International Journal of Water Resources and Environmental Engineering

Volume 9 Number 9 September 2017

ISSN-2141-6613

Academic Journals

ABOUT IJWREE

The International Journal of Water Resources and Environmental Engineering is published monthly (one volume per year) by Academic Journals.

International Journal of Water Resources and Environmental Engineering (IJWREE) is an open access journal that provides rapid publication (monthly) of articles in all areas of the subject such as water resources management, waste management, ozone depletion, Kinetic Processes in Materials, strength of building materials, global warming etc. The Journal welcomes the submission of manuscripts that meet the general criteria of significance and scientific excellence. Papers will be published shortly after acceptance. All articles published in IJWREE are peer-reviewed.

Contact Us

Editorial Office: ijwree@academicjournals.org

Help Desk: helpdesk@academicjournals.org

Website: http://www.academicjournals.org/journal/IJWREE

Submit manuscript online http://ms.academicjournals.me/

International Journal of Water Resources and Environmental Engineering

Table of Contents: Volume 9 Number 9 September 2017

ARTICLES

academicJournals

Vol. 9(9), pp. 191-200, September 2017
DOI: 10.5897/IJWREE2017.0740
Article Number: 414D80C65672
ISSN 2141-6613

http://www.academicjournals.org/IJWREE

International Journal of Water Resources and Environmental Engineering

Full Length Research Paper

Application of stochastic models in predicting Lake Malawi water levels

Rodgers Makwinja[1,3]*, Titus Phiri[2,3], Ishmael B. M. Kosamu[1] and Chikumbusko C. Kaonga[1]

[1]Department of Physics and Biochemical Sciences, University of Malawi, The Polytechnic, Private Bag 303, Chichiri, Blantyre 3, Malawi.
[2]Senga Bay Fisheries Research Unit, P. O. Box 316, Salima, Malawi.
[3]Department of Fisheries, P. O. Box 593, Lilongwe, Malawi.

Received 15 July, 2017; Accepted 3 August, 2017

Stochastic models have proven to be practically fundamental in fields such as science, economics, and business, among others. In Malawi, stochastic models have been used in fisheries to forecast fish catches. Nevertheless, forecasting water levels in major lakes and rivers in Malawi has been given little attention despite the availability of ample historical data. Although previous multichannel seismic surveys revealed the presence of low stands (sediment bypass zone) in Lake Malawi indicating that since the beginning of its formation, important water level fluctuations have been occurring, these previous surveys failed to predict and highlight much more clearly the status of these levels in the future. Therefore, the main objective of the study was to fill these research gaps. The study used *Autoregressive* (AR), *Moving Average* (MA), Autoregressive Moving Average (ARMA) and *Autoregressive Integrated Moving Average* (ARIMA) processes to select the appropriate stochastic model. Based on lowest *Normalized Bayesian Information Criterion* (NBIC), *Root Mean Square Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), *Mean Forecast Error* (MFE), *Maximum Absolute Percentage Error* (MAXAPE), *Maximum Absolute Error (MAXAE)*, and *Mean Absolute Error* (MAE) - ARIMA (0,1,1) model is found suitable for forecasting Lake Malawi water levels which shows negative trend up to 2035. The study further predicted that Lake Malawi water levels will decrease from the current average level of 472.97 m to an average of 468.63 m for the next 18 years (up to 2035).

Key words: Forecasting, Lake Malawi, modelling, stochastic, time series, water levels.

INTRODUCTION

Time series stochastic process is a set of random variables $\{z_t\}$ where the index *t* takes values in a certain set C (Alonso and Garcia-Martos, 2012). The process provides attractive modeling techniques for forecasting and planning because historical data can be used to a reasonable level of certainty (Box et al., 2015). The model deals with a sequential set of data points, measured typically over successive times. It is

*Corresponding author. Email: makwinjarodgers@gmail.com.

mathematically defined as a set of vectors $x(t)t = 0,1,2$ ...where *t* represents the time elapsed (Hipel and McLeod, 1994). The variable *x(t)* is treated as a random variable and the measurements taken during an event in a time series are arranged in a proper chronological order. The principles of stochastic process are to describe and summarize time series data, fit low-dimensional models and make forecast (Box et al., 2015). Time series data have many forms and represent different stochastic processes. According to literature, *Autoregressive (AR)* and *Moving Average (MA)* models have been widely and commonly used in different fields (Box and Jenkins, 1970; Hipel and McLeod, 1994). The combination of *AR* and *MA* models forms *Autoregressive Moving Average* (ARMA). However, ARMA model only works with stationary time series data. Thus, from application viewpoint, ARMA models are inadequate to properly describe non-stationary time series, frequently encountered in practice. For this reason, the *Autoregressive Integrated Moving Average (ARIMA*) model (Box and Jenkins, 1970) is proposed. The ARIMA model is a generalisation of an ARMA model which includes the case of non-stationarity as well (Chang et al., 2012). The model was first proposed by Box and Jenkins in the early 1970s and was often termed as Box-Jenkins models (Stuffer and Dhumway, 2010). Because ARIMA model is relatively systematic, flexible and can grasp more original time series information, it is widely used in meteorology, engineering technology, marine, economic statistics, prediction technology, hydrology and water resources studies (Yevjevich, 1972; Aksoy et al., 2013; Cryer and Chan, 2008; Kantz and Schreiber, 2004).

In Malawi, ARIMA model has been commonly used in fisheries to forecast fish catches (Zindi et al., 2016; Lazaro and Jere, 2013; Singini et al., 2012; Mulumpwa et al., 2016). Nevertheless, forecasting water levels in major lakes and rivers in Malawi has been given little attention despite the availability of ample historical data. On the same note, although previous multichannel seismic surveys (Scholz and Rozendahl, 1988; Johnson and Davis, 1989; De Vas, 1994) revealed the presence of low stands in Lake Malawi indicating that since the beginning of its formation, important water level fluctuations have been occurring, these previous surveys failed to predict and highlight much more clearly the status of these levels in the future. Consequently, the present study was designed to fill these research gaps.

MATERIALS AND METHODS

Study area and physiography

The study was conducted in Lake Malawi, located at the southern end of the Great Rift Valley region. It is an elongated lake surrounded by mountains with highest elevations to the north. Figure 1 shows that the boundaries of Lake Malawi cross Mozambique and Tanzania with an outlet in the southern end. The lake is ranked as the ninth largest and third deepest freshwater lake in the world with an estimated total area of 28,750 km^2 and a volume of about 7725 km^3. The Shire River is the outlet of Lake Malawi and flows approximately 410 km from Mangochi to Ziu Ziu in Mozambique, where it drains into Zambezi River (Shela, 2000). According to Shela (2000), the physiography of upper Shire has offered opportunities for regulating river flows and subsequently lake levels, with possible expansion. The middle section of Shire River is estimated to be 80 km and is very steep characterised by rock bars and outcrops with water falls of about 370 m.

Data collection and time series model description

Lake Malawi has been there over the years. Literature has shown that in early 1924, Dixey attempted to understand the hydrology of Lake Malawi (Dixey, 1924). However, he failed due to lack of hydrological data (Dixey, 1924). Later in the years, the fear of period of no outflow by authorities greatly forced them to seriously monitor the Lake levels (Drayton, 1984). Department of Water Resources seriously embarked on collection of water levels data later in the years; however, the data collected from 1950s to somewhere around 1980s were too complex and the quality was too inconsistent. Similar observation was reported by Kaunda (2015). Because of these past data anomalies, the present study analysed the univariate time series data of Lake Malawi water levels from 1985 to 2016 period. Figure 1 shows that the Department of Water Resources collects water levels data from three stations along the lake shore (Chilumba, Nkhatabay and Monkey Bay). The water level is normally the average of three records ignoring the water level gradient which is between the north and south tip of the lake (Kumambala, 2010).

Application of stochastic models

The study used two linear time series models known as *Autoregressive* (AR) (Box and Jenkins, 1970) and *Moving Average* (MA) (Zhang, 2003) models. These models were combined to form *Autoregressive Moving Average* (ARMA) (Cochrane, 1997). The combination of these two models were based on famous Box-Jenkins principle (Box and Jenkins, 1970) also known as the Box-Jenkins models.

Autoregressive Moving Average (ARMA) Model

An ARMA (p,q) model which is a combination of AR(*p*) and MA (*q*) models was developed. In an AR (*p*) model, the future value of a variable was assumed to be a linear combination of (*p*) past observations and a random error together with a constant term. Mathematically, the AR (*p)* model (Hipel and McLeod, 1994) is expressed as

$$\boldsymbol{\gamma_t = c + \sum_{i=1}^{p} \varphi_i \gamma_{t-i} + \varepsilon_t = c + \varphi_1 \gamma_{t-1} + \varphi_2 \gamma_{t-2} .. \varphi_p \gamma_{t-p} + \varepsilon_t} \quad (1)$$

where γ_t and ε_t are the actual value and random error at time period *t*, respectively,φ_i (i=1, 2 ... *p)* are model parameters and *c* is a constant. Just as an AR (*p*) model regress against past values of the series, an MA (*q*) model uses past errors as the explanatory variables. The MA (*q*) is given by $\boldsymbol{\gamma_t}$ (Hipel and McLeod, 1994) and is expressed as:

$$\boldsymbol{\gamma_t = \mu + \sum_{j-1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} .. + \theta_q \varepsilon_{t-q} + \varepsilon_t} ... \quad (2)$$

Here, μ is the mean of the series, *j =1, 2….* are model parameters

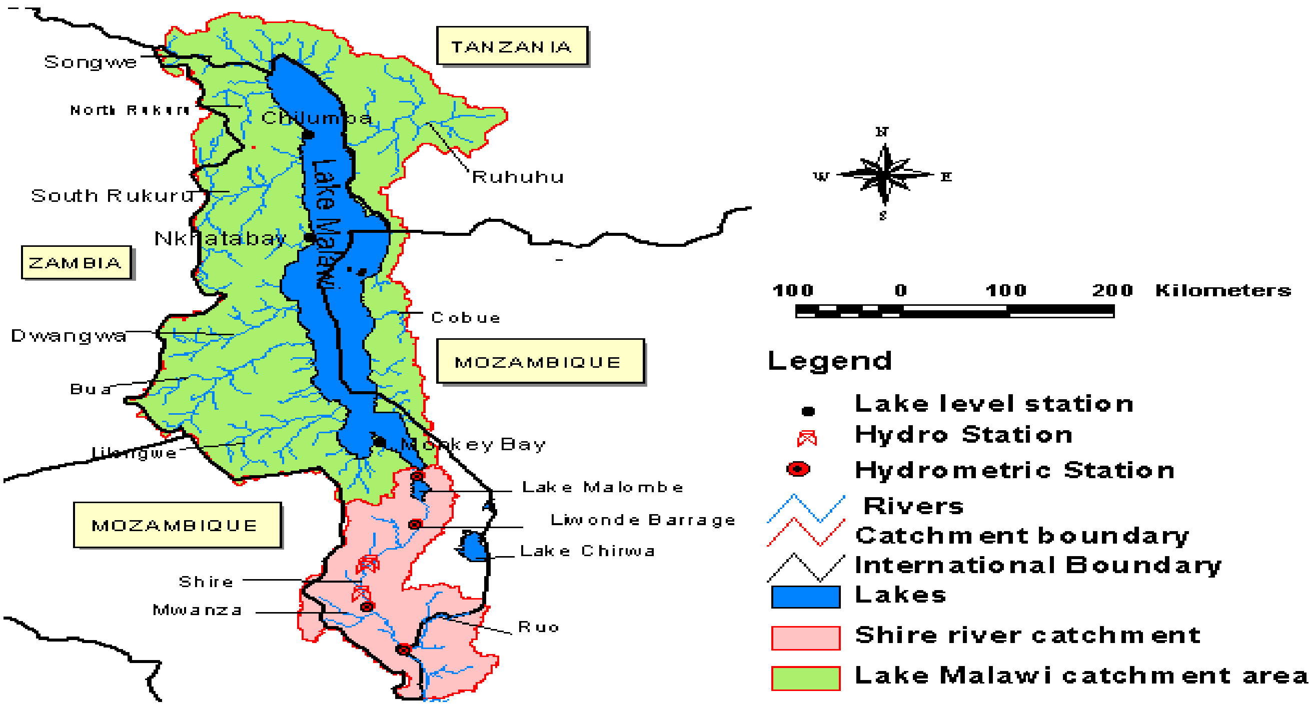

Figure 1. Map of Malawi showing Lake Malawi-Shire River system (GoM, 2005).

