Journal of Mathematics and Informatics
Vol. 24, 2023, 61-73
ISSN: 2349-0632 (P), 2349-0640 (online)
Published 30 March 2023
www.researchmathsci.org
DOI: http://dx.doi.org/10.22457/jmi.v24a06218

Journal of Mathematics and Informatics

Time Series Analysis of Economic Factors Influencing Deforestation in Tanzania

John Gweba[1*]*, *Isambi Sailon Mbalawata[2]* and *Silas Mirau[3]

[1,3]Department of Applied Mathematics
Nelson Mandela African Institution of Science and Technology
Arusha, Tanzania.
[2]African Institute for Mathematical Sciences, Rwanda.
[2]Email: imbalawata@nexteinstein.org
[3]E-mail: silas.mirau@nm-aist.ac.tz
[*]Corresponding author. [1]E-mail: gwebaj@nm-aist.ac.tz

Received 10 January 2023; accepted 28 February 2023

Abstract. Climate change is a significant contributor to environmental harm and the rise in Atmospheric carbon dioxide, which raises the earth's surface temperature. As forests are the primary mechanism for absorbing carbon dioxide gas and protecting the earth from global warming and unpredictable weather patterns, a high rate of deforestation is to blame for this. In this study, the economic drivers causing deforestation in Tanzania include per capita income, per capita purchasing power, inflation rate, per capita purchasing power, poverty rate, and electricity consumption are investigated. Autoregressive models with exogenous variables (VARX (1) – VARX (3)) models are formulated to analyze the effect of economic variables and forecast the rate of deforestation in Tanzania. The time series data used from 1994 to 2014 were collected in Tanzania, nature of the data suggests the increase in the rate of deforestation as time progresses. In this study, the best model VARX (3, 0) was obtained, and the relationship between the variables through granger causality was obtained. Also, forecasting was carried out for the next 10 years using the best model VARX (3, 0) and Kalman Filters. It was observed that economic variables, especially the poverty rate, have an impact on the rate of deforestation in Tanzania. Furthermore, the graph shows the increasing trend in the rate of deforestation in the coming years in Tanzania.

Keywords: Deforestation, Vector autoregressive with exogenous variables, economic factors, Kalman Filters, Granger causality, Forecasting

AMS Mathematics Subject Classification (2010): 37M10

1. Introduction

Forests and their products have been the main source of food and natural medicines for the human population, adding value to the human economy. It is believed that forests cover one-third of the world's land and source of 75% of the world's freshwater [1]. It has been observed that 25% of the world's people turn on the forest for quartering [2]. Although forests are very important to land-based flora and fauna, it has been ruined through human activities such as deforestation. Deforestation is the process of transforming forest land to zero-forest land through different human activities such as agriculture and the urbanization of cities [3]. Globally from 2015 -2020, 10 million hectares of forests were cleared mostly from agricultural activities (80%) and each minute 2,400 hectares of trees are cut down [1]. Forests are the biggest source of energy globally, about 13% of people in Latin America and the Caribbean, 5% in Asia, and more than 27% in Africa use forests and their products as cooking energy. In Africa, the use of forests and their products like charcoal for the household is estimated to increase due to the high population growth [4]. Global forest areas in Africa and Asia are expected to decrease y 477 million hectares between 1999-2030 [5]. Also, 71% of tropical rainforests were lost between 2000 and 2012 due to commercial agriculture, especially in Africa and Asia [6]. Other studies showed the influence of poor income, which leads to poor technology, economy, education, and gender as the factors influencing the deforestation rate [7]. Poor income contributed to deforestation because people cut down trees and use the land for cash crops, charcoal production, firewood, and plant residue for cooking as they cannot afford modern energy sources such as gas, kerosene, and electricity [8].

Deforestation contributes considerable amounts to the increase in global carbon dioxide gas and methyl hydride gases which have massive impacts on climate change [9]. High concentration of greenhouse gases may lead to a raise in global temperature, a rise in sea level, glacial retreats, climatic shifts, and acidification of the oceans, these affects crop yields due to unpredictable weather patterns [1],10].

Forest loss due to deforestation reduces its ability to confiscate carbon dioxide gas as a consequence increases the atmospheric carbon dioxide which leads to global warming. Deforestation is found to be the second largest source of carbon dioxide after fossil fuel combustion and deforestation accounts for 20%of anthropogenic carbon dioxide emissions [9,10] .

Climate change impacts such as an increase in the local temperature, unpredictable rainfall rate, and health problems have been observed in Brazil and Latin America regions and are expected to increase as the carbon dioxide emission rate increase [12].

Forests in Africa act as the major carbon sink in the world but deforestation reduces this ability [13]. For three decades African and Amazon tropical forests absorbed roughly about 46 billion tonnes of carbon dioxide from the atmosphere in the 1990s but in 2010's it is estimated that only 25 billion tonnes were removed from the atmosphere [3].

Deforestation acts as a driving force for climatic change in Africa and is expected to alter weather patterns such as temperature, precipitation, and extreme heat in the region [2,3]. The study by Nogherotto and others in 2013 found that there is a reduction of about 50% of precipitation in the Congo basin and significant warming up to 4 centigrade [14].

