The Relationship Between Urbanization and Precipitation in Kisumu City: Co-Integration Analysis

Benard Ochieng Samba Dr. George Orwa Dr. Joseph Mungatu
Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi, Kenya

Abstract
The work done in this study thesis is by empirical analysis. If two series are integrated of the same order, then the two series are said to be co-integrated and this is shown in the two series where the population series being non stationary is made stationary by third differencing and the stationary rainfall series subjected to the same order of differencing to achieve the co-integration rule. The OLS method was used and then the model parameters tested for adequacy. A linear Error Correction Model was fitted and evidence that a short term relationship between the rainfall and population series is seen to exist. Also, a high threshold value is observed at the second lag. A high R squared value of 0.9881 is an indication that the model fits well to the data. A small p-value also indicates that the model is highly significant. Hence, it is recommended that a close analysis of population growth rates be analyzed to aid in the prediction of the rainfall rates movements.
Keywords: OLS, Co-integrated, differencing, Error Correction Model, p-value.

1. Introduction

Precipitation is defined as any liquid or solid aqueous deposit that falls from the clouds and reach the ground. They include rain, drizzle, snow, snow pellets, ice crystals and hail.

Rain on the other hand is defined as the liquid water in form of droplets that have condensed from atmospheric water vapor and then precipitated (become heavy enough to fall under gravity). Rain water is a major component of the water cycle and is responsible for the deposit of most of the fresh water on earth. It provides suitable conditions for many types of ecosystems. Therefore, rainfall change is considered to be one of the most serious dangers to sustainable development. It has effects on Environment, Human health, Food security, Economic activities, Natural resources and Physical Infrastructure. Some important finding by the Intergovernmental Panel on Climate Change (IPCC) fourth Assessment Report (AR4) include the fact that millions of people globally will be exposed to increased water dilemma and consequently, the access to food in many African countries will be seriously compromised. Kenya is geographically located in Eastern Africa with latitude of $1'00N$ and a longitude of $38'00E$. The variety of relief and the range of altitude in Kenya produce various numbers of distinctive local climates, BBC (2012).

1.1. Urbanization

The urban population is presumed to have grown from less than 30% in 1950 to more than 47% in 2000 and United Nations Report of 2009 on Global Human Settlement indicates that currently, Africa is the fastest urbanizing continent in the world, Asoka et al. (2013).

Urbanization is defined as the population shift from rural to urban areas and the way in which the society adapts to the change. It usually results into the physical growth of urban areas either horizontally or vertically. Urbanization according to Peng et al. (2009) can also be defined as the process by which rural areas become urbanized as a result of economic development and industrialization. The United Nations projected that half of the world's population will live in urban areas by the end of 2008. It also predicted that by 2050, about 64% of the developing world and 86% of the developed world will be urbanized. Urbanization can be seen as a specific condition at a set time (e.g. the proportion of total population or areas in cities or towns) or as an increase in that condition over time.

Urban areas are usually mostly affected by rainfall than in rural areas. Thus, a clear understanding of urban rain could improve the weather forecast for regions with high population and also providing accurate warnings of potential flood producing rains and the possibility of shortages of rain. It has been observed by different scholars

on matters concerned with climate that urbanization is the main cause of carbon dioxide emission as a result of human activities and some of the human activities that enhance the atmospheric CO2 are: combustion of fossil fuels, increase in atmospheric dust content from human activities, and clouds produced by aircraft exhaust, Landsberg (1975)

1.1.1. Urbanization in Kisumu City

Kisumu City is found within Kisumu County. It is the third largest city in Kenya covering 417 square kilometers and is located in the Nyanza Province of Kenya, about 310 kilometers northwest of Nairobi and is on the shores of Lake Victoria. The growth rate of Kisumu is high estimated to be about 2.8% annually. According to the 2009 census, Kisumu City had a population of 390,164 people. This growth is as a result of high fertility which currently stands at 4.8 children per woman, compared to national average of 4.6 per woman. Currently, the country's population is dominated by young people who need support by those in work forces. Three quarters of the population is under 30 years old and 43.5 percent is under 15 years, PopulationAction (2014). Khaemba et al. (2012) did a suitability assessment showing the quantitative analysis of land use in Kisumu municipality indicating the four dominant land use types:

Table 1: Current Land use Allocations

Land use types	Areas in Hectares	Percentage area
Residential	23,440.9	79.3%
Commercial	1,168.6	4.0%
Industrial	189.6	0.6%
Recreational	4779.7	16.1%

1.1.2. Precipitation in Kisumu City

Kisumu records a significant rainfall throughout the year. This occurs even during the driest month. The average annual rainfall here is approximately 1352mm with the driest month being January with a precipitation averaging to about 65 mm. Most precipitation falls in April with an average of 224mm. The average precipitation varies at 159mm between the driest month and the wettest month. According to the report by the Kenya Red Cross Society, floods marooned over 200 homes and over 1500 people were affected in Kisumu County in Nyaori, Kasagam, Otera and Nyamasaria area on 1st of April 2013, RedCross (2013).

1.2. Introduction to Co-integration

Co-integration is a procedure which aims at investigating the relationship between a time series and the other, where the two series share a common drift. It is founded on the principle of identifying equilibrium or long-term relationship between variables. Two data series are said to have a long term relationship if their divergence from the equilibrium are bounded, meaning they move together and are co-integrated. Generally, for two or more series to be co-integrated, two conditions have to be met. The first one is that the series must all be integrated to the same order and secondly, a linear combination of the variables exists which is integrated to a lower order than that of the individual series. If the variables become stationary in a regression equation after first differencing, that is I(1); then the error term for the co-integration regression becomes stationary, I(0); Hansen and Juselius (1995). For a long period, equations involving non-stationary variables used to be estimated in macroeconomic models by using straightforward linear regression. Yet, testing the hypothesis using standard statistical inference sometimes led to different spurious results. It was until Clive Granger and Paul Newbold, Granger and Newbold (1974) in an influential paper called spurious regression in econometrics pointed out that tests of such regressions may often suggest statistically significant relationship between variables even when none exists. In conclusion, they generated independent non-stationary series, precisely random walks. They regressed these series on each other and observed the value of the t-statistic of the coefficient estimate calculated by assuming that the true value of the coefficient equals zero. Despite the variables in the regression being independent, they found that the null hypothesis of a

zero coefficient was rejected more frequently than predicted by standard theories. At the same time, the residuals of the estimated equation displayed strong positive correlation. This work thus formed an initial step in Granger's research agenda of developing methods for building more realistic and useful econometric model.

1.3. Pearson's Correlation Coefficient

This is the mostly used measure of correlation. It measures the linear regression between two random variables. The sign of the correlation coefficient determines whether the correlation is positive or negative. The magnitude of the correlation determines the strength of the correlation. Certain guidelines are offered for describing the correlational strength, Hoffman and John (2010):

$0<|r| < 0:3$ weak correlation

$0.3<|r| <0.7$ moderate correlation

$|r| > 0:7$ strong correlation

i.e. r = -0:849 indicates a strong negative correlation, where r in this case refers to the Pearson's correlation coefficient.

1.4. Objectives of the study

The objective of this study was to determine the effect of urbanization on precipitation in Kisumu City using the Error Correction Model

1.5. Assumptions of the Study

Urbanization was measured using the population rates. We assumed that Urbanization does not contribute to the increasing rates of rainfall. This might not necessarily be true.

2. Literature review

Boansi (2014) did a study on the Yield Response of Rice in Nigeria using co-integration analysis. The Johansen's Full Information Maximum Likelihood test was applied in estimating a yield response model. It was noticed that increasing the yield level for paddy rice in Nigeria and ensuring stability required interplay of biophysical, socio-economic and structural forces. Bridging of the demand-supply was found to be to be possible through initiation of measures to address inefficiencies in the supply chain. This would ensure appropriate transmission of price increments, promotion of local rice consumption, addressing soil fertility challenges through efficient use of fertilizers and regular management of rice fields' fertility and finally increasing access to credit for farmers to help them meet the costs of relevant inputs of production.