and *q* being the order of the model. The random shocks are assumed to be white noise (Hipel and McLeod, 1994) process. *Autoregressive* (AR) and *Moving Average* (MA) models were combined together to form a general and useful class of time series models known as the ARMA model. Mathematically, an ARMA (p, q) model is presented as (Cochrane, 1997):

$$\boldsymbol{\gamma_t = c + \varepsilon_t \sum_{i=1}^{p} \varphi_i \gamma_{t-i} + \sum_{j-1}^{q} \theta_j \varepsilon_{t-j}} \quad (3)$$

where the model orders *p, q* refers to *p autoregressive* and *q moving average* terms. Usually, ARMA models are manipulated using the lag operator notion. The lag operator is defined as $L\gamma_t = Y_{t-1}$. Polynomial of lag operators are used to represent ARMA models as follows:

AR(*p*) model: $\varepsilon_t = \varphi(L)\gamma_t$ (4)

MA(*q*) model: $\gamma_t = \theta(L)\varepsilon_t$ (5)

ARMA (*p, q*) model: $\varphi(L)\gamma_t = \theta(L)\varepsilon_t$ (6)

where

$$\varphi(L) = 1 - \sum_{i=1}^{p} \varphi_i L^i \text{ and } \emptyset(L) = 1 + \sum_{j=1}^{q} \theta_j L_j \quad (7)$$

Stationary analysis

When an AR (*p*) process is presented as: $\varepsilon_t = \varphi(L)\gamma_t$, the $\varphi(L) = 0$ is known as the characteristic equation for the process. Box and Jenkins (1970), proved that a necessary and sufficient condition for the AR (p) process to be stationary is that all roots of the characteristic equation must fall outside the unit circle. It is very important to note that ARMA models can only be used for stationary time series data. The fact that Lake Malawi water levels data was non-stationary, led to proposition of the *Autoregressive Integrated Moving Average* (ARIMA) model which is a generalization of ARMA model. In ARIMA model, non-stationary time series data is made stationary by applying finite differencing of data points (Cochrane, 1997). The mathematical formulation of the ARIMA (*p, d, q*) using lag polynomials is given below (Lombardo and Flaherty, 2000).

$$\varphi(L)(1-L)^d \gamma_t = \theta(L)\varepsilon_t, \quad (8)$$

$$\left(1 - \sum_{i=1}^{p} \varphi_i L^i\right)(1-L)^d \gamma_t = \left(1 + \sum_{j=1}^{q} \theta_j L^i\right)\varepsilon_t \quad (9)$$

here *p, d* and *q* are integers greater than or equal to zero and refer to the order of the autoregressive, integrated and moving average parts of the model, respectively. The integer *d* controls the level of differencing.

Autocorrelation (ACF) and Partial Autocorrelation (PACF)

To determine a proper model for fitting time series data, ACF and PACF analysis was carried out. These statistical measures reflected how observations in a time series data are age-related to each other. For modelling and forecasting purposes, ACF and PACF against consecutive time lags were plotted. These plots helped to determine the order of AR and MA terms. Below are the mathematical models: For a time, series $\{x(t), t = 0,1,2 \ldots.\}$ the autocovariance at lag k is defined as:

$$\gamma_k = Cov(x_t x_{t+k}) = E[(x_t - \mu)(x_{t+k} - \mu)] \quad (10)$$

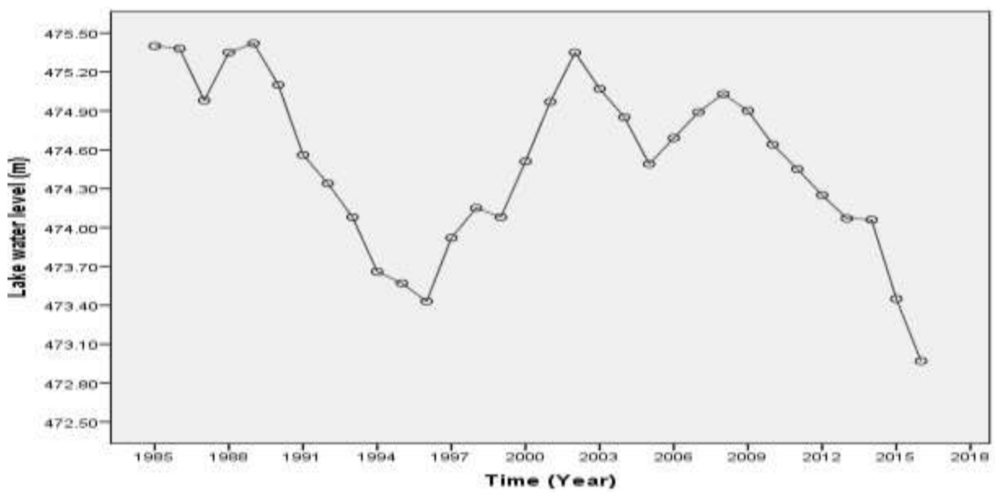

Figure 2. Water levels of Lake Malawi from the period of 1985 to 2016.

The autocorrelation coefficient at lag *k* is defined:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \tag{11}$$

where μ is the mean of the time series, that is, $\mu = E[x_t]$. The autocovariance at lag zero, that is, γ_0 is the variance of the time series. Another measure, known as the *Partial Autocorrelation Function* (PACF) is described by Box and Jenkins (1970). It is used to measure the correlation between an observation *k* past period and present observation after controlling observations at intermediate lags.

Trend model fitting

Conducting various diagnostic tests is an important step in time series modeling (Chung, 2009). The famous Box-Ljung Q-statistics as described by Box and Jenkins (1970) was used to transform the non-stationary data into stationary and to check adequacy for the residuals. In practice, the Box-Ljung Q-statistics was computed (Ljung and Box, 1978) as

$$Q = n(n+2)\sum_{k=1}^{m}\frac{\hat{r}_k^2}{n-k} \tag{12}$$

where $\hat{r}_k$ is the estimated autocorrelation of the series at lag *k* and *m* is the number of lags being tested. Box and Jenkins (1970) developed a practical approach to build ARIMA model, which best fit a given time series and also satisfy the parsimony principle. According to Box and Jenkins (1970), the three-step approach of *model identification*, *parameter estimation* and *diagnoststic checking* to determine the best persimonious model from general class of ARIMA models (Zhang, 2003) were applied. The three-step process was repeated several times until a satisfactory model was finally selected. The appropriate model selection step is very critical. It is based on the fact that sample ACF and PACF, calculated from the training data should match with the corresponding theoretical or actual values (Chatfield, 1996). In this case, various model fitting statistics like Root Mean Square Error (RMSE), Maximum Absolute Percentage Error (MAXAPE) and Maximum Absolute Error (MAXAE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Forecast Error (MFE) and Bayesian Information Criterion (BIC) were employed to evaluate the adequacy of AR, MA and ARIMA processes. Based on Normalized BIC, the principle is that the lower the value, the better the model. Fit statistics such as MAPE, MAE, MFE, BIC and RMSE were calculated as shown below:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\gamma_i - \bar{y}_i}{\gamma_i}\right| \tag{13}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\gamma_i - \bar{y}_i| \tag{14}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\gamma_i + \bar{y}_i)^2} \tag{15}$$

$$MFE = \frac{1}{n}\sum_{t=1}^{n} e_t \tag{16}$$

$$BIC(p) = n\ln\left(\hat{\sigma}_e^2/_n\right) + P + P\ln(n) \tag{17}$$

where, γ_i and $\bar{y}_i$ are actual observed and predicted values respectively, while *n* is number of predicted values. In BIC model, *n* is the number of effective observations used to fit the model, *p* is the number of parameters in the model and $\hat{\sigma}_e^2$ is the sum of sample squared residuals. Upon identification of optimum model, forecast of the Lake Malawi water levels from 2017 to 2035 were made.

All inferential and descriptive statistics were performed using International Business Management Statistical Package for Social Scientists software (IBM SPSS 20) (IBM Corp, 2011).

RESULTS AND DISCUSSION

Model selection

The stationarity of a stochastic process was visualized in form of a data plot as shown in Figure 2. According to Hipel and McLeod (1994), identification of stationarity in time series data is a necessary condition for building a time series model that is useful for forecasting. Sankar (2011) defined time series stationarity as a set of values that vary over time around a constant mean and constant variance.

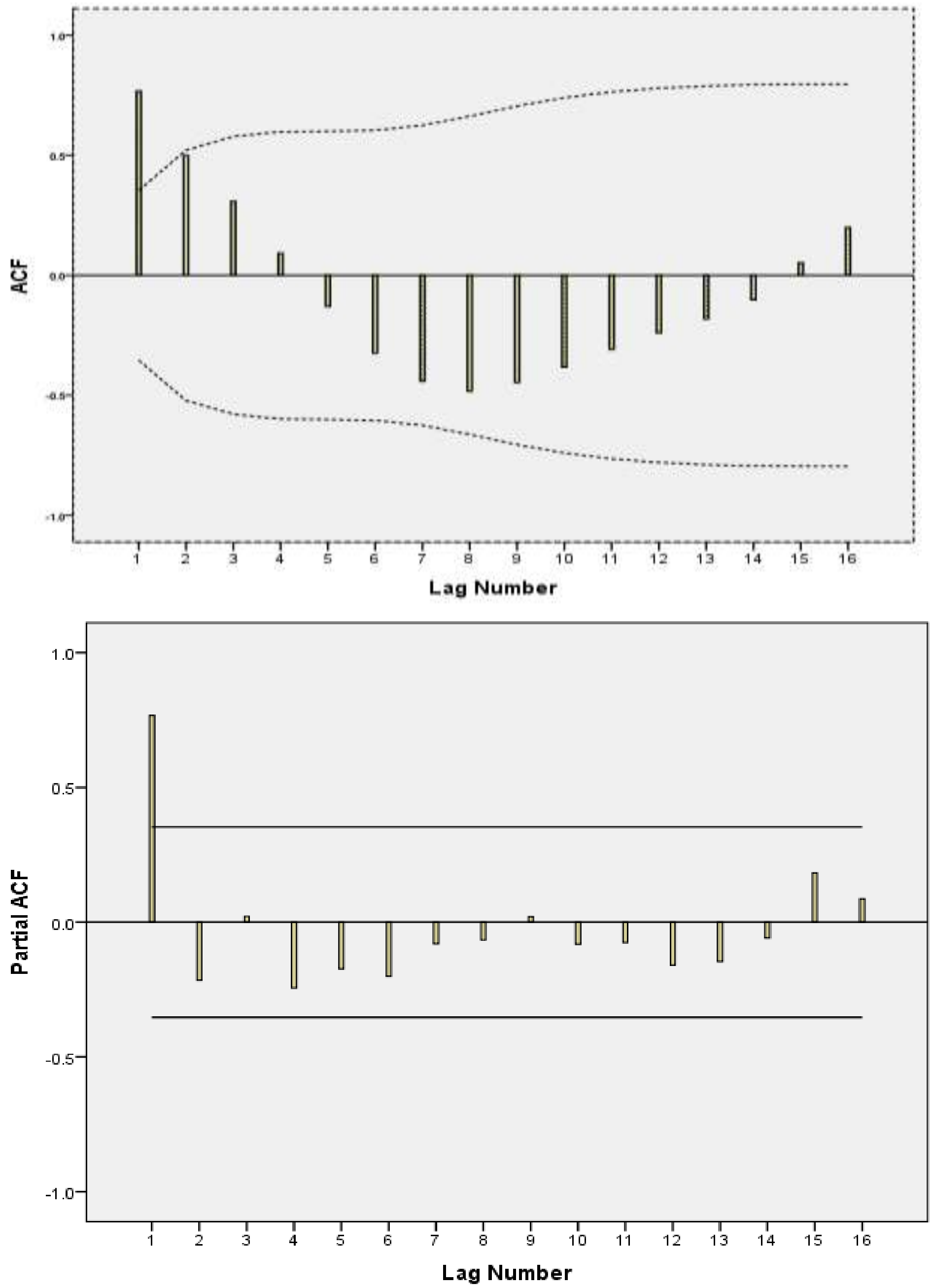

Figure 3. Autocorrelograms and partial autocorrelograms of first order differenced data.