Tanzania specifically is highly affected by deforestation, the rate of deforestation increases each year due to population growth and cities urbanization. Statistics show that the rate of deforestation increased from 1% to 1.47% between 1999 and 2013 [15-17] argues that deforestation in Tanzania is influenced directly by agriculture, overgrazing, wildfires, and charcoal production. These factors are mainly associated with the population increase rate and the poverty rate. Most poor people live in woodland areas and depend on forests and forest products for their survival, also the population is yearly increasing while natural resources are the same [16].

Globally Tanzania is a spotlight in the sub-region concerning massive forest loss and recent reports from Tanzania Forests Services (TFS) agency and FAO show forest cover to be reduced by 372 thousand hectares per year [2,18].

Forests in Tanzania contribute a lot to the mitigation of climate change and are estimated to absorb over 1 billion tonnes of carbon dioxide in above and below-ground biogas, with the woods storing 73% of total carbon [19].

In Tanzania, deforestation reduces the carbon stocks and at the same time contributes to the atmospheric carbon dioxide gas; carbon stocks in Eastern Arcs mountains decrease at the rate of 1.47 billion tonnes yearly, which is equivalent to 2% of the carbon stocks yearly [13].

This paper forecasts economic factors influencing deforestation in Tanzania using the Vector Autoregressive with exogenous variables (VARX) model. VARX(p,q) model explains the vigorous existing relationship between dependent variables and independent variables or that of dependent variables only [20].

In this study, the rate of deforestation as an endogenous variable and per capita income (X1), per capita purchasing power (X2), inflation rate (X3), poverty rate (X4), and electricity consumption (X5) as exogenous variables were studied.

2. Literature surveys

Time series data involves a logical ordering of data points which is usually computed o successively over a certain amount of time, numerically can be denoted as $Y(t) = 0,1,2,3,\ldots,$ whereby **t** stands for time passed on and $\boldsymbol{Y(t)}$ stands for a random variable. Time series events are usually computed while arranged in correct sequential order [21,22]. A time series with one variable is called a Univariate and with several variables, it is called Multivariate time series and can be either continuous or discontinuous. Continuous time series involves recording observations for every occasion of time, while discontinuous involves recording events at discrete time points [21,23].

Types of models

Autoregressive models AR (q)

Autoregressive is the model which presents observations of a variable as linear combinations of previous values **p** and the random shock with a constant term. This model can generally be represented with the following equation;

$$\boldsymbol{y_t = C + \mathrm{A}_1 y_{-1} + \mathrm{A}_2 y_{t-2} + \cdots + \mathrm{A}_p y_{t-p} + \in_t} \quad (1)$$

where by $\boldsymbol{y_t}$, is the series under study, $\boldsymbol{\mathrm{A}_i}$**(1,2,3…p)** are the autoregressive parameters that explain the results of small changes in two successive time series values and $\in_{\boldsymbol{t}}$ is the disturbance term that is assumed to be ordinarily distributed and free distributed with mean and variance equal to zero [24].

Moving average models MA (q)

As in **AR (p)** models, the moving average model of order **q** uses past errors as a variable that describes the series. Moving average models of order q can be illustrated generally in this equation;

$$\boldsymbol{y_t = \mathrm{A}_1 \varepsilon_{t-1} + \mathrm{A}_2 \varepsilon_{t-2} + \cdots + \mathrm{A}_q \varepsilon_{t-q} + \in_t} \quad (2)$$

where, $\boldsymbol{\mathrm{A}_i}$ **(1,2,3…q)** stands for the model specifications, **q** is the degree of the model, $\in_{\boldsymbol{t-j}}$ **(0,1,2.3…q)** are the random error which is presumption to be white noise [24]. Rando m error is usually assumed to follow the statistical distribution with the mean zero and constant variance. The moving average model with order **q** shows the linear relationship between the series' present observations and the residual of one or more past values [22].

Autoregressive moving average models ARMA (p,q)

When an AR model and an MA model are combined, the ARMA model is formed with order terms and moving average terms. Usually, it is written as ARMA(p,q) with **p^{th}** representing the order of terms and **q^{th}** stands for the order of moving average part [25].

Generally, ARMA (p,q) can be illustrated using the following equation;

$$y_t = C + \epsilon_t + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + B_1 \varepsilon_{t-1} + B_2 \varepsilon_{t-2} + \cdots + B_q \varepsilon_{t-q} \quad (3)$$

whereby **C** stands for constant term, $\mathbf{A}_i$ **(I = 1, 2.3…p)** are the parameters of the autoregressive model B_j**(j=1, 2, 3…q)** are the degrees of the moving average model and ϵ_t represent the white noise disturbance.

Vector autoregressive models (VAR)

VARs are among the most successful, dynamic, and undemanding models for the investigation of time series involving more than one quantity. VARs in economics were made popular by Sims (1980) and are the normal extension of the univariate autoregressive model (AR) [26].