Kuwornu et al. (2011) did a study in the agricultural field on the Supply Response of Rice in Ghana by applying co-integration analysis. The annual time series data of aggregate output, total land area cultivated, yield, real prices of rice and maize, and rainfall were used in the analysis. The Augmented-Dickey Fuller test was applied in testing the stationarity of the individual series and Johansen maximum likelihood criterion was used in estimating the short-run and long-run elasticities. It was discovered that the cultivated land area of rice was significantly dependent on output, rainfall, real price of maize and real price of rice. The elasticity of lagged output in the short-run was significant at 1%. The error correction model was found to be the expected negative coefficient of -0.434 which was significant at 1% significance level. The empirical results also showed that the aggregate output of rice in the short run was found to be dependent on the acreage cultivated, the real prices of rice, rainfall and previous outputs. Real prices of rice and area cultivated were found to be significant at 10% while rainfall and lagged output at 5% significance level. In general, the analysis showed that short run responses in rice production were lower than long term responses.

Arouri et al. (2014) did a study on the Environmental Kuznets Curve in Thailand by applying Co-Integration and Causality analysis. The study aimed at exploring the existence of Environmental Kuznets Curve (EKC) in Thailand over the period 1970-2010. From the EKC relationship, as the economy grew, as measured by the per capita income, at the initial stage pollutants increased; but started falling after a certain threshold after income had been achieved. The postulated region produced an inverted U-curve. The study applied the Auto Regressive Distributed Lag (ARDL) bounds testing approach to co-integration in the presence of structural breaks for a long run relationship among the series. The Error Correction Mechanism was implemented for the short run dynamics. The results confirmed co-integration among the economic growth, energy consumption, trade openness, urbanization and energy pollutants and vindicated the presence of an EKC for Thailand. It was also noticed from the study that energy consumption and trade openness added to energy emissions while urbanization lowered it.

Solarin and Shabhaz (2013) did a study on the Trivariate Causality between Economic Growth, Urbanization and Electricity Consumption in Angola using Cointegration and Causality Analysis. The objective of the study was to investigate the causal relationship between economic growth, urbanization and energy consumption in the case of Angola by using the data over the period 1971-2009. The Lee and Strazicich (2003, 2004) unit root test was applied to test the stationarity properties of the series. In examining the direction of the causality between economic growth, urbanization and electricity consumption, the Variance Error Correction Method (VECM) was used. Granger causality test was subsequently used. The results indicated the existence of a long run relationship. There was also evidence on the bidirectional causality between Electricity Consumption and Economic Growth. A feedback hypothesis was also found between Urbanization and Economic Growth. It was also noted that Urbanization and Electricity Consumption Granger caused each other.

Ruxanda and Botezatu (2008) also did a study on co-integration and spurious regression while modelling the Romanian's M2 money demand. He confirmed that co-integration was the simplest way of eliminating the illogical correlation established between time series due to the presence of trends. In the study, the Romanian M2 money demand was modelled through co-integration and Vector Error Correction Mechanism. The Johansen's co-integration technique was also applied and an adequate equation obtained.

As far as climate change and economic growth is concerned, several studies have also been done by applying co-integration analysis. Nuru (2012) did an assessment on how climate change may impact economic growth through rainfall variability. Simple growth model were applied to demonstrate that the adverse impacts of rainfall variability on economic growth depended on the rate of expansion of the amplitude of rainfall variability and frequency of occurence of extreme events. A co-integration analysis using time series data from Ethiopia showed that both inter-annual and within-annual rainfall variations had negative effect on growth. Also, the simulation results on the past growths due to rainfall variability for the last five decades implied that mitigation and adaptation strategies towards climate change that would reduce the impact of rainfall variability would put Ethiopia on a higher trajectory of growth.

Li et al. (2013) also did a study on the long term relationship between population growth and vegetation cover basing their study on the panel data of 21 cities in Guangdong Province in China. From the research, they were basing their hypothesis on the fact that an inverse relationship existed between population growth and vegetation cover hence the study was done to disapprove this hypothesis. This was as a result of the fact that reports about vegetation protection and reforestation had been increasing continuously in recent decades. The Panel Error Correction Method (PECM) was used to investigate the causality direction between population growth and vegetation cover. The final results showed that not only will the consuming destruction effects and planting construction effect induced by the population growth have a great impact on the vegetation cover changes, but vegetation cover changes will in turn also affect the population growth in the long term.