According to Hipel and McLeod (1994), time series data showing seasonal patterns are usually non-stationary in nature. From Figure 2, it is very apparent that the time series data from Lake Malawi water levels is non-stationary due to unstable means which increase and decrease at some points throughout 1985 to 2016. Similar observation was reported by several authors in Lake Malawi (Lazaro and Jere, 2013; Singini et al., 2012; Mulupwa et al., 2016; Zindi et al., 2016). Given these difficulties in Lake Malawi water levels time series data, first order differencing of the data and stationary test were conducted on the newly constructed series of the data. Since the newly constructed data was stationary in mean, the next issue was how to select an appropriate model that can produce accurate forecast based on the description of historical pattern in the data and how to determine the optimal model order. In this case, the values of p and q in the ARIMA model were identified by plotting autocorrelogram and partial autocorrelogram presented in Figure 3.

Figure 3 illustrated that autoregressive model of order p(AR (q)) was stationary and moving average model of

Table 1. ACF and PACF for time series data of Lake Malawi water levels.

Lag	ACF	Std Error	Box-Ljung Statistic			PACF	Std Error
			Value	df	Sig		
1	0.767	0.177	20.647	1	0.000*	0.767	0.177
2	0.499	0.261	29.694	2	0.000*	-0.216	0.177
3	0.309	0.289	33.267	3	0.000*	0.021	0.177
4	0.092	0.299	33.598	4	0.000*	-0.244	0.177
5	-0.130	0.300	34.276	5	0.000*	-0.173	0.177
6	-0.323	0.302	38.646	6	0.000*	-0.200	0.177
7	-0.441	0.313	47.112	7	0.000*	-0.080	0.177
8	-0.483	0.331	57.673	8	0.000*	-0.066	0.177
9	-0.446	0.353	67.094	9	0.000*	0.020	0.177
10	-0.383	0.370	74.331	10	0.000*	-0.082	0.177
11	-0.308	0.382	79.239	11	0.000*	-0.076	0.177
12	-0.239	0.390	82.349	12	0.000*	-0.159	0.177
13	-0.181	0.394	84.235	13	0.000*	-0.147	0.177
14	-0.102	0.397	84.869	14	0.000*	-0.059	0.177
15	0.052	0.398	85.044	15	0.000*	0.182	0.177
16	0.200	0.398	87.753	16	0.000*	0.086	0.177

[ns]: Non-significant, *, **: Significant at P<0.01, and P < 0.05, respectively.

Table 2. Fit statistics for various competing ARIMA models.

ARIMA (p,d,q)	RMSE	MAPE	MAXAPE	MAE	MAXAE	MFE	NBIC
ARIMA (1,1,0)	0.28	0.04	0.12	0.19	0.53	0.38	-0.21
ARIMA (1,1,2)	0.29	0.05	0.13	0.21	0.63	0.52	-1.87
ARIMA (1,1,1)	0.29	0.05	0.12	0.22	0.58	0.46	-2.00
ARIMA (0,1,1)	0.29	0.05	0.12	0.22	0.59	0.54	-2.12
ARIMA (2,1,2)	0.31	0.05	0.13	0.22	0.13	0.49	-1.70

order q(MA (q)) was good. Guti´errez-Estradade et al. (2004) explained that a good autoregressive model of order p(AR (q)) has to be stationary and a good moving average model of order q(MA (q)) has to be invertible. The invertibility and stationarity gives a constant mean, variance and covariance which is a necessary condition for forecasting (Singini et al., 2012). Following Hipel and McLeod (1994), autocorrelation and partial autocorrelation coefficients (ACF and PACF) of up to 16 lag were considered. The type and order of the adequate model required to fit the series was determined. As the ACF values diminished rapidly with increasing lags, it was assumed that lynx series was stationary. The autocorrelation and partial autocorrelation coefficients (ACF and PACF) of various orders of differenced series of data were computed and presented in Table 1. The basic principle of model parsimony states that the model with smallest number of parameters is to be selected so as to provide an adequate representation of the underlying time series (Chatfield, 1996). In other words, out of a number of suitable models, it is very important to consider the simplest model while still upholding an accurate description of inherent properties of the time series (Zhang, 2007).

As discussed by Hipel and McLeod (1994), a number of ARIMA models were competed in order to select the simplest one as shown in Table 2. Hipel and McLeod (1994) observed that the more complicated the model, the more possibilities will arise for departure from actual model assumptions. In other words, with the increase of model parameters, the risk of model overfitting also subsequently increases. Although over fitted time series models describe the data very well, it may not be suitable for future forecasting. Therefore, genuine attention was given to select the most parsimonious model among all other possibilities. Using the coefficients in Table 1, various ARIMA models were identified and the models together with their corresponding fit statistics are presented in Table 2. The Root Mean Square Error (RMSE) which measured how much dependent series varies from its model-predicted level was lowest (0.28) in ARIMA (1,1,0) model which according to Cao and Francis

Table 3. Lake Malawi water levels estimated ARIMA model.

Parameter	Estimate	Std Error	t-value	p-value
Constant	10.56	0.41	0.59	0.56^{ns}
AR	0.38	0.18	2.13	0.04**

^{ns}: Non-significant, *, **: Significant at P<0.01 and P < 0.05, respectively.

(2003), indicated a good forecast of the model. Similarly, Mean Absolute Error (MAE) also known as Mean Absolute Deviation was lowest (0.19) in ARIMA (1,1,0) model which indicated a good forecast of the model. In other words, the magnitude of overall error occurring due to forecasting was very small.

It was further noted that Mean Absolute Percentage Error (MAPE) was lowest and smallest in ARIMA (1,1,0) meaning that the percentage of average absolute error occurring was very small. In other words, the opposite signed errors did not offset each other. It was further interesting to note that Maximum Absolute Percentage Error (MAXAPE) and Maximum Absolute Error (MAXAE) expressed as percentage was very small in ARIMA (1,1,0) model indicating overall good model fit. According to Czerwinski et al. (2007), the best model should have adequate accuracy measures (RMSE, MAE) and lowest Normalised BIC for it to have accurate forecasts. Therefore, ARIMA (1,1,0) model was selected because it had lowest RMSE, MAE, MFE and Normalized Bayesian Information Criterion (NBIC). It was further observed that the coefficients of the parameters of ARIMA (1,1,0) model were significant. According to Czerwinski et al. (2007), the model which indicate lowest normalized BIC and is significant (p<0.05) is a better model in terms of forecasting performance than with large normalized BIC. Estimates of the selected ARIMA (1,1,0) model are presented in Table 3.

Based on the study findings, the most suitable model for forecasting Lake Malawi water levels was confirmed to be ARIMA (1,1,0).

Model systematic checks

The basic model verification is concerned with checking the residues to see if they contain any systematic pattern which could still be eliminated to improve the performance of the selected model. Therefore, the selected ARIMA (1,1,0) model was subjected to autocorrelations and partial autocorrelations of residues of various orders. Various autocorrelations of up to 24 lags were computed and plotted as shown in Figure 4. The results showed that none of the autocorrelation was significantly different from zero at any reasonable level. This implied that the selected ARIMA (1,1,0) model was an appropriate model for forecasting Lake Malawi water levels.

It is very apparent from Figure 4 that autocorrelations of the coefficients are within 95% confidence interval, suggesting that the selected model was well fitted in time series model and had an accurate forecast.

Forecasting

Using the selected ARIMA (1,1,0) model, the forecast of Lake Malawi water levels was made from 1985 to 2035.

For preciseness and accurateness sake, only observations from 2005 to 2016 were compared with the forecasted values as shown in Table 4. Figure 5 on the other hand, indicates the forecasted value from 1985 to 2035. Czerwinski et al. (2007) explained that the forecasted and actual values need to be very close, meaning that the forecasting error must be very low for the model to qualify as good. As observed in Table 4, the noise residues were a combination of positive and negative errors indicating that the model had a good performance of forecasting. It was further interesting to note that the magnitude of the difference between the forecasted and actual values were very low indicating a good forecasting performance. In Figure 5, it is very apparent that Lake Malawi water levels are fluctuating with a negative trend. Such negative trend will continue up to 2035.

Figure 5 further indicated that values for water levels increased during 2006 to 2010 and decreased up to 2015 when compared to values of 2005. However, the trend declined continuously up to 2035. The basic principle of ARIMA model assumes that time series data is linear and follows a particular known statistical distribution such as normal distribution (Cochrane, 1997). Therefore, it may be concluded that the trend in this study behaved in a manner consistent with ARIMA principle which is assumed to follow a certain probability model described by joint distribution of random variable. It is also interesting to note that time series is non-deterministic in nature such that it cannot predict with certainty what will occur in the future. Based on this observation, the study indicated that there is high probability that Lake Malawi water levels will decrease as far as up to 468.63 m by 2035. Kidd (1983) had similar observation in 1915 and recorded the lowest lake level of 469 m above sea level. Drayton (1984) in the 1980s reported that Lake Malawi water levels have been unstable over the years with notable events occurring in 1890s where unusual low water levels (112 m) were recorded. He further noted that the Lake water levels were near cessation of outflows for more than 20 years (from 1890s to 1935) and experienced high levels and outflows in 1970s and 1980s which caused flooding of lakeshore communities and areas immediately downstream. Kidd (1983) earlier noted that a small decrease in the ratio resulted in the basin being closed with no outflow as occurred between 1915 and 1937. Recently, Shela (2000) observed unusual low level (115 m) and outflows in the 1990s which was further

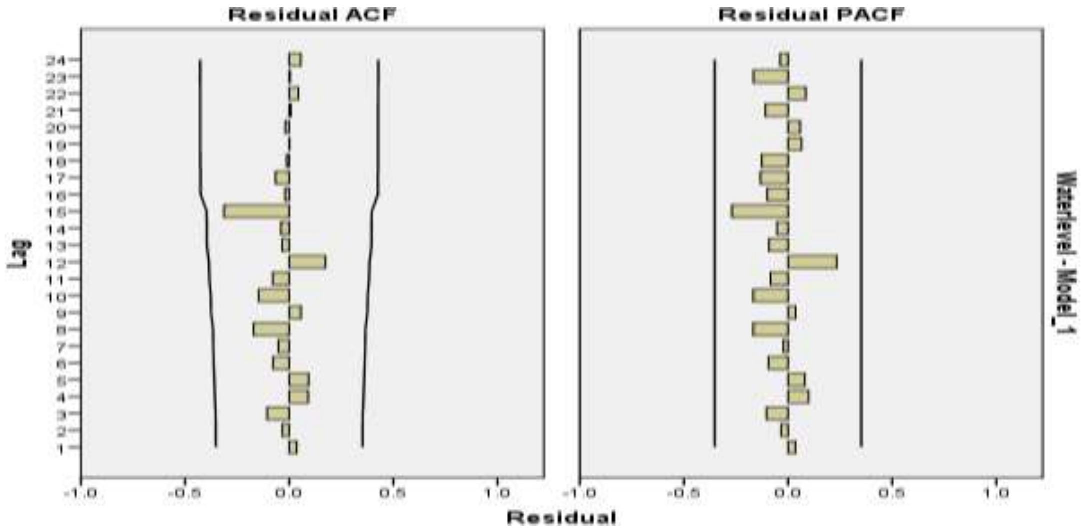

Figure 4. ACF and PACF residue.

Table 4. Forecasted Lake Malawi water levels.

