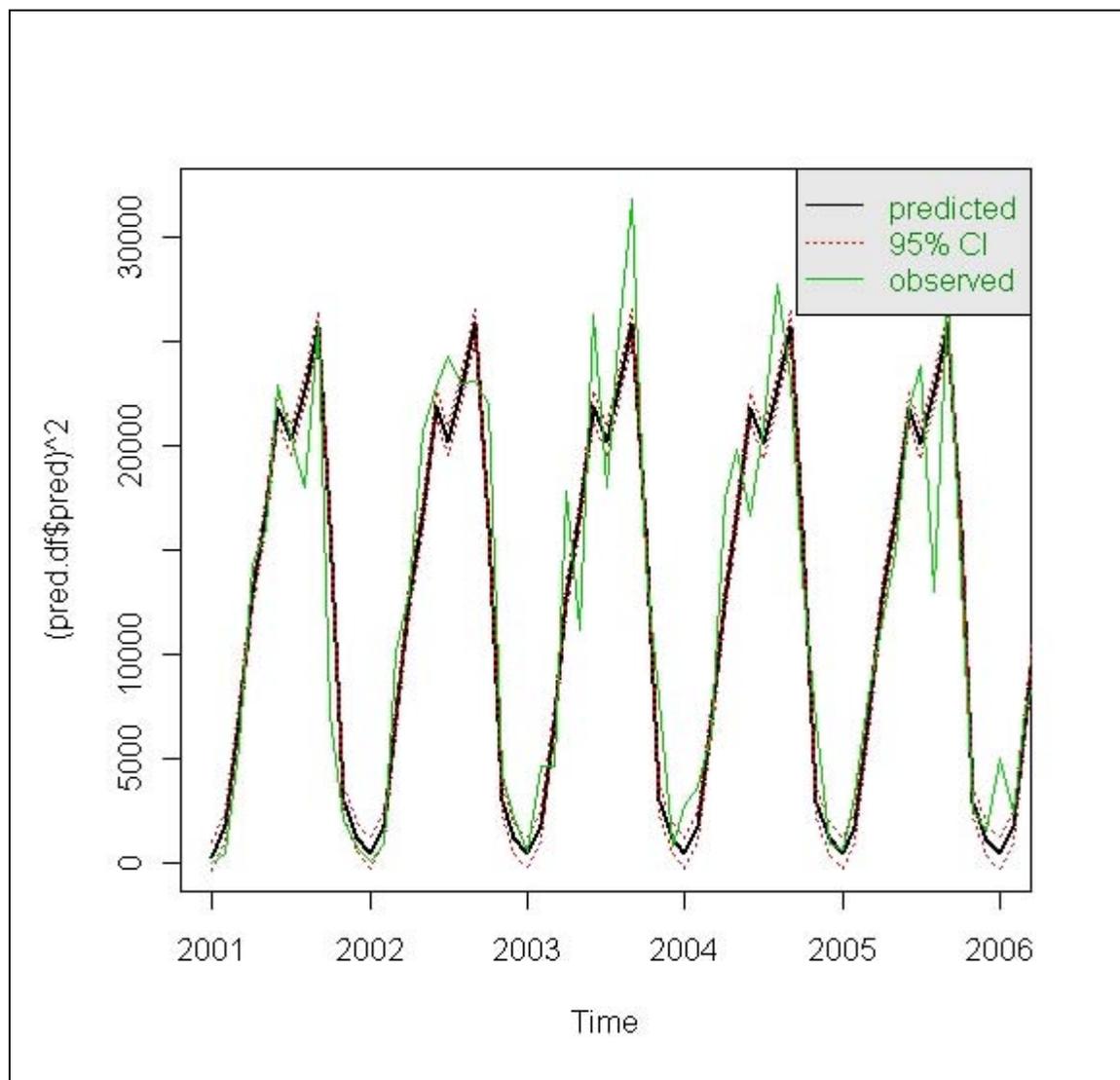
$$(1 - \Phi_1 B^{12})(1 - \phi_1 B - \phi_2 B^2)(1 - B^{12})(1 - B)\Delta Y_t = (1 - \Theta_1 B^{12})(1 - \theta_1 B)\epsilon_t$$

$$(1 + 0.0468 B^{12})(1 - 0.1413 B - 0.0860 B^2)(1 - B^{12})(1 - B)\Delta Y_t = (1 + 0.9356 B^{12})(1 + 0.9813 B)\epsilon_t$$

$$(1 + 0.0468 \Delta Y_{t-12})(1 - 0.1413 \Delta Y_{t-1} - 0.0860 \Delta Y_{t-2})(1 - \Delta Y_{t-12})(1 - \Delta Y_{t-1})\Delta Y_t = (1 + 0.9356 \epsilon_{t-12})(1 + 0.9813 \epsilon_{t-1})\epsilon_t$$

$$\Delta Y_t = (1 + 0.9356 \epsilon_{t-12})(1 + 0.9813 \epsilon_{t-1})\epsilon_t / [(1 + 0.0468 \Delta Y_{t-12})(1 - 0.1413 \Delta Y_{t-1} - 0.0860 \Delta Y_{t-2})(1 - \Delta Y_{t-12})(1 - \Delta Y_{t-1})]$$

After applying a non-seasonal and seasonal differencing of the time series data it was rendered stationary. The stationarity effect is shown in autocorrelation graph as spike cuts off quickly. Plot of the residuals, theoretical quantile plot and histogram plot after applying non-seasonal and seasonal autoregressive (AR) and moving average (MA) models to the transformed data indicate good fit of the ARIMA model. The Box-Jenkins statistics also clearly indicates the goodness of fit of the model as the p values are highly significant from zero (0). A plot of the six-year model prediction (from 2000 to 2006) against observed recordings over the same period at 95% confidence interval clearly shows an excellent performance of the model (figure 2).