3. Model Development

3.1. Introduction

Let $Y_t = (y_{1t},...,y_{nt})'$ denote an (n×1) vector of I(1) time series. Y_t is said to be co-integrated if there exists an (n×1) vector $\beta = (\beta_1,...,\beta_n)'$ such that: $\beta' Y_t = \beta_1 y_t + ... + \beta_n y_{nt} \sim I(0)$

This means that the non-stationary time series in Y_t are co-integrated if there is a linear combination of them that is stationary or I(0): If some elements of β are equal to zero then only the subset of the time series in Y_t with non-zero coefficients are co-integrated. The linear combination of $\beta' Y_t$ is often triggered by economic theory and referred to as a long-run equilibrium relationship. Intuitionally, the I(1) time series with a long-run equilibrium relationship cannot drift too far apart from the equilibrium because the economic forces will act to restore the equilibrium relationship. The co-integration vector β described above is not unique since for any scalar c the linear combination:

$c\beta' Y_t = \beta^{*'} Y_t \sim I(0)$. For this case, some normalization assumption is essential to uniquely identify β: A typical normalization is:

$\beta = (1, -\beta_2, ..., -\beta_n)'$ so that the co-integration relationship may be expressed as:

$\beta' Y_t = y_{1t} - \beta_2 y_{2t} - ... - \beta_n y_{nt} \sim I(0)$ or

$$y_{1t} = \beta_2 y_{2t} + ... + \beta_n y_{nt} + e_t \tag{3.1}$$

where $e_t \sim I(0)$ and the error term, e_t in 3.1 is often referred to as the disequilibrium error or the co-integrating residual. In the long run, the disequilibrium error is zero and the long term equilibrium relationship is:

$$y_{1t} = \beta_2 y_{2t} + ... + \beta_n y_{nt}. \tag{3.2}$$

3.2. Co-integration tests

Testing for the order of integration is important in several econometric works. The first purpose behind a unit root test is finding the order of integration and thus setting up an econometric model and does the inferencing. In this situation, it is vital to perform very detailed tests and take great care in finding exact critical values.

The second motive behind unit root tests is to investigate the properties of the time series prior to the construction of an econometric model. In this case, unit root tests acts as a descriptive tool performed to classify the series as either stationary or non-stationary. The following recommendations are usually made, if a data series appears to be non-stationary, then it is hypothetically assumed to be to be non-stationary and integrated. This hypothesis is rejected if and only if there is enough evidence of rejection. Once the variables are adequately classified as integrated, stationary or even deterministic trend stationary, the long-run and short-run effects in the model can be easily sorted out and a model set to give a meaningful statistical inference. The first step in co-integration analysis is to test for the stationarity of the two series. For the two series to be co-integrated, it is a condition that they must be non-stationary. The tests that will be applied will be ADF, KPSS and PP tests.

3.2.1. Augmented Dickey Fuller test

The ADF is the generalized form of the Dickey-Fuller test. It assumes that the residuals are independent and identically distributed. The ADF's idea is to include enough lagged dependent variables to rid the residuals of serial correlation. The ADF Model is estimated as:

$$\Delta x_t = \alpha + \lambda t + \pi x_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta x_{t-i} + \varepsilon_t \tag{3.2.1}$$

Where $\propto$ is a constant, λ is the coefficient on a time trend series, π is the coefficient of x_{t-1}, k is the lag of the autoregressive process, $\Delta x_t = x_t - x_{t-1}$ are first differences of x_t ; x_{t-1} are lagged values of order one of x_t, Δx_{t-i} are the changes in lagged values, and ε_t is the white noise.

3.2.2. Kwiatkowski-Philips-Schmidt-Shin test

The KPSS introduced by Kwiatkowski et al. (1992) tests the null hypothesis that the variable is stationary contrary to the ADF test that has the null that the variable is non-stationary. It tests the null for level or trend stationarity. A normal regression model with a linear combination of a deterministic trend, a random walk and a stationary residual series is given by:

$$X_t = \alpha + \beta t + \lambda \sum_{i=1}^{t} \varepsilon_i + \delta_t \tag{3.2.2}$$

is applied where δ_t is a stationary process, β_t is the trend component while $\sum_{i=1}^{t} \varepsilon_i$ is the random walk.