Year	Actual water level (m)	Predicted water level (m)	95% confidence interval
2005	474.49	474.70	(476.0, -475.6)
2006	474.69	474.28	(-475.6, 474.4.4)
2007	474.89	474.69	(-474.5, 373.4)
2008	475.03	474.89	(-134.31, 124.32)
2009	474.90	475.00	(-402.1, 470.29)
2010	474.64	474.77	(-373.69, 374.88)
2011	474.45	474.45	(-171.1, 170.29)
2012	474.25	474.29	(-804.29, 815.48)
2013	474.07	474.08	(102.41, 105.6)
2014	474.06	473.90	(-604.17, 575.05)
2015	473.45	473.96	(-073.86, 075.05)
2016	472.97	473.11	(-408.69, 414.88)
2017		472.68	(-106.48, 124.68)
2018		472.46	(-102.31, 114.5)
2019		472.26	(-401.36, 404.55)
2020		472.07	(-102.52, 203.71)
2021		471.87	(-108.09, 103.27)
2022		471.68	(-071.45, 073.47)
2023		471.47	(-570.9, 573.62)
2024		471.27	(-010.42, 013.78)
2025		471.05	(-069.97, 073.780
2026		470.84	(-409.55, 407.81)
2027		470.61	(-468.14, 473.81)
2028		470.38	(-1068.75, 1473.8)
2029		470.15	(-106.36, 107.76)
2030		469.91	(-401.98, 423.71)
2031		469.67	(-132.08, 132.4)
2032		469.42	(-246.21, 246.2)
2033		469.16	(-187.63, 179.2)
2034		468.90	(-465.03, 472.7)
2035		468.63	(-1464.6, 1473.7)

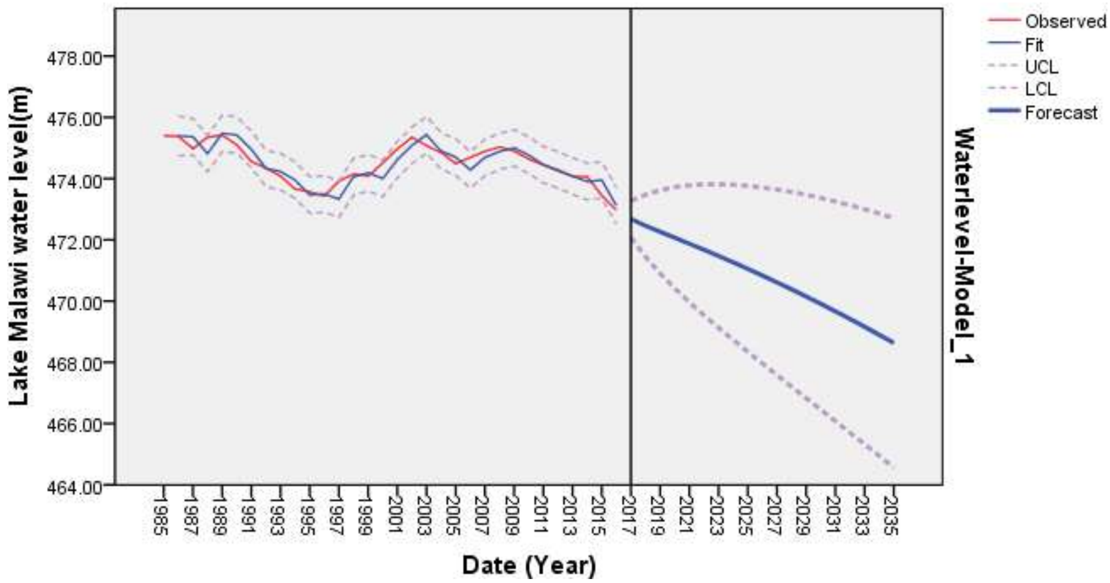

Figure 5. Actual and forecasted Lake Malawi water level.

associated with a widespread regional drought. The study by Neuland (1984) also revealed that there is little risk of the lake level exceeding 477.8 m above mean sea level. Using the most recent observed climatic parameters of the lake, the predicted level by Neuland (1984) remains below 477 m and further indicated a high probability of negative trend of future water levels as reported in the present study. Kumambala and Ervine (2010) further added that it is very unlikely for the water level to increase to a maximum height of 477 m as it was in 1980. Recent prediction by Kaunda (2015) indicated that near future and far future projects show that water yield will decrease by 8.84% and therefore Lake Malawi water level is expected to drop. However, Kaunda findings were thus on short term from 2017 to 2020. Following the dramatic rise in lake level in 1979, Drayton (1979) made a statistical analysis of lake levels and recommended a “safe" static level of 477.6 m ASVD for the next 30 years. Nonetheless, the negative trend of Lake Malawi water levels predicted in the present study is worrisome. With such future prediction, deliberate effort has to be made to find appropriate policy options and strategies for sustaining Lake Malawi water levels.

Conclusion

The study selected the ARIMA (0, 1, 1) model for forecasting Lake Malawi water levels. The ARIMA (0, 1, 1) had lowest Normalized Bayesian Information Criterion (NBIC), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Forecast Error (MFE) and Mean Absolute Error (MAE) which indicated a good forecast of the model. Based on the selected model, it is very apparent that Lake Malawi water levels fluctuation is showing a negative trend. Such negative trend is predicted to continue up to 2035. The model further predicted that Lake Malawi water levels will decrease up to 468.63 m by 2035. This study provides critical information for future policy making and formulation of intervation strategies for sustaining Lake Malawi water levels.

RECOMMENDATION

The major limitation of ARMA and ARIMA models in this study was that they only capture short-range dependence (SRD). In other words, they belong to the conventional integer models. In practice, several time series exhibit long range dependence (LRD) in their observations. To overcome this difficulty, it is recommended that a similar study should be conducted using *Autoregressive Fractionally Integrated Moving Average* (ARFIMA) model with ability to capture long range property of the fraction system accordingly and project extended period of more than 18 years.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENT

The authors wish to thank Department of Water Resources of the Ministry of Agriculture, Irrigation and

Water Development for the kind support in providing historical data for Lake Malawi water levels. Many thanks to Mayamiko Chaweza for assisting in data analysis.

REFERENCES

Alonso A, Garcia-Martos C (2012). Time series and stochastic processes. Madrid: Universidad Carlos III de Madrid. Available online at www.etsii.upm.es/ingor/estadistica/Carol/TSAtema3petten.pdf

Aksoy H, Unal N, Eris E, Yuce M (2013). Stochastic modeling of Lake Van water level time series with jumps and multiple trends. Hydrol. Earth Syst.Sci. 17:2297-2303.

Box G, Jenkins G (1970). Time Series Analysis, Forecasting and Control. San Francisco: Holden-Da. Available online at www.library.wiley.com/doi/10.1002/9780470316566.refs/pdf

Box G, Jenkins G, Reinsel G, Ljung G (2015). Time Series Analysis: Forecasting and Control. Hoboken, NJ, USA: John Wiley and Sons.

Cao L, Francis E (2003). Support Vector Machine with Adaptive Parameters in Financial Time Series Forecasting. IEEE Trans. Neural Networks 14(6):1506-1518.

Chang X, Gao M, Wang Y, Hou X (2012). Seasonal Autoregressive Integrated Moving Average Model For Precipitation Time Series. J. Maths. Stat. 8(4):500-505.

Chatfield C (1996). Model uncertainty and forecast accuracy. J. Forecasting pp. 495-508.

Chung M (2009). Lecture on time series diagnostic test. Taipei 115, Taiwan: Institute of Economics, Academia Sinica. Available online at www.ntu.edu.tw/~ckuan/pdf/Lec-DiagTest_0902.pdf

Cochrane H (1997). Time Series for Macroeconomics and Finance. Chicago: Graduate School of Business, University of Chicago, Spring. Available online at www.bseu.by/russian/faculty5/stat/docs/4/Cochran,TimeSeries.pdf

Cryer J, Chan K. (2008). Time Series Analysis with Application in R 2nd ED. New Yolk: Springer.

Czerwinski I, Guti´errez-Estrada J, Hernando-Casal J (2007). Short-term forecasting of halibut CPUE: Linear and non-linear univariate approaches. Fisheries Res. 86:120-128.

De Vas A (1994). Pliocene en Kwartaire evolutie van het Livingstone bekken (Malawi rift, Tanzanie) afgeleid uithoge-resolutie reflectieseismische profielen.Licenciaat theis. Belgium: s, Universiteit Gent. Avaliable online at https://www.gadventures.com/trips/victoria-falls-serengeti-adventure.../DVN/

Dixey F (1924). Lake level in relation to rainfall and sunspots. Nature 114:659-661.

Drayton R (1979). A study of the causes of the abnormally high levels of Lake Malawi. Lilongwe: Wat. Resour. Div. Tech. Pap.No 5.

Drayton R (1984). Variation in the level of Lake Malawi. J.Hydrol.Sci. 29:1-12.

GoM (2005). National Spatial Data Center, Ministry Lands Housing physical Planning and Surveys, Lilongwe, Malawi. Available online at www.lands.gov.mw/index.php/contacts/physical-planning.ht

Guti´errez-Estradade J, L´opez-Luque E, Pulido-Calvo I (2004). Comparison between traditional methods and artificial neural networks for ammonia concentration forecasting in an eel (Anguilla Anguilla L.) Intensive rearing system. Aquat. Eng. 31:183-203.

Hipel K, McLeod A (1994). Time Series Modelling of Water Resources and Environmental Systems. Amsterdam: Elsevier 1994.

IBM Corp (2011). IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp. Available online at www.sciepub.com/reference/159832

Johnson T, Davis T (1989). High resolution profiles from Lake Malawi, Africa. J. Afr. Earth. Sci. 8:383-392.

Kantz H, Schreiber T (2004). Nonlinear Time Series Analysis 2nd Edn. Cambridge: Cambridge University Press.

Kaunda P (2015). Investigating the Impacts of Cliamte Change on the Levels of lake Malwi Thesis, Department of meteology, university of nairobi, Kenya.

Kidd C (1983). A water resources evaluation of Lake Malawi and the Shire River. Geneva:UNDP project MLW/77/012, World Meteological Organisation.

Kumambala P, Ervine A (2010). Water Balance Model of Lake Malawi and its Sensitivity to Climate Change. Open. Hydrol. J. 4:152-162.

Kumambala P (2010). Sustainability of water resources development for Malawi with particular emphasis on North and Central Malawi. PhD thesis. Glasgrow, UK: University of Glasgrow. Available on line at https//theses.gla.ac.uk/1801/1/2010Kumambalaphd.pdf

Lazaro M, Jere W (2013). The Status of the Commercial Chambo (Oreochromis Species) fishery in Malaŵi: A Time Series Approach. Intl. J. Sci. Technol. 3:1-6.

Lombardo R, Flaherty J (2000). Modelling Private New Housing Starts In Australia. Pacific-Rim Real Estate Society Conference Sydney: University of Technology Sydney (UTS) pp. 24-27.

Ljung G, Box G (1978). On a measure of lack of fit in time series models. J. Bio. 65:297-303.

Mulumpwa M, Jere W, Mtethiwa A, Kakota T, Kang'ombe J (2016). Application Of Forecasting In Determining Efficiency Of Fisheries Management Strategies Of Artisanal Labeo mesops Fishery Of Lake Malawi. Int. J. sci. Tech. Resea 5(09):28-40.

Neuland H (1984). Abnormal high water levels of Lake Malawi? -An attempt to assess the future behaviour of the lake water levels. Geo. J. 9:323-334.

Sankar I (2011). Stochastic Modelling Approach for forecasting fish product export in Tamilnadu. J. R .Resea. Sci.Technol. 3(7):104-108.

Scholz E, Rozendahl B (1988). Low lake stands in Lakes Malawi and Tanganyika, East Africa, delineated from multifold seismic data. J. Sci. 240:1645-1648.

Shela O (2000). Naturalisation of Lake Malawi Levels and Shire River Flows: Challenges of Water Resources Research and Sustainable Utilisation of the Lake Malawi-Shire River System. Sustainable Use of Water Resources Maputo: 1st WARFSA/WaterNet Symposium, pp. 1-12. Available online at https://floodobservatory.colorado.edu/SiteDisplays/Shire2SHELA.PDF

Singini W, Kaunda E, Kasulo V, Jere W (2012). Modelling and Forecasting Small Haplochromine Species (Kambuzi) Production in Malaŵi-A Stochastic Model Approach. Intl. J. Sci. Technol. Resea 1:69-73.

Stuffer D, Dhumway R (2010). Time series Analysis and its Application 3rd Ed. New Yolk: Springer.

Yevjevich V (1972). Probability and Statistics in Hydrology. Colorado: Water Resources Pub. Available online at https://www.abebooks.com/.../probability-and-statistics-in-hydrology/.../yevjevich-vu

Zhang G (2007). A neural network ensemble method with jittered training data for time series forecasting. J. Infor. Sci. 177:5329-5346.

Zhang G (2003). Time series forecasting using a hybrid ARIMA and neural network model. J. Neuroc. 50:159-175.

Zindi I, Singini W, Mzengeleza K (2016). Forecasting Copadichromis (Utaka) Production for Lake Malaŵi, Nkhatabay Fishery- A Stochastic Model Approach. J. Fish. Livest. Prod.4:162.

academicJournals

Vol. 9(9), pp. 201-210, September 2017
DOI: 10.5897/IJWREE2015.0580
Article Number: FCC90B165676
ISSN 2141-6613

http://www.academicjournals.org/IJWREE

International Journal of Water Resources and Environmental Engineering

Full Length Research Paper

Assessment of groundwater suitability for irrigation in three sub catchments in Upper Athi River Basin, Kenya

Philip Kwabla Attibu[1*], Ebenezer Ashun[1] and George Lartey Young[2]

[1]Department of Geography, Kenyatta University, P. O. Box 43844 - 00100, Nairobi, Kenya.
[2]Department of Environmental Science, University College of Agriculture and Environmental Studies, P. O. Box 27, Bunso, Eastern Region, Ghana.