Vector autoregressive with exogenous variables model VARX (p)

The VARX model is the extension of the VAR model with the incorporation of exogenous variables. This model gives a critical investigation of the relationship between multiple influencing variables. VARX (p) model incorporates several AR models which form a vector between variables influencing each other. Mathematically the model can be denoted using the following equation;

$$y_t = c + \sum_{i=1}^{p} \Phi_i y_{t-1} + \sum_{i=0}^{q} \Theta_i x_{t-1} + \epsilon_t \quad (4)$$

whereby $y_t = (y_{1t}, y_{2t} \dots y_{kt})^/$ is a vector of multiple variables in time series and $x_t = (x_{1t}, x_{2t} \dots x_{rt})^/$ is the vector of independent t variables, Φ_i and Θ_i are the matrix coefficients and, x_t **are (m*1**) and **(r*1**) column vectors, Φ_i, Θ_i are **(m*m**) and (**m*r**) matrices respectively and $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t} \dots \epsilon_{kt})^/$ is a noise process vector that has zero means and is independent during time t [27].

VARX (p,q) model is very popular in modelling the economics and finances time series data. Currently, due to its flexibility and usability, it is also used to model climate change and other social issues which faces by the human population. For example, the VARX model was used to "forecast effects of cultivation of main crops in Saudi Arabia" and "an empirical study of climate change on carbon dioxide emission in ASEAN 4" [12,28].

This study involves five exogenous (independent) variables which are economic factors and an endogenous (dependent) variable which is the rate of deforestation. In this case, the VAR model is extended to a VAR model with exogenous variables (VARX), where X represents all the exogenous variables present in the time series data of this study [13].

2.1. Condition for stationary

Multivariate time series data can be tested for stationarity by looking at the nature of the graphs if there is a trend or if data fluctuate around a certain number. For non-stationary data, a statistical approach such as the Augmented Dickey-Fuller test (ADF) or unit root is used to test for the stationarity of the series, if the series is not yet stationary differencing is done to make the series stationary.

In the unit root test with lag **p**, the model with constant **C** can be illustrated as;

$$\Delta X_t = C + \mathrm{A}X_{t-1} + \sum_{i=1}^{p-1} \mathrm{A}_1 \Delta X_{t-1} + U_t \tag{5}$$

where, $\Delta X_t = X_t - X_{t-1}$, U_t stands for the white noise, $H_0 : \mathrm{A} = 0$ and $H_a : < 0$ stands for the null hypothesis and alternative hypothesis respectively. The test statistics is τ (tau) with the distribution approximately t-distribution. For the level of significance $\alpha = 0.05$, the null hypothesis is rejected if

$\tau < -2.57$ **or if** $p - value < 0.05$ [20], [29], [30].

The test statistics are given by:

$$ADF\ \tau = \frac{\Phi}{Se(\Phi)} \tag{6}$$

2.2. Granger causality test

Granger causality tests the consequential relationship between the time series variables, it was initiated by Granger in 1969. Granger variables influence other variables to increase if the coefficients are positive in the model [31]. In time series involving more than one variable, observed variables are included in autoregressive vector models (VAR) for **x**. The test of **x** granger causes **y** can be represented in the following model:

$$y_t = C + \mathrm{A}_1 y_{t-1} + \mathrm{A}_2 y_{t-2} + \mathrm{A}_3 y_{t-3} + \Theta_1 x_{t-1} + \Theta_2 x_{t-2} + \Theta_3 x_{t-3} + \Theta_4 x_{t-4} + \Theta_5 x_{t-5} + \epsilon_t \tag{7}$$

2.3. Forecasting using Kalman filters (KF)

Forecasting is the main objective of the analysis of time series data. Multivariate model forecasting usually resembles that of the univariate model, but due to the nature of the data best VAR model can be fitted and the estimation can be done using Kalman Filter under maximum likelihood and henceforth forecasting.

A Kalman filter is a repetitious process that allows the recovery of least squares solutions through adding data successively. Usually, the data adjustment procedure is continuous and updated with new data acquired with the probability distribution of the given model [32]. This technique has got several practices such as building a Global Navigation Satellite System (GNSS), and solving time series problems because can filter even a small amount of data acquired and produce the required results compared to other methods [33,34].

In the dynamic system prediction and filtering may encounter a lot of limitations that have an impact on their practical application. Kalman filter method provides an efficient iterative means to approximate the state space and is very influential in solving linear filtering problems [35]. This implies that the model presumes that the state of a system at time **t** is grounded on the prior state at **t-1** and gets an observation y_t of the true state x_t at time **t** according to the state space mode:

$$x_T = A_t x_{t-1} + w_t \qquad \textbf{(System equation)} \tag{8}$$

$$y_t = H_t x_t + v_t \qquad \textbf{(Observation equation)} \tag{9}$$

where $\boldsymbol{x_t}$ stands for the state vector at time step **t**, A_t illustrates the state transition matrix of the system state parameter at step **t-1,** $\boldsymbol{w_t}$ is the process noise vector at time step **t,** $\boldsymbol{y_t}$ shows the vector computations, $\boldsymbol{H_t}$ denotes the transformation matrix that maps the state vector parameters into the measurement domain, $\boldsymbol{v_t}$ is the measurement noise term at time step **t.**