**Fig. 2: A plot of six year model prediction against observed values at 95% confidence interval**

#### **4.2 Prediction result**

A forty-year time step ahead (2007 to 2047) prediction of rainfall over the study area by the model at 95% confidence interval indicates a steady decline in precipitation. It is observed from the decomposed plot that precipitation over the study area will decline within the next forty year period by approximately 9%.

### **DISCUSSION**

#### **5.1 Rainfall pattern**

The result of the analysis clearly showed that precipitation over the Volta catchment area is on the steady decline. The total rainfall within the catchment over the last forty years has declined by approximately seven and half percent (7.5%). The model prediction for the next forty years indicated a further decline of approximately nine percent (9%). The model performed very well as the predicted values corresponded well with the observed values. The results of the analysis corroborated other earlier reports of the decline of the catchment's precipitation (Leehuis *et al.*, 2009; Long *et al.*, 2000). Similar downward trend was reported by Nicholson and Palao (1993) in the Sahel region of West Africa. Nicholson (1992) reported that the decade of 1980s was the driest of the twentieth century for the continent as a whole and for West Africa in particular. Rainfall fluctuations in the continent are similar as observed in the 1980s. There were periods when rainfall within West Africa was high as opposed to the situation in East and Southern Africa. For instance, in the 1990s and 1950s when rainfall throughout East Africa was low, the reverse occurred in West Africa (Nicholson, 1999).

### **CONCLUSIONS AND RECOMMENDATIONS**

#### **6.1 Conclusion**

Integrating Geographic Information System (GIS) with Autoregressive Integrated Moving Average (ARIMA) model for time series phenomena with spatial dimension could account for the spatial component. In this study, GIS assessment of monthly spatial distribution of From the model prediction it was observed that the catchment's rainfall will decrease further approximately by 9% within the next forty years. This indicates that the catchment's water resources would be significantly threatened and calls for prudent and pragmatic water resources management strategies. The study counteracted the findings that the end as well as the total amount of rainfall will remain more or less fixed.

#### **6.2 Recommendations**

It is recommended that the government should continuously review the catchment's precipitation to assess any variability which might not be accounted for in the current model prediction.

Similar models of the catchment's surface temperature, evapotranspiration and insolation should be built so that relationship between rainfall and these variables could be developed through regression modelling.

The Ghanaian Government should further promote and intensify storage of surface run-off in small reservoirs in the northern part of the country as part of climate change adaptation strategies as the study showed further decrease in total precipitation depth in the coming years.

## REFERENCES

- Barry BE, Obuobie E, Andreini M, Andah W, Pluquet M. 2005. Comprehensive Assessment of Water Management in Agriculture (Comparative Study of River Basin Development and Management).
- Gordon C, J. A. Amatekpor JA. (editors), 1999. The Sustainable Integrated Development of the Volta Basin in Ghana. Volta Basin Research Project, Accra, 159p.
- Leehuis C, Jung G, Kasei R, Liebe J. 2009. The Volta Basin Water Allocation System: assessing the impact of small-scale reservoir development on the water resources of the Volta basin, West Africa. *Adv. Geosci.*, 21, 57 – 62, 2009.
- Long M, Entekhabi D, Nicholson SE. 2000. Interannual variability in rainfall, water vapour flux, and vertical motion over West Africa. *Journal of Climate*, Volume 13.
- Nicholson SE, Palao IM, 1993. A Re-evaluation of Rainfall Variability in the Sahel. Part I. Characteristics of Rainfall Fluctuations. *INTERNATIONAL JOURNAL OF CLIMATOLOGY*, VOL. 13,371-389 (1993).
- Nicholson SE. 1992. An Overview of African Rainfall Fluctuations of the Last Decade *JOURNAL OF CLIMATE*, VOLUME 6
- Nicholson SE. 1999. Historical and Modern Fluctuations of Lakes Tanganyika and Rukwa and their Relationship to Rainfall Variability. *Climatic Change* 41: 53-71, 1999.
- Nicholson SE. 2001. Climatic and Environmental Change in Africa During the Last Two Centuries. *Clim Res*, Vol., 17: 123 – 144
- Nicholson SE, Barcion AI, Challa M, Baum J. 2007. Wave Activity on the Tropical Easterly Jet. *Journal of the Atmospheric Sciences*, Volume 64.
- Nicholson SE, Grist JP. 2003. The seasonal evolution of the atmospheric circulation over West Africa and Equatorial Africa. *Journal of climate*, Vol. 16, No. 7.
- Opuku-Duah S, Donoghue DNM, Burt TP. 2008. Intercomparison of Evapotranspiration over the Savannah Volta Basin in West Africa, Using Remote Sensing Data. *Sensors* 2008, 8, 2736 – 2761.
- Philippon N, Fontaine B. 2002. The relationship between the Sahelian and previous 2<sup>nd</sup> Guinean rainy seasons: a monsoon regulation by soil wetness? *Annales Geophysicae* (2002) 20, 575 – 582.
- Sabry M, Abd-El-Latif H, Badra N. 2007. Comparison between Regression and ARIMA models in forecasting traffic volume. *Aust. J. Basic & Appl. Sci.*, 1(2): 126 – 136, 2007.
- Van de Giesen N, Liebe J, Jung G. 2010. Adaptation to climate change in the Volta Basin, West Africa. *CURRENT SCIENCE*, Vol. 98, NO. 8, 25 April 2010.

Zheng X, Eltahir EAB. 1997. The response to deforestation and desertification in a model of West African monsoons. *GEOPHYSICAL RESEARCH LETTERS*, Vol. 24, No.2, Pages 155 - 158, January 15, 1997.