Received 28 April, 2015; Accepted 1 March, 2017

Over the years, erratic rainfall pattern coupled with increasing population has led to the reliance on groundwater as an alternative and sustainable source for crop irrigation to meet increasing food demand. Irrigation of crops though essential, if not done with care through good practices and use of quality water can lead to soil salinization and ecological unsustainability. This study was carried out to assess the salinity of groundwater used for irrigation in three sub catchments in the Upper Athi River Basin of Kenya. Stratified random sampling technique was used to select representative boreholes and shallow wells for the study. In all, water from 17 boreholes and 17 shallow wells spread across the study area were sampled and analysed for selected physico-chemical properties. Standard methods were used for all the laboratory analysis; temperature, pH and electrical conductivity (EC) of water samples were measured in the field. The results obtained were compared with FAO Water Quality Guidelines for Irrigation. pH ranged from 4.2 to 7.13 indicating weak acidity with about 75% samples falling below lower guideline value. EC values ranged from 467 to 1328 µS/cm which were within FAO and NEMA permissible limits for irrigation purposes. All salts ions were within permissible irrigation water suitability standards except CO_3^-, Cl and K^+. 97% samples had above the recommended carbonate concentrations while 80% had more potassium than the recommended value with the remaining 20% being boreholes. In relation to chloride concentrations, samples from shallow wells are not suitable for sprinkler irrigation since they were above the recommended levels; however, 58% were suitable for surface irrigation. For the boreholes, chloride concentrations were suitable for both surface and sprinkler irrigation. In conclusion, the boreholes had less ions as compared to shallow wells. This research may serve as a preliminary study to provide baseline information that may direct future water quality assessment studies in the study area.

Key words: Irrigation, groundwater quality, boreholes, shallow wells, physico-chemical quality, salinity, sodicity.

INTRODUCTION

Water scarcity for crop irrigation has been considered as a global problem. Irrigation water, irrespective of its source whether diverted from springs, pumped from wells or diverted from streams directly, do contain some

*Corresponding author. E-mail: attibuphil@gmail.com.

amount of chemical substances in solution which has the potential to degrade the quality of soils and reduces crop yield (Banderi et al., 2012). All irrigation waters contain soluble salts and this has become a noted problem associated with farm lands irrigated in arid regions of the world with about one third of agricultural lands becoming saline and this extends to more than one hundred countries of different climates (Squires and Glenn, 2004). Land degradation has been estimated to be about 65% of agricultural lands in Africa, 45% in South America, 74% in Central America and 35% in Asia (CGIAR, 2003). According to Ghassemi et al. (1995), an estimated 1.2 tonnes of salt is added to every hectare of soil per year from irrigation activities.

In Kenya, approximately 40% (25 million hectares) of the land have high salt concentrations (Wanjogu et al., 2004). According to Mugwanja et al. (1995), 26,000 hectares of irrigated soils in Kenya are considered 'salts-affected' due to poor quality of water, poor drainage and irrigation management systems, especially in areas with high or increasing ground water tables. Arid and semi-arid lands cover approximately 80 to 83% of the country and receive less than 700 mm rainfall per year which is erratic, poorly distributed and cannot reliably support rain fed agriculture (Ndegwa and Kiiru, 2009). Irrigation schemes in the country have been abandoned after less than twenty years of operation due to salinization and many more will follow if preventive and mitigation measures are not put in place (Stein and Schulze, 1978) as cited by Wanjogu et al. (2004). Agriculture in Kenya contributes to 55% of GDP, providing about 80% of employment. It accounts for 60% of export and generates about 40% of Government revenue (Blank et al., 2002). Irrigation fed agriculture directly contributes 3% of the total GDP and provides 18% of the value of all agricultural produce (ROK, 2009).

Nevertheless, the agricultural sector has been constrained with the following: salinity, sodicity, drain ability, effective rooting depth, availability of oxygen for root growth, workability of the soil, water retention capacity, moisture availability, soil fertility, conditions of germination, ease of land clearing and freedom for layout of field plans (Wanjogu et al., 2004). Because of the erratic rainfall patterns in the study area, both large and small scale farms are exploiting groundwater resources for irrigation without paying attention to the quality and associated impacts on land and crop losses. This research sought to investigate the quality of groundwater used for irrigation purposes in three catchment areas (Theta, Thiriika and Rwabura) of the Upper Athi River basin.

Study area

Geographic settings

The study was carried out in three sub catchments: Theta, Thiririka and Rwabura within the Upper Athi River Basin. These three adjacent sub catchments are located in Gatundu south constituency within Kiambu County of Kenya. The study area lies between Latitudes 0° 51' 22'' and 01° 09' 25'' S and Longitudes 36° 34' 59'' and 37° 02' 10'' E. Its covers approximately 165 km^2 and is bounded by Kikuyu escarpment to the North and Kiaora estate to the South. Location map of the study area is presented in Figure 1.

Topography

The altitude of the study area ranges from 1600 m above sea level at the lower zone to 2200 m above sea level in the middle and upper zones. It is hilly to the north and west and scattered hilly in the central and southern parts, while the East and South East has gentle plains. There are several valley bottoms scattered all over the centripetal drainage system, draining into the Athi River Basin. The terrain at lower parts favours gravity irrigation system which is currently being exploited on a gradual basis.

Climate

The sub catchments lie within the humid and semi humid agro-climate zones of Kenya. The upper parts of Thiririka and Rwabura sub-catchments lay within the Kikuyu escarpment forest comprising the humid zone and are the source of river Thiririka and Rwabura while that of Theta sub catchment lies within the Aberdare forest and also a humid zone and a source to river Theta and other streams. The middle parts of the sub catchments comprise the sub humid and semi humid zones and provides agricultural land where small scale agricultural activities are undertaken. The rainfall pattern is bimodal with two distinct rainy seasons with long and high rainfalls in March and May with short rain fall between October and November. The rainfall received ranges between 800 and 2000 mm with the highest being at the tea production zones. The maximum and minimum rainfall received is 257 and 33.4 mm in April and July, respectively (Gatundu Agricultural office and Rwabura Irrigation Project Report, 2013). The mean monthly evaporation experienced in the sub catchments range from 1.6 to 6.6 mm/day. On the average, the minimum evaporation rate is 75 mm in July, while the maximum is 166.6 mm in March (Rwabura Irrigation Project Report, 2013). Temperature distribution varies from humid to semi humid with the upper zone experiencing a mean annual temperature between 14^{o}C and 18^{o}C and the lower zones 18 ^{o}C and 22 ^{o}C. The maximum temperatures range from 25 to 34°C in August and March, whereas the minimum ranges from 9.8 to 15.4°C in February and April, respectively (Gatundu Agricultural

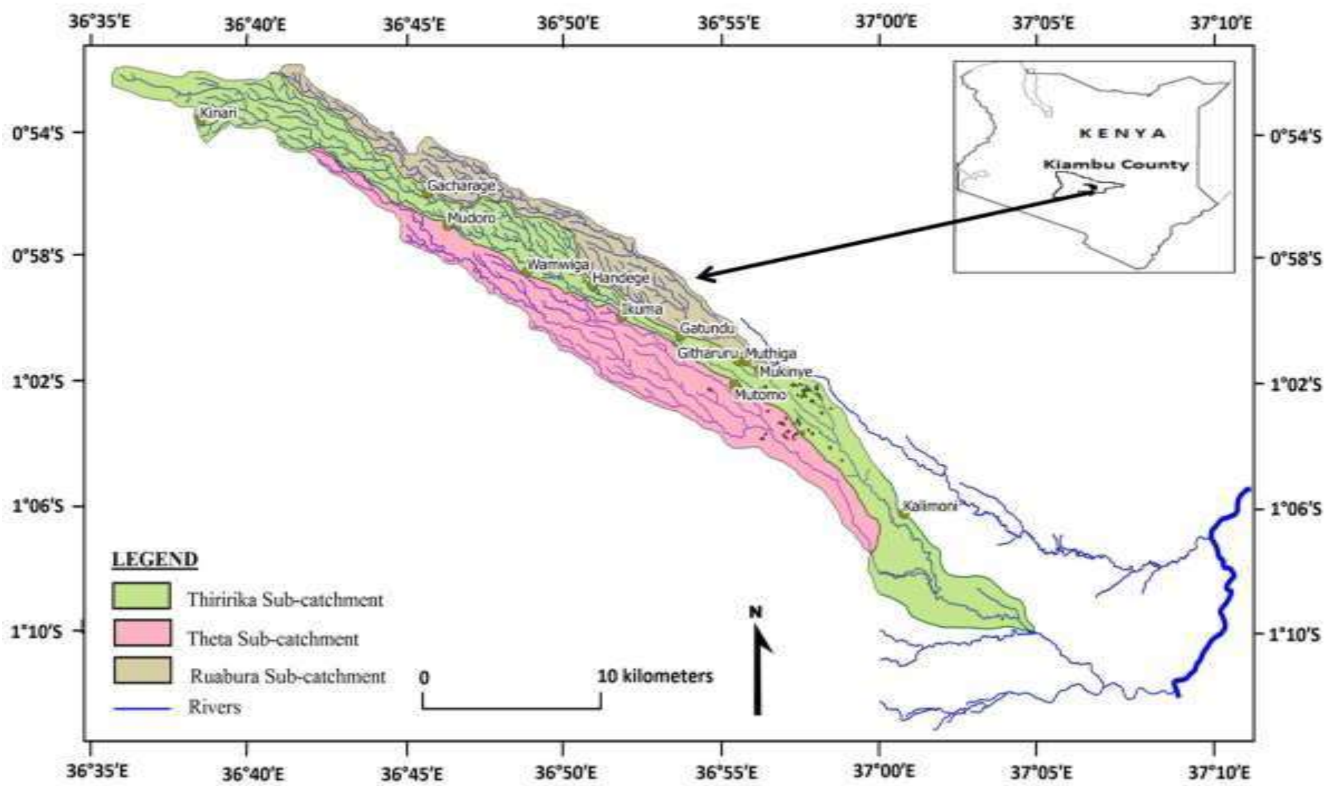

Figure 1. Location map of the study area.

office, 2013). Relative humidity ranges between 52 and 74.5% for dry and rainy seasons, respectively, with a monthly average of 66.7 per cent. The months of April, August and November are noted for peak humidity rises, while the least is experienced in February and March (Rwabura Irrigation Project Report, 2013). The physical features together with the climatic conditions which characterize the study area create a favourable environment for the cultivation of high value crops like coffee, tea, cereals and horticultural crops.

Geology and soils

The study area lies in the tertiary volcanic rocks region of central Kenya. Its geology can be classified as Kerichwa valley tuffs along the river valleys and the middle and upper Kerichwa valley tuffs found on the higher grounds. The study area is characterized b y one soil unit nitrisols which comprises two soil types, that is, humic and rhodic nitisols. Other types that are presents but occupy small areas are umbric andosols, haplic nitisols and rhodic ferralsols. Humic nitisols are found on the upper parts of all the three catchments and rhodic nitisols found in the lower parts (Rwabura Irrigation Project Report, 2013).

Socio-economic conditions

The catchments have a total population of about 114,180 people, representing a density of approximately 593.5/km^2 (KNBS, 2009). The population is denser in the lower zones with over 550 persons per square kilometre while the upper zones have 450 persons per square kilometre. There are clustered settlements around towns as a result of influx of people who migrate to live in urban areas due to employment opportunities and better infrastructural facilities. Administratively, Theta town is a ward under Gatundu north constituency, while Thiririka and its environs falls under Gatundu south constituency. The main economic activities in the study area are horticultural farming involving the growing of fruits and vegetables and cash crop farming involving the growing of coffee, tea and tissue culture bananas. In addition, most farmer also practice livestock keeping. There are small scale irrigation practices along river banks and valleys bottoms which are either pump-fed or bucket-fed. Irrigation is by furrows, sprinkler or tied- basins mainly for horticulture crops like tomatoes, kales, onions, cabbages, bananas, spinach and French beans. Most commercial farms grow tea and coffee especially in Theta sub-catchment own dams and ponds which they use for irrigation. Irrigation water management involves direction of water into the fields and specific intervals per crop. Irrigation normally varies depending on the kind of crops being irrigated. Regular irrigation is required for shallow rooted crops and those grown in shallow and light soils. Long period interval for irrigation is required for the deep rooted crops and in moderate to loamy sands.