3. Results and discussion

The time series data used in this paper are the rate of deforestation as an endogenous variable denoted by **Y** and five exogenous variables (explanatory variables) which are economic factors influencing deforestation in Tanzania denoted by $\mathbf{X_i}$**, i=(1,2.3…5)**. The economic factors influencing the rate of deforestation in Tanzania are presented in table 1 below with their computations as follows; **Y** denotes the rate of deforestation in Tanzania (ha/year) and is the depletion in total areas of the forest cover yearly with a negative sign, $\mathbf{X_1}$ is the per capita income computed in US dollars at the market price, $\mathbf{X_2}$ represents per capita purchasing power over time (constant at 2011US dollars), X_3 is inflation rate over time (% change in consumer price index), $\mathbf{X_4}$ is the poverty rate over time (computed as household consumption% per capita) and $\mathbf{X_5}$ is the electricity consumptions (as per electricity price) per population over time (KWh)

Table 1: Data for economic factors influencing deforestation in Tanzania from 1994-2014

YEAR	X1	X2	X3	X4	X5	Y
1994	170.26	1363.304	34.1	81.24	49.259	-400400
1995	171.31	1370.84	29.4	83.12	57.184	-400320
1996	200.48	1394.497	28	82.81	60.715	-400300
1997	250.44	1407.133	18.1	87.26	55.211	-400280
1998	277.21	1423.74	15.8	80.28	60.079	-400211
1999	301.2	1456.2	9.9	81.15	55.651	-400210
2000	308.41	1489.65	9.9	78.27	58.25	-400206
2001	306.24	1538.05	9.15	74.97	61.979	-400101
2002	320.21	1604.37	10.55	71.93	67.298	-400090
2003	325.25	1667.92	9.55	69.73	67.162	-400073
2004	338.05	1747.9	9.15	66.94	78.634	-400062
2005	456.16	1836.15	9.35	66.8	78.355	-400060
2006	495.21	1864.77	10.45	64.02	64.206	-400043
2007	503.17	1961.28	9.05	61.36	80.126	-400035
2008	607.23	2006.64	8.25	63.96	84.351	-400024
2009	685.34	2048.89	12.15	66.29	70.553	-400001
2010	788.22	2111.2	7.2	68.37	93.869	-372816
2011	790.38	2206.91	7.65	65.22	84.651	-372701
2012	897.23	2247.86	10	65.58	94.602	-372670
2013	930.38	2335.96	3.9	64.77	89.478	-372231
2014	945.14	2421.21	3.13	64.57	89.111	-372000

Usually, a VAR process is influenced by the presence of exogenous variables which can be stochastic or non-stochastic. Also, it's highly affected by the lags of the present exogenous variables.

The VARX (p,q) model is expressed in equation (4) above, when p=3 and q=0, then VARX (3,0) with one endogenous variable and five exogenous variables becomes;

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \\ y_{5t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix} + \begin{bmatrix} \Phi^1{}_{11} & \Phi^1_{12} & \Phi^1_{13}\,\Phi^1_{14} & \Phi^1_{15} \\ \Phi^1_{21} & \Phi^1_{22} & \Phi^1_{23}\,\Phi^1_{24} & \Phi^1_{25} \\ \Phi^1_{31} & \Phi^1_{32} & \Phi^1_{33}\,\Phi^1_{34} & \Phi^1_{35} \\ \Phi^1_{41} & \Phi^1_{42} & \Phi^1_{43}\,\Phi^1_{44} & \Phi^1_{45} \\ \Phi^1_{51} & \Phi^1_{52} & \Phi^1_{53}\,\Phi^1_{54} & \Phi^1_{55} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \\ y_{4t-1} \\ y_{5t-1} \end{bmatrix} +$$

$$\begin{bmatrix} \Phi^2{}_{11} & \Phi^2_{12} & \Phi^2_{13}\,\Phi^2_{14} & \Phi^2_{15} \\ \Phi^2_{21} & \Phi^2_{22} & \Phi^2_{23}\,\Phi^2_{24} & \Phi^2_{25} \\ \Phi^2_{31} & \Phi^2_{32} & \Phi^2_{33}\,\Phi^2_{34} & \Phi^2_{35} \\ \Phi^2_{41} & \Phi^2_{42} & \Phi^2_{43}\,\Phi^2_{44} & \Phi^2_{45} \\ \Phi^2_{51} & \Phi^2_{52} & \Phi^2_{53}\,\Phi^2_{54} & \Phi^2_{55} \end{bmatrix} \begin{bmatrix} y_{1t-2} \\ y_{2t-2} \\ y_{3t-2} \\ y_{4t-2} \\ y_{5t-2} \end{bmatrix} +$$