3.2.3. Philip-Perron's test

This is a unit root test by Philip and Perron (1988). It involves fitting the regression model and the results are used to calculate the test statistics. It modifies the test statistic so that no additional lags of the dependent variable are needed in the presence of serially-correlated errors. The PP test corrects for any serial correlation and heteroscedasticity in the errors ε_t non-parametrically by modifying the Dickey-Fuller test statistic. An advantage with the test is that it assumes no functional form for the error process for the variables (i.e. it is non-parametric).

The Philip and Perron's test statistic can be viewed as Dickey-Fuller statistics that have been made robust to serial correlation by using the Newey and West (1987) heteroscedasticity and autocorrelation-consistent covariance matrix estimator.

3.3. The Engle-Granger two step method for co-integration

If y and x are both I(1) and have a long run relationship, then there must be some forces which tends to pull the equilibrium error back to zero. Engle and Granger (1987) recommended a two-step procedure for co-integration analysis.

Step 1

We ensure that the individual series are I(1) which is the basic definition of co-integration
i.e. two variables must be integrated of the same order. This is done by the Augmented Dickey Fuller (ADF), the Philip-Perron and KPSS unit root tests to infer the number of unit roots (if any) in each of the variables under investigation. The model applied for the testing procedure of the ADF unit root test is as given in equation 3.2. After rejecting the hypothesis of unit root test, the long-run equilibrium relationship is estimated in the form of an OLS regression line. If the variables co-integrate, the OLS regression yields a super consistent estimator, Enders (2004).

3.3.1. Error Correction Model

The second step is the estimation of the Error Correction Model by OLS method. An error correction model is defined as a dynamic model in which the movement of a variable in any period is related to the previous period's gap from the long-run equilibrium. Although it may be possible to estimate the long run or co-integrating relationship, $y_t = \beta x_t + \varepsilon_t$, economic systems are rarely in equilibrium as they are affected by institutional and/or structural changes that are either temporary or permanent. Since equilibrium is hardly observed, the short-run evolution of variables (short-run dynamic adjustment) is important. A simple dynamic model of a short-run adjustment model is given by:

$$\Delta y_t = \alpha_0 + \pi_0 \Delta x_t - (1-\alpha)[y_{t-1} - \beta_0 - \beta_1 x_{t-1}] + \varepsilon_t \tag{3.3.1}$$

which is the ECM with –(1-α) as the speed of adjustment and $\varepsilon_{t-1} = y_{t-1} - \beta_0 - \beta_1 x_{t-1}$ as the error correction mechanism which measures the distance of the system away from the equilibrium.

3.4. Granger Causality

X_t is said not to Granger cause Y t if for all h > 0 :

$$F(Y_{t+h} \backslash \Omega_t) = F(y_{t+h} \backslash \Omega_t - X_t) \quad (3.4)$$

where F denotes the conditional distribution and $\Omega_t - X_t$ is all the information in the universe except series X_t : In plain words, X_t is said not to Granger cause Y_t if X cannot help predict to the future Y; Lin (2008).

3.4.1. Theoretical Granger Causality

If X_t and Y_t are two series, X_t is said to Granger cause Y_t if the lagged values of X_t has statistically important information about the future values of Y_t and it is calculated for stationary series. In testing for the absence of Granger-Causality, consider the VAR model below:

$$Y_t = a_0 + a_1 Y_{t-1} + ... + a_p Y_{t-p} + b_1 X_{t-1} + ... + b_p X_{t-p} + u_t \quad (3.4.1)$$

where u_t are the residuals. The hypothesis tests H_0 : b1 = b2 = ::: = b_p = 0 against H_A

Rejection of the null hypothesis implies that there is Granger Causality. The lagged values of X_t in equation 4.2.1 are retained if they add an explanatory power to the regression equation. F-tests are used to determine the retained lagged values of X_t : Also, the null hypothesis of no Granger causality is rejected if there exists at least one lagged value of X_t in equation (3.4.1)

3.5. Model Selection Criteria

One of the main challenges in statistical modelling is selecting a suitable model from a candidate family to characterize the underlying data. Model selection criteria provide useful tools with this regard. The selection criteria determine whether a fitted model offers an optimal balance between goodness-of-fit and parsimony. Ideally, a criteria will identify the model as either being too simplistic to accommodate the data or unnecessarily complex. The most common model selection criteria that will be discussed here is the AIC:

3.5.1. The Akaike's Information Criterion

The Akaike Information Criterion (AIC) was the first model selection criterion introduced in 1973 by Hirotogu Akaike as an extension of the maximum likelihood principle. It became the first model to gain widespread acceptance. Conventionally, the maximum likelihood is applied in this criterion to estimate the parameters of a model once the structure of the model has been specified. The Akaikes idea is to combine the estimation and structural determination into a single procedure. Assuming that a statistical model of M parameters is fitted to the data, to assess the quality of the model fitting, Akaike (1976) introduced an information criterion. The criterion is the Akaike's Information Criterion (AIC) in the literature defined as:

$$AIC = -2\ln[ML] + 2N \quad (3.5.1)$$

where N is the effective number of parameters, and ML is the Maximum Likelihood under the fitted model, Wei (2006). The AIC aims at identifying the best possible model to the unknown true data generating process.

4. Data analysis and results

4.1. Statistical analysis

After exploring the general theory of Co-integration analysis in the preceding chapter, this chapter is aimed at fitting the Error Correction Model to the Rainfall-Population data in Kisumu City. Study results was interpreted to be significant at P<0.05. Much of the analysis will be applied using R statistical software and the result will be graphically displayed.

4.2. Results

4.2.1. Error Correction Model

After building the Granger model as in the equation below:

$$Y_t - 1235.4419 + \varepsilon_t = 0 \quad (4.2.1)$$

an ECM is built by considering the residuals of the model in the Equation 4.2.1.
The ECM involves fitting a regression of the differenced series and the residuals of the fitted Granger Causality model. Thus, the dynamic linear model is used as a result of the inclusion of the residuals. The fitted model will include the residual part considered to be more stable than the differenced part hence the choice of the dynamic linear model. The output of the estimated ECM is as shown in Table (1) below:

Table 2. An Estimation of an ECM for the Two simulated series

	Estimate	Std. Error	t-value	$pr(>\|t\|)$
Intercept	1230	13.04	94.36	2.0×10^{-16}
Differenced Population series	0.01811	0.001565	11.57	2.2×10^{-12}
Residual	1.000	0.021	47.62	2.0×10^{-16}
Residual Standard Error: 18.35 on 29df				
Multiple R^2 : 0.9881				
Adjusted R^2 : 0.9872				
P-value = 2.2×10^{-16}				

All the parameters are significant at 5% level of significance. R^2 value of 0.9881 shows that the overall model fits well to the data. Therefore, the fitted ECM model becomes:

$$Y_t = 1230 + 0.01811\Delta X_t + 1.000\Lambda \quad (4.2.2)$$

where Y_t represents the Rainfall series, X_t the Population series while Λ denotes the Error Correction Component. 1.000 is the speed of adjustment meaning that the system could easily diverge from the equilibrium path.

5. Discussion

Tests done on the rainfall shows stationarity. However, the population series returns non-stationarity. This can be explained by the following facts: From the inspection of the rainfall series, it can be seen that the annual rainfall levels fluctuate about a constant mean level and this trend has been constant though out the recorded years. Though over the years there has been a shift in the rainfall recorded such that a small increment has been observed, it is not significant. The population on the other hand has been observed to be so non-stationary owing to the level of population growth in Kisumu. This is observed by the highly trended almost linear shaped curve. This is so because population is majorly affected by growth rates and this includes birth rates, immigration levels with a relatively low death and emigration rates. Thus, it is expected that the rainfall series becomes the most stable of the two.

The ECM has also been built from the Granger Causality model and the results are shown in section (4.2), Table 2. The model equation is presented in Equation (4.2.1). From the output, the R^2 value is 0.997, an indication that the model fits well to any other data of similar characteristics. The speed of adjustment is however recorded to be positive meaning that the system could easily move away from the long-run equilibrium path, as evidenced by the results. The model can hence be used to analyze data with similar characteristics.