MATERIALS AND METHODS

Sampling procedure

Sampling techniques based on United States Environmental Protection Agency Method 1669 (USEPA, 1996) and Standard Operating Procedures (Joy, 2006; Ward, 2007) were followed. Sampling bottles were cleaned by washing with detergent and tap water, followed by sequential rinsing with tap water and then soaked in 10% nitric acid and finally rinsed with deionised water and openly dried for 24 h. At the sampling points, bottles were rinsed thrice with the source of water being sampled prior to sample collection.

Boreholes were purged for five minutes prior to sampling. Samples were collected by pumping directly into the sample bottles. The bottles were filled just enough, not allowing or introducing bubbles in the water and caped soon enough to prevent exposure to air and cross contamination. The bottles were labelled using permanent ink markers. *In situ* field parameters, that is, pH, temperature and electrical conductivity were recorded by pumping sample water into a beaker and immersing the Multi electrode water testing kit (portable OakTon 510 series). Measurements were recorded in a field note book with unique number and exact sampling location. Shallow wells were sampled using teflon bailers attached to a rope. Each bailer was disposed after use in order to prevent cross contamination. *In situ* field parameters were measured by collecting sample water in a beaker and immersing the Multi electrode water testing kit (portable OakTon 510 series). Prior to field sampling, the water testing kit was calibrated before field use and after every field activity prior to the next field activity.

The samples were kept in ice coolers at a temperature of 4°C and transported to Kenyatta University for analysis. All the electrodes were calibrated and checked prior to field measurements. The bulb end of the temperature electrode was carefully placed into the beaker of water and the temperature was determined after 2 min of waiting for reading to stabilize. The pH electrode was immersed in the sample and stirred gently, it was then allowed for 1 to 2 min for a stable reading and recording. Water EC was measured using EC electrode of the multi electrode water testing kit (portable OakTon 510 series). The conductivity cells and beaker was rinsed with a portion of the sample. The beaker was filled completely. The cell was inserted into the beaker. The temperature control was adjusted to that of the sample and the probe was inserted into the vessel and the conductance read.

Laboratory analyses

The samples were analysed for calcium (Ca^{2+}), magnesium (Mg^{2+}), potassium (K^{+}), sodium (Na^{+}), carbonate (CO_3^{2-}), bicarbonate (HCO_3^{-}), chloride (Cl^{-}), sulphate (SO_4^{2-}), nitrate (NO^{3-}), and phosphate (PO_4^{-3}). Besides, pH and electrical conductivity in µS/cm at 25°C were also measured. The following analysis was also done:

1. Bicarbonate and carbonate were determined by titration
2. Calcium and magnesium were determined by ETDA titration
3. Sulphate was determined by spectrophotometer
4. Nitrate was determined by hydrazine reduction
5. Potassium and sodium were determined by flame photometer
6. Chloride was determined by argentometric precipitation method (Mohr's Method).

Data processing and analysis

Data collected was first evaluated for accuracy by ensuring the sum of total of all cations is equal to sum of all anions in milli equivalent per litre (meq/L) (Deutsch, 1997). The accuracy of laboratory results was estimated using balance error equation in Equation 1. A balance error is acceptable when it is less than 5% (Deutsch, 1997). All the water samples were within the acceptable error level. This was done to ensure precision, reliability and relevance of the data to the study. The summary statistics such as mean and standard deviation for all independent variables with respect to sample identities were generated. Descriptive statistics such as bar graphs was used to display and explain results and compared to FAO (1985) standards. Electrical conductivity (EC) measured on the field was also used to assess salinity hazards.

$$\frac{Cation}{Anion}\,balance = \frac{\sum cation\left(\frac{meq}{L}\right) - \sum anion\left(\frac{meq}{L}\right)}{\sum cation\left(\frac{meq}{L}\right) + \sum anion\left(\frac{meq}{L}\right)} \qquad (1)$$

RESULTS AND DISCUSSION

Temperature, pH and total alkalinity of water sources

Temperature readings for both boreholes and shallow wells ranged from 20.3 to 25.4°C with an average of 23.11°C. The maximum temperature was recorded from shallow well (SW003) while the minimum from a borehole (BH008) in the Rwabura sub catchment. All the water sources fell within FAO (5 to 30°C) depending on the crop being grown) recommended temperature for irrigation. The results show that the pH of water sources was below the FAO maximum limits and samples fell within the normal pH ranges of natural waters (4 to 9) as stated by Maral (2010). The pH of sampled water in the study area ranges from 4.48 to 7.71 with the minimum from a shallow well and the maximum for both shallow well and borehole, though, the average pH (6.16) of water used for irrigation in the study area is less than the lower limits of FAO standards indicating slight acidity of groundwater sources in the catchments. This will however not have any significant effects on the soils pH in a short term since soils are high buffer systems (Maral, 2010). Nonetheless, long term use will cause excess iron and manganese in the soil which can lead to plant toxicity (Faust and Will, 2002). About 95% of samples from shallow wells and 35% of borehole had pH less than the minimum FAO recommendation of 6.5 (Table 1 and 2). This shows that, shallow wells in the study area are acidic as compared to the boreholes. Alkalinity of water sources in the sub catchments range from 99.3 to 189.9 mg/L and an average of 150.29 mg/L with the maximum and minimum recorded for a shallow well and

Table 1. Shallow well water quality.

Well ID	Temperature (°C)	pH	EC	TDS	Nitrate	Chloride	Potassium	Sodium	Calcium	Magnesium	Bicarbonate	Carbonate	Alkalinity	Sulphate	SAR
			(µs/cm)	(mg/L)											
SW 001	21.7	5.73	770	492.8	18.1	374.81	5.3	25.9	42.65	36.82	42.14	6.02	174.2	48.9	4.1
SW 002	22.7	5.25	685	438.4	16.16	184.87	6.4	20.1	44.68	36.69	122.1	6.24	148.8	50.7	3.15
SW 003	25.4	6.13	749	479.4	27.7	139.88	11.3	22	48.08	36.51	196.63	28.1	168.4	62.9	3.38
SW 004	23.5	5.53	1248	798.7	14.95	399.8	17.4	14.9	41.97	35.88	30.51	12.34	150.5	69	2.39
SW 005	21.7	5.16	587	375.7	25.8	324.82	16.7	42.8	46.04	37.09	102.23	8.42	182.6	60.7	6.64
SW 006	21.8	5.37	860	550.4	14.72	349.81	13.7	28.9	28.39	37.19	30.88	4.12	168.7	69.8	5.04
SW 007	23.6	5.93	467	304	16.4	269.84	12.6	35.1	42.65	37.22	80.48	4.44	148.6	58.7	5.55
SW 008	23.5	5.73	704	450.6	67.5	364.81	30.2	34	41.29	38.89	25.68	8.28	106.8	56.4	5.67
SW 009	23.6	6.04	688	440.3	19.74	399.77	12.1	24	37.9	28.8	7.39	4.12	122.5	125.2	4.16
SW 010	23.2	6.23	547	350.1	15.79	399.8	15.2	25	43.33	27.8	8.88	4.22	120.8	87.9	4.19
SW 011	24	4.48	937	599.7	70.98	289.83	11.7	47.5	39.93	27.16	35.2	10.13	114.9	78.8	8.2
SW 012	23.3	6.23	845	540.8	7.63	149.88	16.1	43.9	41.97	41.28	10.34	2.32	181.6	77.4	6.8
SW 013	22.9	6.43	703	449.9	8.57	349.78	15.9	34.2	34.5	39.51	101.56	24.05	150.2	49.9	5.62
SW 014	23.8	7.01	609	389.7	4.76	244.85	19.7	19.6	40.61	43.1	14.74	4.71	176.8	45.3	3.03
SW015	21.9	5.47	542	346.8	5.35	149.88	14.4	48	33.14	38.2	95.41	8.26	173.9	52.5	8.04
SW016	23.6	4.86	487	311.68	20.3	359.81	14.7	37.9	43.33	36.23	46.62	4.17	177.9	69.8	6.01
SW017	22.7	5.49	509	325.7	13.6	399.8	10.3	23.2	40.61	36.59	62.59	10.49	189.9	70.7	3.73

BH, Borehole; SW, Shallow well; EC, Electrical conductivity; TDS, Total dissolved solids; SAR, Sodium absorption ratio.

Table 1. Borehole water quality.

Well ID	Temperature (°C)	pH	EC	TDS	Nitrate	Chloride	Potassium	Sodium	Calcium	Magnesium	Bicarbonate	Carbonate	Alkalinity	Sulphate	SAR
			(µs/cm)	(mg/L)											
BH 001	23.9	7.71	572	366.1	0.29	162.8	4.25	15.05	35.86	36.82	79.17	4.81	173.9	41.77	2.49
BH 002	24.1	7.62	485	310.4	0.55	158.3	3.48	15.83	44.01	36.69	87.45	18.01	149.9	36.97	2.46
BH 003	22.4	7.31	632	404.5	0.61	45	1.96	16.13	40.61	34.51	12.29	8.21	146.9	42.09	2.63
BH 004	23.7	7.15	921	589.4	0.79	46.2	1.86	19.07	41.97	36.88	87.54	18.01	147.3	45.91	3.04
BH 005	23.5	6.74	665	425.6	0.94	37.8	1.05	23.38	31.11	42.09	10.51	8.21	175.3	31.01	3.86
BH 006	23	6.88	802	513.3	1.23	42.3	2.63	25.64	42.65	27.19	20.07	10.44	99.6	27.43	4.34
BH 007	23.5	5.81	755	483.2	1.41	45.1	2.72	25.44	41.29	27.22	5.07	6.05	99.3	20.8	4.53
BH 008	20.3	6.25	781	499.8	1.48	53.8	4.83	17.3	39.25	36.89	45.92	14	142.7	31.68	2.8
BH 009	22.7	6.62	670	428.8	1.51	65.7	1.29	27.11	43.33	35.8	29.22	12.06	140.4	26.91	4.31
BH 010	23.6	6.17	652	417.3	1.52	69.8	5.88	28.19	41.29	35.97	45.23	14.65	140.2	33.2	4.54
BH 011	21.8	5.37	870	556.8	2.01	62.8	0.81	16.22	44.01	47.16	35.76	15.68	176.3	41.38	2.4
BH 012	24	6.11	891	570.2	2.58	60.6	1.48	24.26	47.4	36.28	58.31	5.09	132.7	39.88	3.75
BH 013	23.2	7.31	621	397.4	2.67	47.6	2.24	16.62	45.36	26.51	75.26	8.4	120.6	38.99	2.78
BH 014	24.6	5.87	754	482.6	2.95	129	0.72	14.56	45.36	37.1	94.76	10.45	138.9	39.74	2.27
BH 015	22.5	6.39	622	398.1	3.14	78.9	1	17.3	43.33	35.2	40.27	3.74	145.3	30.99	2.76
BH 016	23.2	6.65	1328	849.9	3.45	55.8	1.38	28.26	38.57	35.23	65.14	13.69	147.8	29.77	4.65
BH017	22.9	6.34	505	323.2	3.54	50.2	2.53	28.97	45.36	39.59	18.63	6.98	175.8	30.8	4.45

borehole, respectively. The average for the shallow wells and boreholes were 156.30 and 144.29 mg/L respectively indicating slight alkalinity of the former as compared to the latter. All the samples are within the FAO limit and therefore will have no significant effect on soils and plant growth in the sub catchments.

Electrical conductivity (EC), total dissolved salts (TDS) and sodium absorption ratio (SAR)

The most influential irrigation water quality guideline is salinity. Water with high EC or TDS is toxic to plants and causes soil degradation and therefore must be controlled. The EC value of sampled groundwater in the study area range from 467 to 1328 µS/cm with a mean of 719.5 µS/cm which according to Wilcox (1955) cited by Islam and Shamsad (2009) falls within the irrigation Water quality standard 'excellent to good' and 'slight to moderate'. 50% of the sample can be classified as having excellent to good quality and the rest having slight to moderate quality. Electrical conductivity of water is a direct function of its total dissolved salts (Harilal et al., 2004). Both EC and TDS values are indicative of saline water in the absence of non-ionic dissolved constituents (Michael, 1992). Electrical conductivity correlates with estimation of TDS. However, every water resource has unique and variable dissolved salts and therefore in addition to EC, it is very important to consider TDS. The TDS of groundwater in the study area range from 304 to 849.9 mg/L with a mean of 460.63 mg/L. The maximum and minimum occurred for BH016 and SW007, respectively. The maximum SAR value permissible for irrigation water according to FAO is 9 meq/L. From this, all the water samples had values less than 9 meq/L and therefore will not cause any sodicity hazards when used for irrigation.