$$\begin{bmatrix} \Phi^3{}_{11} & \Phi^3_{12} & \Phi^3_{13}\,\Phi^3_{14} & \Phi^3_{15} \\ \Phi^3_{21} & \Phi^3_{22} & \Phi^3_{23}\,\Phi^3_{24} & \Phi^3_{25} \\ \Phi^3_{31} & \Phi^3_{32} & \Phi^3_{33}\,\Phi^3_{34} & \Phi^3_{35} \\ \Phi^3_{41} & \Phi^3_{42} & \Phi^3_{43}\,\Phi^3_{44} & \Phi^3_{45} \\ \Phi^3_{51} & \Phi^3_{52} & \Phi^3_{53}\,\Phi^3_{54} & \Phi^3_{55} \end{bmatrix} \begin{bmatrix} y_{1t-3} \\ y_{2t-3} \\ y_{3t-3} \\ y_{4t-3} \\ y_{5t-3} \end{bmatrix} + \begin{bmatrix} \Theta_{11} \\ \Theta_{12} \\ \Theta_{13} \\ \Theta_{14} \\ \Theta_{15} \end{bmatrix} x_t + \varepsilon_t \qquad (10)$$

3.1. Stationarity test

The stationary test was done using the Augmented Dickey-Fuller Unit root test (ADF). The results of the ADF test showed that all the variables were not stationary as the probability was < 0.05. Since all the variables were not stationary, first differencing is performed to make all the data stationary.

After first differencing all the variables were stationary with the value of probability amounting to < 0.05. Table 2 shows the results of the ADF test before and after differencing.

Table 2: ADF stationarity test results

Levels			**1st difference**	
Var trend	**t-stat**	**prob**	**t-stat**	**prob**
Per capita income	1.0206	0.9945	-2.9443	0.0404
Capita purchasing power	-1.173	0.6851	-3.5004	0.008
Inflation rate	-2.1431	0.2275	-3.1675	0.0219
Poverty rate	0.6952	0.9897	-4.4661	0.0002
Electricity consumptions	-0.492	0.8936	-4.19	0.0007
Rate of deforestation	-2.0521	0.2642	-4.4075	0.0003

3.2. Building VARX (p,q) model

To find the best model, the correct lag selection is required through AIC and BIC. The selected lag value was obtained by using the smallest value of AIC and BIC. VAR analysis as shown in Table 3, compares (VARX (1, 0) - VARX (3, 0)) models, and using AIC and BIC values it was observed that the smallest value correlated with VARX (3, 0) model. The minimum values of AIC and BIC were 355.876 and 363.889 respectively.

Statistical test results for parameters estimation in VARX (3, 0) model is given in Table 3.

Table 3: Statistical test results for fitting VARX (3, 0) model

Parameters	Values	Standard Error	t -values	p-values
Constant(1)	$-3.1343*10^5$	$1.2588 * 10^5$	-2.4899	0.012776
AR(1)(!,1)	0.54463	0.1459	3.7329	0.00018927
AR(2)(1,1)	-0.30143	0.18499	-1.6295	0.10321
AR (3)(!,1)	0.57777	0.22	0.2662	0.79284
Theta(1,1)	53.437	26.057	2.0508	0.040288
Theta(1,2)	30.221	28.882	-1.0464	0.29539
Theta(1,3)	-573.89	420.36	-1.3652	0.17219
Theta(1,4)	521.19	439.04	1.1871	0.23518
Theta(1,5)	486.59	148.98	3.2661	0.0010904

From Table 3, AR1 to AR3 parameters are consequential with – and + coefficients in the VARX (3, 0) model. Therefore, VARX (3, 0) is chosen as the best model for the prediction of the rate of deforestation in Tanzania.

Equation (4) gives the general form of the VARX (p, q) model equation, therefore VARX (3, 0) model univariate regression equation can be written as;

$$\boldsymbol{y_t = C + \Phi_1 y_{t-3} + +\Theta_1 x_{t-1} + \Theta_2 x_{t-2} + \Theta_3 x_{t-3} + \Theta_4 x_{t-4} + \Theta_5 x_{t-5} + \epsilon_t} \quad (6)$$

The coefficients thetas from equation (6) above act for the change in deforestation for a unit in the change of all exogenous variables while keeping other variables constant. Substituting the statistical test parameters results from Table 3 into equation 6, the univariate regression equation becomes:

$$\boldsymbol{y_t = -3.1343 * 10^5 + 0.57777 y_{t-3} + 53.437 x_{t-1} + 30.221 x_{t-2} - 573.89 x_{t-3} + 521.19 x_{t-4} + 486.59 x_{t-5} + \epsilon_t}$$

The constant term **-3.1343 *10^5** denotes the mean rate of deforestation which always is obtained after assuming the values of other variables to be zero. The minus sign shows the forest cover change which logically stands for the deforestation rate.

The probability value of each variable test assumes the coefficient to be zero or no effect. A probability value < **0.05** indicates the existence of a strong relationship or correlation between variables and the hypothesis in question may not describe the observations sufficiently, so required to deny the null hypothesis.

The probability value> **0.05** indicates the existence of a weak relationship or correlation which is against the assumptions of a null hypothesis, hence, the null hypothesis should not be rejected.