6. Conclusion

The study has presented us with an opportunity to have a wide understanding of the theory of co-integration analysis in time series in non-stationary systems and its application to real life situations. The model building processes have been explored and utilized. Using the AIC criterion, the best lag to be used was selected that acts as the best fit to our model which was lag 2 with the difference between the null and the actual differenced population series being 0.9. After estimation of the parameter of the model, we checked the direction of the causality and found that population Granger causes rainfall at 10% significance level. We then fitted the Error Correction Model and the observed Error Correction Model was built from the Granger Causality Model. The ECM fitted well to the data with an R^2 value of 0.9881 and with the plot of the residuals ascertained to be stationary. Thereafter, the Co-integration model was fitted and its parameters were also found to be significant hence the model fitted well to the data.

Acknowledgement

Much thanks goes to the Kenya Meteorological Department of Kisumu for making data available and also to my supervisors for the assistance they provided throughout the duration of writing this project.

Conflict of Interest
The authors declare that there is no conflict of interest regarding the publication of this paper.

References
Akaike, H. (1976). Canonical correlation analysis of time series and the use of an information criterion in system identification.

Arouri, M., Shabhaz, M., Onhang, R., Islam, F., and Teulon, F. (2014). Envronmental kuznets curve in thailand:co integration and causality analysis. Ipag Business School, 204:1–35.

Asoka, G., Thuo, A., and Bunyasi, M. (2013). Effects of population growth on urban infrastructure and services: a case of eastleigh neighborhood nairobi,kenya. Anthropology and Archaelogy, 1:41–56.

BBC (2012). Kenya. weather.

Boansi, D. (2014). Yield response of rice in Nigeria: a co-integration analysis. American Journal of Agriculture and Forestry, 2:15–24.

Enders, W. (2004). Applied econometric time series. Wiley series in probability and statistics.

Engle, R. and Granger, C. (1987). Co-integration and error correction: representation, estimation and testing. Econometrica, 55:251–276.

Granger, C. and Newbold, P. (1974). spurious regression in econometrics. Journal of econometrics, 2:111–120.

Hansen, S. and Juselius, K. (1995). Cats in rats: Co-integration analysis of time series. Evanson,Illinois.

Hoffman and John (2010). Linear Regression Analysis,Applications and Assumptions. Brigham Young University, Provo, U.T.

Khaemba, A., Mwasi, B., and Opata, G. (2012). Urban landuse suitability assessment using geoinformation techniques for kisumu municipality in kenya. IJJRAS, 13:522–530.

Kuwornu, K., Izideen, P., and Asare-Osei, B. (2011). Supply response of rice in Ghana: A co-integration analysis. Journal of Economics and Sustainable development, 2:1–15.

Kwiatkowski, D., Philips, P., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unt root. journal of Econometrics, 54:78–159.

Landsberg, H. (1975). Man-made climatic changes. The changing global environment,pages 197–234.

Lin, L. (2008). Notes on testing causality. Institute of Economic:Academics sinica, National Chengchi University.

Li, C., Kuang, Y., Huang, N., and Zhang, C. (2013). The long term relationship between population growth and vegetatation cover: An empirical analysis based on the panel data of 21 cities in guangdong province, china. Int.J.Environ. Res.Public Health, 10:660–667.

Newey, W. and West, K. (1987). A simple positive-definite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica, 55:703–708.

Nuru, S. (2012). Climate change and economic growth in a rain fed economy: How much does rainfall variability cost Ethiopia? Ethiopian Economics Association, pages 1–38.

Peng, X., Chen, X., and Cheng, Y. (2009). Urbanization and its consequences. Encyclopedia of life support systems, 2:1–27.

Philip, P. and Perron, P. (1988). Testing for a unit root in time series regression. Biometrica, 75:335–346.

PopulationAction (2014). Kisumu County. Population Action Int., pages 1–4.

RedCross, K. (2013). Emergency appeals kenya:floods 2013. Saving lives changing minds, 38:1–15.

Ruxanda, G. and Botezatu, A. (2008). Spurious regression and cointegration. numerical example:romania's m2 money demand. Romanian Journal of Economic Forecasting, 3:51–62.

Solarin, A. and Shabhaz, M. (2013). Trivariate causality between economic growth, urbanization and electricity consumption in angola:co-integration and causality analysis. MPRA, pages 1–38.

Wei, W. (2006). Time Series Analysis:Univariate and Multivariate Methods. Pearson Addison Wesley, Fox School of Business and Management:Temple University.

Map Showing the Rainfall distribution of Kenya