Major salts ion concentrations in water sources

Toxicity problems occur when ions in irrigation water accumulate within the soil or plant or plant at concentrations that causes damage and reduced yield. The degree of damage depends on the concentrations of ion present. Results of the various cations and anions (Na^+, Ca^{2+} and NO_3^-) that form salts in the water sources were compared with FAO standards to assess their possible effect on either plant and or soil.

Bicarbonate and carbonate concentrations

Bicarbonate and carbonate ions in high concentrations in soils affect plant mineral nutrients uptake and metabolism (Phocaide and FAO, 2007). According to Kalbasi (1995), chlorosis in plants is as a result of iron nutrition disorder due to high concentrations of bicarbonate and carbonate. It has also been established that irrigation water with high bicarbonate/carbonate concentrations gives rise to the precipitation of calcium as calcium carbonate causing soil solution to become relatively enriched with sodium (DWAF, 1996).

Bicarbonate concentrations

Figure 2 indicate the levels of bicarbonate concentration in water from the various sampling points. The results show that all the water sources have bicarbonate levels lower than the FAO upper limit except for the SW003 with a bicarbonate concentration of 196.63 mg/L. The bicarbonate concentrations in the three sub catchments range from 5.07 to 196.63 mg/L (BH007 and SW003, respectively). The average concentration for the study area is 53.67 mg/L and below the mid value of FAO (1985) permissible level. However, about 35% of water sources (SW002, 003, 005, 007, 013, 015 and BH 001, 002, 004, 013, 014) have above 75 mg/l bicarbonate concentration. Prolong application will increase the tendency for calcium and magnesium to precipitate as insoluble carbonates in soils in which they are used. The low bicarbonate content of all the water samples accounted for the low alkalinity and pH levels in the study area.

Carbonate concentrations

Carbonate in irrigation water becomes significant as water pH increases beyond 8.0 and exceeds 10.3. Water containing appreciable carbonates will exceede desirable bicarbonate levels as well (Maral, 2010). pH, alkalinity and bicarbonate levels were very minimal and therefore carbonate will not be an exception as confirmed in the table. The concentration of carbonate in the sub catchments ranged from 2.32 (a shallow well - SW012) to 28.1 mg/L (SW003) with an average of 9.67 mg/L. Though these values are very low, the maximum limit as per FAO (1970) cited by Banderi et al. (2012) is 3 mg/L and therefore only the water from shallow well SW012 is good for irrigation.

Calcium concentrations

Some quantities of calcium impact positively by countering the negative effects of sodium and helps maintain good soil properties (Fipps, 2004). According to Gebauer and Ebert (2005), calcium preserves the structural and functional integrity of a plants membrane, stabilizes cell wall structure, regulates ion transport and selectivity and control ion exchange behaviour as well as cell wall enzyme activities. The calcium concentration of

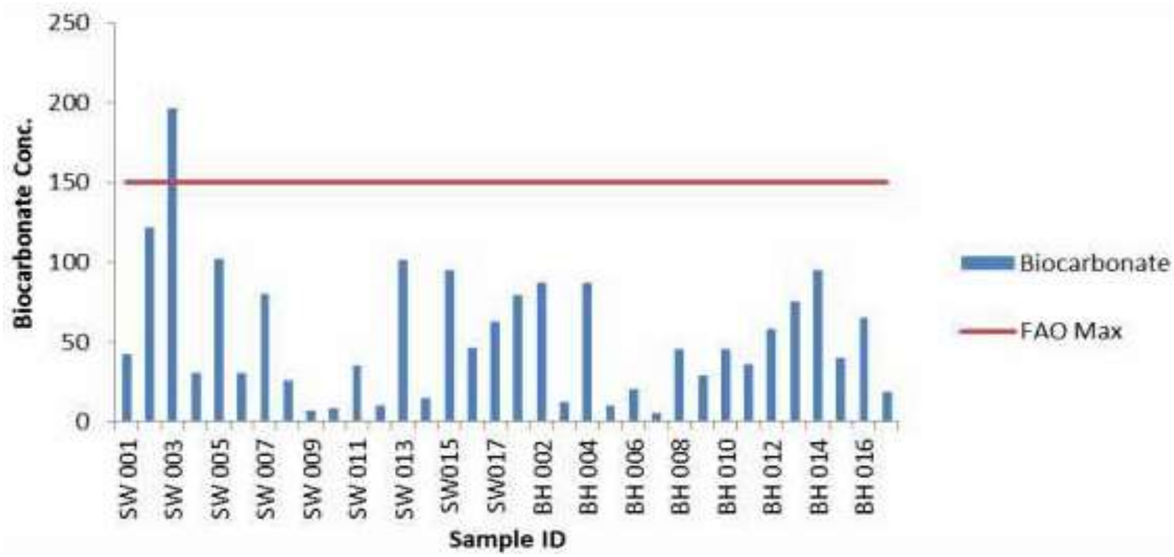

Figure 2. Bicarbonate concentration.

water samples in the study area range from 28.39 to 48.08 mg/L with an average of 41.23 mg/L with both the maximum and minimum occurring for shallow well SW003 and SW006, respectively. Calcium concentration was nearly the same across the study area. These concentrations fell within the acceptable FAO (1985) standards and therefore will not inhibit the uptake of iron from soil and induced chlorosis (Padmore, 2009) (Figure 3).

Chloride concentrations

Chlorides are important anions effective in the formation of saline soils and are highly water soluble and thereby possess severe toxicity (Banderi et al., 2012). These anions cannot be precipitated at concentrations usually present in water but are readily transported through plants' roots and conveyed to the leaves where they accumulate (DWAF, 1996). Its toxicity is common in water logged areas due to the rapid rates of its ion transport.

Concentration of chloride in all the 34 irrigation water samples varied considerably from 37.8 to 399.8 mg/L with an average value of 187.17 mg/L within the study area. From Figure 4, all the samples from the shallow wells and that of boreholes (BH001, BH002, and BH014) are not conducive for sprinkler irrigation as it may damage the plant leaves. For surface irrigation, about 21% (SW 001, SW004, SW008, SW009, SW010, SW016 and SW017) are not conducive, while the rest are conducive for irrigation.

Magnesium concentrations

Water with magnesium increases the potential effect of sodium on the soil when used for irrigation and its extreme deficiency also causes acidic soils which can lead to low levels of essential plant nutrients such as phosphorus and molybdenum (Maral, 2010). Excessive magnesium in water also causes water hardness which when used for irrigation deposits mineral residues on plant surfaces or foliage. As per FAO recommendations, the magnesium concentrations in samples from the study area are within the acceptable limit for irrigation purposes. It ranged from 26.51 to 47.16 mg/L for BH 013 and BH 011, respectively. The values as indicated will therefore have no effect on both the soil and plant. However, its use in conjunction with magnesium rich fertilizer must be avoided so it does not exceed soil the magnesium ratio (MR).

Nitrate concentrations

The nitrate concentration varied within the study area (Figure 5). The nitrate values varied from 0.3 to 71.0 and a mean of 11.7 mg/L. The maximum was observed for a shallow well and the minimum for a borehole. All the sampled shallow well in the study area have concentrations way above FAO (1985) limit, while all the boreholes are within the limit.

Sodium concentrations

Sodium in minute quantities is beneficial to plants, however in excess can cause toxicity problems for some crops, especially when sprinkler irrigation is applied. The concentrations of sodium in all the samples were within the FAO permissible levels of 50 mg/L and pose no

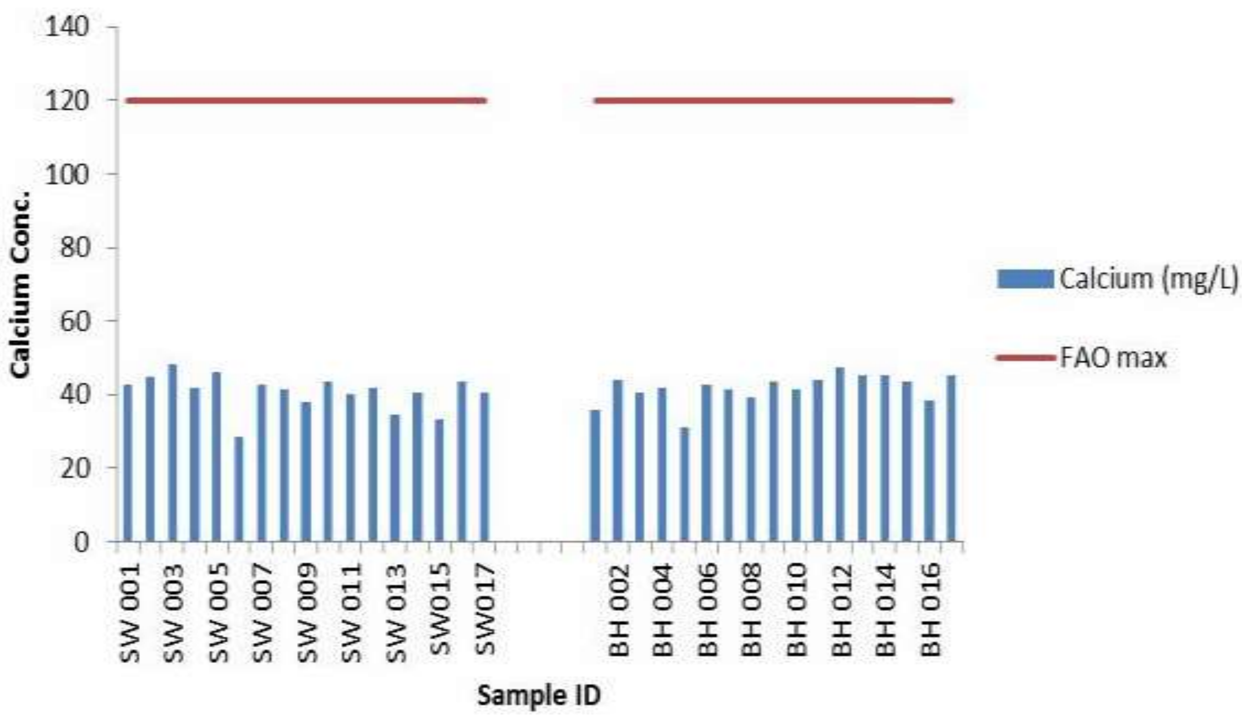

Figure 3. Calcium concentrations.

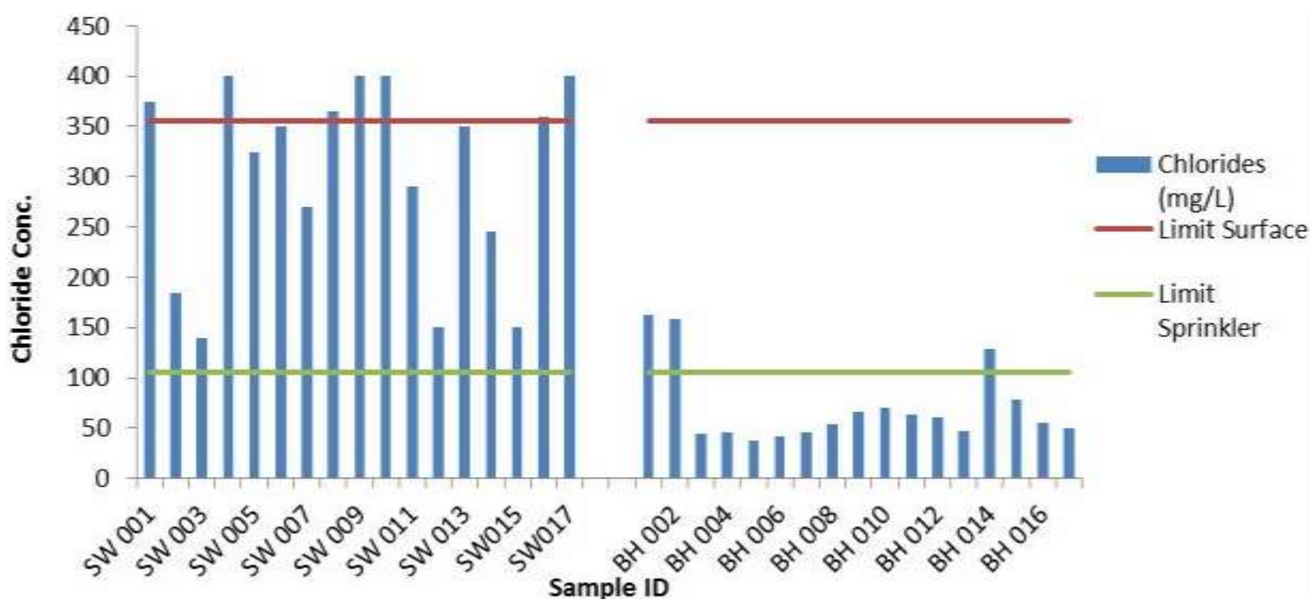

Figure 4. Chloride concentration.

toxicity problems. Nonetheless, the continuous and long term use of water from points SW 005 (42.8 mg/L), SW 011 (47.5 mg/L) and SW 012 (43.9 mg/L) is not advisable or should be used with care since they have high levels of sodium. The sodium concentration across the study area ranges from 14.56 to 48 mg/L.