In this paper, probability values of **p=0.05** were used to test the results of all exogenous variables $\boldsymbol{x_i}$ to the endogenous variable **y.** All exogenous variables $\boldsymbol{x_1, x_2, x_3, x_4}$ ***and*** $\boldsymbol{x_5}$ show a strong association with the deforestation rate in Tanzania due to their probability values being less than **0.05.** Poverty rate specifically $(\boldsymbol{x_4})$ with a positive coefficient bigger than all other coefficients indicating great associations with the deforestation rate. This explains the fact that many people cannot afford to buy other alternative sources of energy such as kerosene and liquefied gases, so they turn to the cheapest sources like charcoal which continue to accelerate the deforestation rate in Tanzania.

Table 4: Analysis of Variance (ANOVA)

	Sum Squares	Df	Mean Squares	R^2	F	P-value
Model	2.5409e+09	5	5.0819e+08	0.871122064	20.278	3.4409e-06
Residual	3.7592e+08	15	2.0819e+08			
Total	2.9169e+09	20	1.4584e+08			

Table 4 above shows the model results which are very significant with a p-value less than **0.05** and an R square of **0.8711**, this indicates that 87% of the deforestation rate is influenced by the exogenous variables incorporated in the VARX (3, 0) model. With the highest R –square values, VARX (3, 0) model fits correctly with the time series data used. Also, the remaining 13% describes other factors which were not considered in this study The results of the analysis of variance shown in Table 4 suggest using VARX (3, 0) model in forecasting the deforestation rate in Tanzania. Results are numerically significant and indicate all the factors have contributed to the deforestation rate with a p-value less than **0.05** of recommended conditions.

3.3. Granger causality

In Table 5 below, each column represents predictor variable x and the rows represent responses y and **p**-values of each pair of the variables. The values in row 4 column 1 and row 4 column 5 have a great value than **0.05**, this shows that per capita income and electricity consumption do not necessarily cause poverty rate. All other variables p –values are less than **0.05** except for diagonals. For those values < **0.05,** the null hypothesis is rejected and the time series data is better for VARX (p,q) modeling and forecasting.

Table 5: Granger causality test

	Per capita income_x	Percapitapurchasing power_x	Inflation rate_x	Poverty rate_x	Electricity consumptions_x	Rate of deforestation_x
Per capita income_y	1.00000	0.00000	0.00052	0.01025	0.00053	0.00005
Per capita purchasing power_y	0.00000	1.00000	0.00002	0.09419	0.00002	0.00175
Inflation rate_y	0.00000	0.00000	1.00000	0.00000	0.00000	0.00000
Poverty rate_y	0.19564	0.00004	0.00304	1.00000	0.23179	0.00000
Electricity consuptions_y	0.00000	0.00000	0.04695	0.00223	1.00000	0.03067
Rate of deforestation_y	0.00000	0.00000	0.00000	0.00000	0.00008	1.00000

3.4. Forecasting

In time series data forecasting allow the approximation of the unknown upcoming values that are specifically used for predicting the forecasted values. In this specific paper, the VARX (3, 0) model was used to predict the rate of deforestation in Tanzania for the next 10 years. In Fig 3 VARX (3, 0) with the blue line is a bit compatible with the original data with the bolded green line. According to the rate of deforestation prediction data for the next 10 years appears to form an increasing trend