Potassium concentration

Potassium levels of the study area range from 0.72 to 30.2 with a mean of 8.35 mg/L. The maximum observed for a shallow well (SW008) and the minimum for a borehole (BH014) as well. It is also observed that

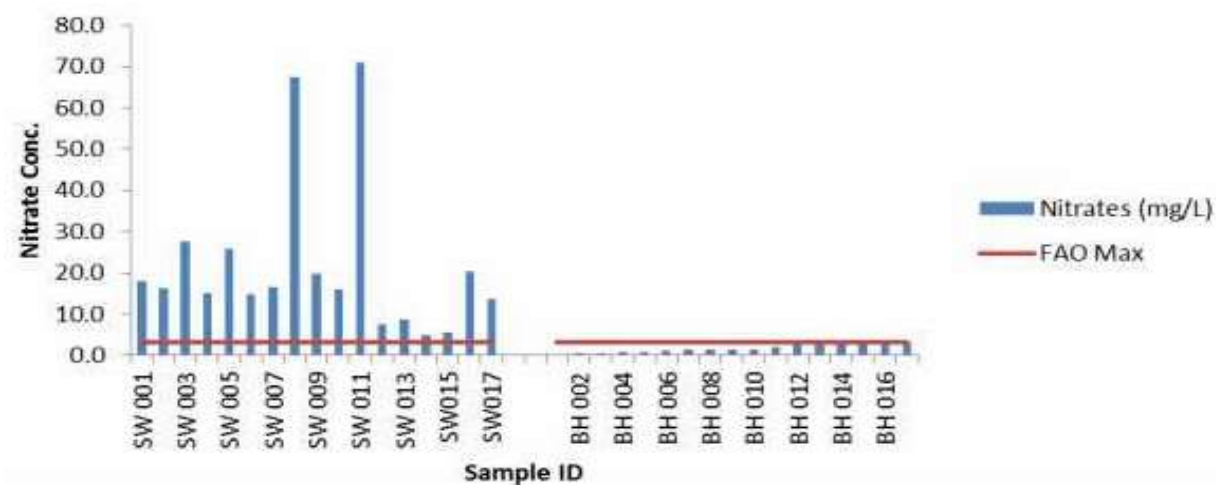

Figure 5. Nitrate concentration.

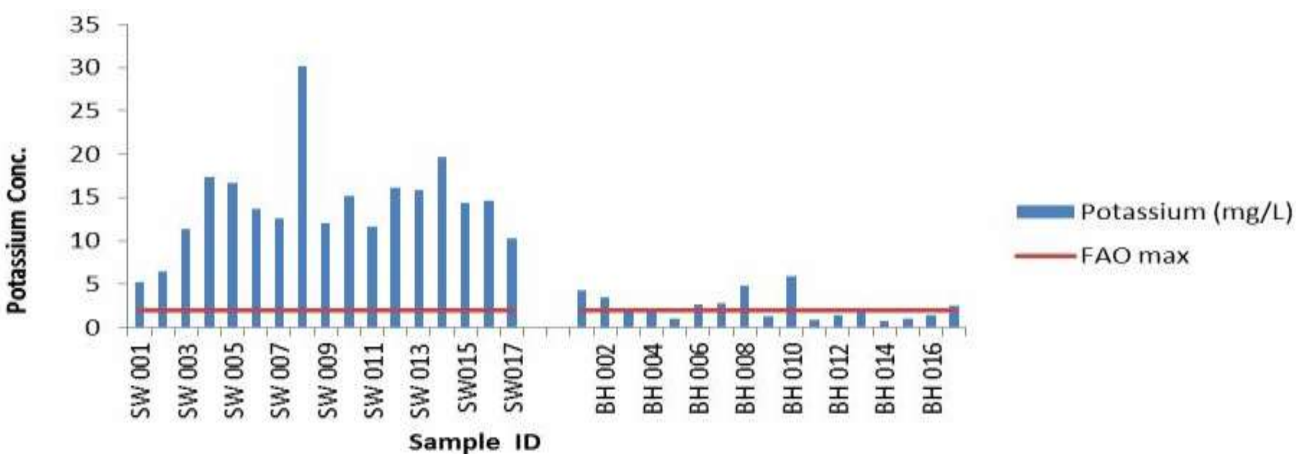

Figure 6. Potassium concentration.

samples from all the shallow wells were above the required FAO limit with 50% of samples from boreholes also having concentration above the FAO required limit (2 mg/L) (Figure 6). This means that, there will be a possibility of magnesium deficiency and iron chlorosis in the study area when used for irrigation as suggested by Fipps (2004).

Sulphate

Sulphate is a major anion in many irrigation waters. Its toxicity is not much of a problem; however, extremely high concentration may interfere with uptake of other nutrients by plant. According to FAO (1985), water with a sulphate concentration of 0 to 20 meq/L is considered as a usual range in irrigation water. This value when converted in mg/L is very high (above 200 mg/L). In the study area, the highest concentration was 125.2 mg/L, indicating that the water samples in relation to sulphate are safe for irrigation purposes without restriction.

Conclusion

Groundwater samples were assessed for their quality in terms of their potential for irrigation. The results show that groundwater in the study area were slightly acidic. Majority of samples also recorded TDS and EC values less than FAO maximum allowable concentration. Chloride ion concentration was generally high as compared to other ions. Calcium ion concentration was generally high as compared to other cations. Bicarbonate, calcium, magnesium, sodium and sulphate ions of groundwater were within the acceptable limit for irrigation, while the rest have no sodium adsorption ratio (SAR) values, suggesting suitability of groundwater from the study area for irrigation. It can also be derived from the results that, the boreholes were of ample quality as compared to the shallow well. The boreholes had less concentration of ions as compared the shallow wells, however, the results in general indicate that groundwater in the three sub catchments is suitable for irrigation purpose in a short term. This research may serve as a preliminary study to provide baseline information that may direct future water quality assessment studies in the study area.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Banderi A, Mohammad MS, Hamdi H, Shomeli M, Ali SD (2012). Water quality assessment of Amir Kabir sugarcane agro-industry with respect to agricultural and drinking views. J. Environ. Ecol. 3(1):11-28.

Blank HG, Mutero CM, Rust HM (2002). The changing face of irrigation in Kenya: Opportunities for anticipating change in Eastern and Southern Africa. International Water Management Institute. pp. 14-16.

Consultative Group on International Agricultural Research (2003). Water and People in catchments. Challenge program on water and food. Cali, Colombia.

Department of Water Affairs and Forestry (1996). South African Water Quality Guidelines (second edition) Volume 4: Agricultural Use: Irrigation.

Deutsch WJ (1997). Groundwater Geochemistry: Fundamentals and applications to contamination. Lewis Publishers, Boca Raton, FL (United States).

USEPA (1996). Method 3050B Acid digestion of sediments, sludge's and soils, 2nd revision U.S environmental protection agency, office of solid waste www.epa.gov/epawaste/hazard/testmethods/sw846/pdfs/3050b accessed on 04/12/2012.

FAO (1985) Water Quality for Agriculture, Irrigation and drainage paper No. 29, Rev.1. Food and Agriculture Organization of the United Nations, Rome.

Faust JE, Will E (2002). Irrigation Water Quality for Greenhouse Production, Agricultural Extension Service, University of Tennessee. http://www.utextension.utk.edu/publications/pbfiles/pb1617.pdf. Accessed on 17th may 2013.

Fipps G (2004). Irrigation water quality standards and salinity management strategies, Texas cooperative extension, Texas, USA.

Gatundu Agricultural Office (2013). Gatundu south constituency, Kiambu County, Kenya.

Gebauer J, Ebert G (2005). Soil salinity in Sudan and the possibilities of irrigation. WaWi 3:47-49.

Ghassemi F, Jakeman AJ, Nix HA (1995). Pakistan. In: Salinisation of land and water resources: human causes, extent, management and case studies. University of New South Wales Press Ltd, Sydney. pp. 369-395.

Harilal CC, Hashim A, Arun PR, Baji S (2004). Hydro geochemistry of two rivers of Kerala with special reference to drinking water quality. J. Ecol. Environ. Conserv. 10(2):187-192.

Islam MS, Shamsad SZKM (2009). Assessment of Irrigation water quality of Bogra District in Bangladesh. Bangladesh J. Argil Resour. 34(4):597-608.

Joy J (2006). Standard Operating Procedure (SOP) for Manually Obtaining Surface Water Samples, Version 1.0. Washington State Department of Ecology, Olympia, WA, SOP Number EAP015.

Kalbasi M (1995). Iron chlorosis in fruit trees. Isfahan University of Technology Press. P 13.

Kenya National Bureau of Statistics (2009). Economic survey 2009, Government Printer, Nairobi: KNBS.

Maral N (2010). Soil and water analysis techniques for agricultural production a thesis submitted to the graduate school of natural and applied sciences of Middle East technical university.

Michael AM (1992). Irrigation theory and practices, Vikash Publishing House Pvt. Ltd., New Delhi-India. pp. 686-740.

Mugwanja PN, Mochiemo GO, Osoro CM (1995). Management of salt affected soil in Taveta sub-district, Coast Provinces. Proceedings at the International Workshop integrated soil management for sustainable use of salt affected soils. Manila Philippines.

Ndegwa GM, Kiiru I (2009). Investigation on soil and water quality as affected by irrigation in Turkana District, Kenya. J. Agric. Sci. Technol. 1(12):17-36.

Phocaide A, Food and Agriculture Organization of the United Nations (2007). Handbook on pressurized irrigation techniques, Rome: FAO.

Padmore C (2009). Irrigation salinity – causes and impacts, PRIMEFACT 937. Wagga Wagga Agricultural Institute, Department of Primary Industries, Orange Wagga Wagga NSW 2650.

Republic of Kenya (2009). National Irrigation and Drainage Policy, Ministry of water and irrigation, Nairobi, Kenya.

Rwabura Irrigation Project Report (2013). Gatundu Agriculture Office. Kenya.

Squires VR, Glenn EP (2004). Salination, Desertification, and Soil erosion, In: The Role of Food, Agriculture, Forestry and Fisheries in Human Nutrition, [Ed. Victor R. Squires], in Encyclopaedia of Life Support Systems (EOLSS), Developed under the Auspices of the UNESCO EOLSS Publishers Oxford UK.

Stein C, Schulze C (1978). Land use and Development Potential in the arid regions of Kenya. In: Applied Science and Development. Institute for Scientific Cooperation 12:47-64.

Wanjogu SN, Gicheru PT, Maingi PM, Nyamai M (2004). Saline and sodic soils in the Drylands of Kenya, Nairobi, Kenya.

Ward B (2007). Standard Operating Procedure (SOP) for the collection and field processing of metals samples, version 1.3. Washington State Department of Ecology, Olympia.www.ecy.wa.gov/programs/eap/quality.html.

Wilcox LV (1955). Classification and use of irrigation waters. U.s. Department of Agriculture cire, 969, washington, d.c. P 19.

International Journal of Water Resources and Environmental Engineering

Related Journals Published by Academic Journals

- *International Journal of Computer Engineering Research*
- *Journal of Chemical Engineering and Materials Science*
- *Journal of Civil Engineering and Construction Technology*
- *Journal of Electrical and Electronics Engineering Research*
- *Journal of Engineering and Computer Innovations*
- *Journal of Engineering and Technology Research*
- *Journal of Mechanical Engineering Research*
- *Journal of Petroleum and Gas Engineering*

academicJournals