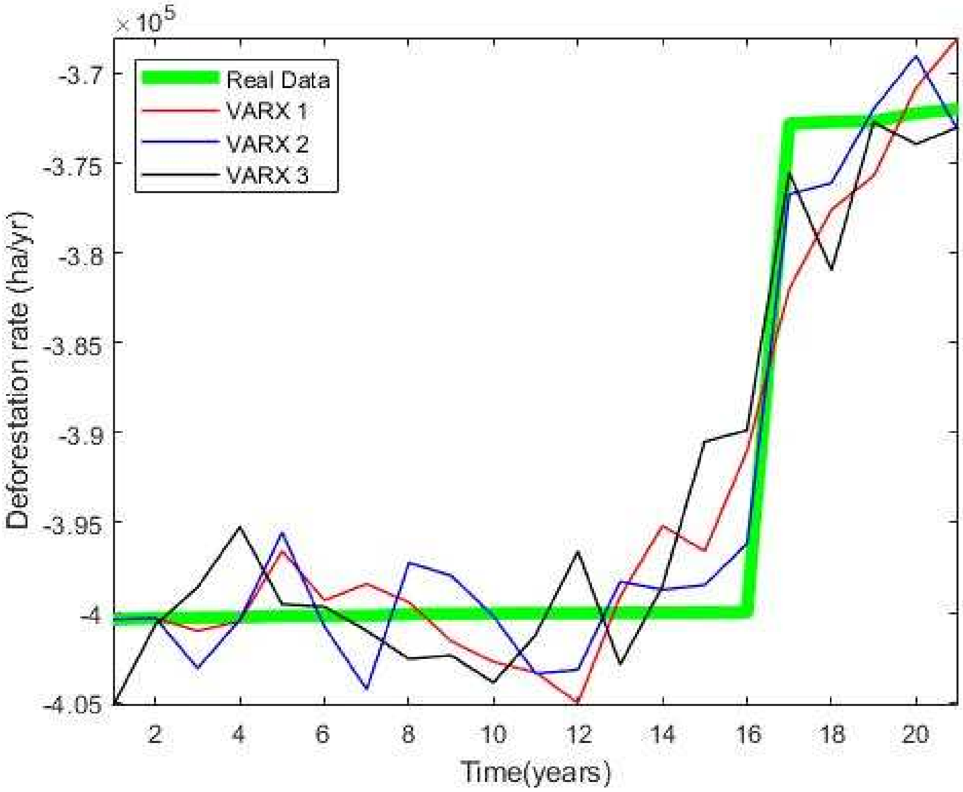

Figure 1: Model fitting for Real data and VARX (1)-VARX (3) models

4. Conclusion

Results on the relationship between the endogenous variable (rate of deforestation) and all five explanatory or exogenous variables indicate that VARX (3, 0) is the best model for the relationship between these variables. The univariate regression model equation from VARX (3, 0) indicates the suitability of the model in the prediction through the coefficients of thetas and constants. The Granger causality test shows that all variables contribute significantly to the rate of deforestation in Tanzania. The VARX (3, 0) model for prediction is very consequential and the predictive value is very near to the real observation. This explicitly explains the model as suitable and reliable for forecasting the rate of deforestation in Tanzania for 10 years. As the forecasted values show in the graph, there is an increasing trend in the rate of deforestation in Tanzania, this must be taken seriously as the impacts of deforestation on climate change and the human population at large have been explained explicitly in this study. Also, there is a need of controlling the population using forests and their products as sources of food and source of energy, to preserve forests from deforestation. The Government of Tanzania should also take into consideration raising the living standards of its citizens as the poverty rate continues to strong contributions to deforestation since many people run to the cheapest sources of energy, which are mostly sources of deforestation

Acknowledgement. The author is very thankful to the reviewer for the studious effort to improve this manuscript. Also, we would like to thank the Tanzanian National Bureau of Statistics and National Forest Resources Monitoring and Assessment of Tanzania for giving researchers access to the data used in this study, the African Institute for Mathematical Sciences (AIMS) Rwanda for supporting this study financially, and the National Institute of Transport (NIT) for sponsoring the author's studies at NM-AIST. Lastly, the author thanks the co-authors for their directives and assistance.

Conflicts of Interest. The authors declare no conflicts of interest.

Authors' contributions. All authors contributed equally to this work.

REFERENCES

1. S.E.Pacheco *et al.*, Climate change and global issues in allergy and immunology, *J. Allergy Clin. Immunol.*, 148(6) (2021) 1366–1377. doi: 10.1016/j.jaci.2021.10.011.
2. H.O.Bergesen, G.Parmann, and O.B.Thommessen, *Food and Agriculture, Organization (FAO)*. 2019.
3. IPCC, *Climate Change 2014 Part A: Global and Sectoral Aspects*. 2014.
4. FAO, The state of Food and Agriculture in 2014 in Brief, 2014.
5. J.A.Turner, J.Buongiorno and S.Zhu, An economic model of international wood supply, forest stock and forest area change, *Scand. J. For. Res.*, 21(1) (2006) 73–86. doi: 10.1080/02827580500478506.
6. S.Lawson, A.Blundell, B.Cabarle, N.Basik, M.Jenkins and K.Canby, Consumer goods and deforestation: an analysis of the extent and nature of illegality in forest conversion for agriculture and timber plantations, *For. Trends Rep.*, 2012, 2014.
7. S.W.H.Reetz, S.Schwarze and B.Brümmer, Poverty and tropical deforestation by smallholders in forest margin areas: evidence from central Sulawesi, Indonesia, 2012.
8. FAO, *Analysis of forests and climate change Analysis of forests and climate change in Eastern Africa*. 2017.
9. G.R.Van Der Werf *et al.*, CO_2 emissions from forest loss, *Nat. Geosci.*, 2(11) (2009) 737–738, doi: 10.1038/ngeo671.
10. D.Gerten *et al.*, Asynchronous exposure to global warming: Freshwater resources and terrestrial ecosystems, *Environ. Res. Lett.*, 8(3) (2013). doi: 10.1088/1748-9326/8/3/034032.
11. O.Hoegh-Guldberg, D.Jacob and M.Taylor, Impacts of 1.5°C of Global Warming on Natural and Human Systems, *Spec. Report, Intergov. Panel Clim. Chang.*, (2018) 175–181.
12. P.T.N.Putra, L.Yuliati and E.K.Lestari, An empirical study of climate change on carbon dioxide emission in ASEAN 4: Vector Autoregression Exogenous (VARX)

Approach, *Eko-Regional J. Pengemb. Ekon. Wil.*, 14(2) (2019) 109–120. doi: 10.20884/1.erjpe.2019.14.2.1335.

13. W.Hubau *et al.*, Asynchronous carbon sink saturation in African and Amazonian tropical forests, *Nature*, 579(7797) (2020) 80–87.doi:10.1038/s41586-020-2035-0.
14. R.Nogherotto, E.Coppola, F.Giorgi and L.Mariotti, Impact of Congo Basin deforestation on the African monsoon, *Atmos. Sci. Lett.*, 14(1) (2013) 45–51. doi: 10.1002/asl2.416.
15. T.O.T.H.E.Unfccc, The united republic of tanzania tanzania ' s forest reference emission level submission to the UNFCCC, November, 2017.
16. J.R.Kideghesho, Realities on deforestation in Tanzania - trends, drivers, implications and the way forward, *Precious For. - Precious Earth*, 2015, doi: 10.5772/61002.
17. P.M.Luhunga, A.L.Kijazi, L.Chang'a, A.Kondowe, H.Ng'ongolo and H. Mtongori, Climate change projections for Tanzania based on high-resolution regional climate models from the coordinated regional climate downscaling experiment (CORDEX)-Africa, *Front. Environ. Sci.*, 6 (2018) 1–20.
doi: 10.3389/fenvs.2018.00122.
18. Tanzania Forest Service Agency, National forest resources monitoring and assessment of Tanzania mainland, 124 (2015).
19. FAO, Global Forest Resources Assessment 2020; Report United Republic of Tanzania, (2020) 15–16.
20. E.Russel and M.Usman, Vector autoregressive with exogenous variable model and its application in modeling and forecasting energy data : case study of PTBA and HRUM Energy, 9(2) (2019) 390–398.
21. S.Farag, Application of time series models to sets of environmental data, *Ed. Cowan Univ.*, (1994).
22. R.Adhikari and R.Agrawal, An introductory study on time series modeling and forecasting, *arXiv Prepr. arXiv1302.6613*, vol. 1302.6613, pp. 1–68, 2013.
23. J.-X.Deng and G. Gan, A survey of stock forecasting model based on artificial intelligence algorithm, *J. Math. Informatics*, 7 (2017) 73–78, doi: 10.22457/jmi.v7a9.
24. K.W.Hipel, A.I.McLeod and W.C.Lennox, Advances in Box-Jenkins modeling: 1. Model construction, *Water Resources Research*, 13(3) (1977) 567–575, doi: 10.1029/WR013i003p00567.
25. J.Mohamed, Time series modeling and forecasting of somaliland consumer price index: a comparison of ARIMA and regression with ARIMA errors, *Am. J. Theor. Appl. Stat.*, 9(4) (2020) 143, doi: 10.11648/j.ajtas.20200904.18.
26. H.Lütkepohl, Vector autoregressive models, *Handb. Res. Methods Appl. Empir. Macroecon.*, (2016) 139–164, 2013, doi: 10.4337/9780857931023.00012.
27. Y.Kim and S.Kim, Electricity load and internet traffic forecasting using vector autoregressive models, *Mathematics*, 9(18) (2021), doi: 10.3390/math9182347.
28. O.S.Alnashwan and M.M.Alderiny, Using VARX model to forecast the effect of cultivation of main crops on water security in Saudi Arabia, *J. Exp. Biol. Agric. Sci.*, 5(2320) (2017) 126–135. doi: 10.18006/2017.5(spl-1-safsaw).s126.s135.
29. A. T. Jebb, L. Tay, W. Wang, and Q. Huang, "Time series analysis for psychological research: Examining and forecasting change," *Front. Psychol.*, 6 (2015) 1–24, doi: 10.3389/fpsyg.2015.00727.
30. A.R.Putri, M.Usman, Warsono, Widiarti and E.Virginia, Application of vector

autoregressive with exogenous variable: case study of closing stock price of PT INDF.Tbk and PT ICBP.Tbk, *J. Phys. Conf. Ser.*, 1751(1) 2021, doi: 10.1088/1742-6596/1751/1/012012.

31. C.J.W.Granger, Investigating causal relations by econometric models and cross-spectral methods authors (s): C.W.J.Granger Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/1912791 Accessed : 25-03-2016 19 : 26 UTC Your use of the JS, *Econometrica*, 37(3) (1969) 424–438.
32. M.Pulido, P.Tandeo, M.Bocquet, A.Carrassi and M.Lucini, Stochastic parameterization identification using ensemble Kalman filtering combined with maximum likelihood methods, *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, 70(1) (2018) 1–17. doi: 10.1080/16000870.2018.1442099.
33. M. Dalaison and R. Jolivet, A Kalman filter time series analysis method for InSAR, *J. Geophys. Res. Solid Earth*, 125(7) (2020) 1–21. doi: 10.1029/2019JB019150.
34. J.P.Muhirwa, I.S.Mbalawata and V.Grace Masanja, Sensitivity and uncertainty analysis of variable-volume deterministic model for endothermic continuously stirred tank reactor, *J. Math. Informatics*, 20 (2020) 73–89. doi: 10.22457/jmi.v20a08189.
35. J.Wu, Q.Fang, Y.Xu, J.Su and F.Ma, Kalman filter based time series prediction of cake factory daily sale, *Proc. - 2017 10th Int. Congr. Image Signal Process. Biomed. Eng. Informatics, CISP-BMEI 2017*, vol. 2018-Janua, no. July 2019, pp. 1–7, 2018, doi: 10.1109/CISP-BMEI.2017.8302108.