VECTOR AUTOREGRESSIVE MODELING OF CROP PRODUCTION INDEX – PERMANENT CROPLAND RELATIONSHIP IN NIGERIA

Saheed Busayo Akanni[1], Kola Yusuff Kareem[2], Adewole Oluranti Grace [1], Muideen Ojo Alabi[1], Saheed Olalekan Jabaru[3], Muhammed Tahir Muhammed[4], Ekundayo Samuel O[5], Olakiitan Ibukun Adeniyi[1], Magdalene Peter[6], Oluwole Joshua Oyerinde[7]

[1] **University of Ilorin, Ilorin - Nigeria, Department of Statistics**
[2] **University of Ilorin, Ilorin - Nigeria, Department of Agricultural and Biosystems Engineering**
[3] **Al-Hikmah University, Ilorin – Nigeria, Department of Physical Sciences (Statistics Option)**
[4] **University of Abuja, Abuja – Nigeria, Department of Statistics**
[5] **University of Ilorin, Ilorin – Nigeria, Department of Finance**
[6] **Kaduna State University, Kaduna – Nigeria, Department of Mathematical Sciences, Statistics Unit**
[7] **University of Ilorin, Ilorin– Nigeria, Department of Mechanical Engineering**

Corresponding author: Saheed Busayo Akanni, 08-55eg131pg@students.unilorin.edu.ng

ABSTRACT: There is currently no study that focuses solely on the causal relationship between Crop Production Index (CPI) and Permanent Cropland (PCL) in developed and developing countries, especially in Nigeria. However, understanding the causal relationship between CPI and PCL is crucial to both food security and economic growth of any nation. In this paper, we investigated the causal relationship between CPI and PCL time series variables using unrestricted Vector Autoregressive (VAR) modeling techniques. Pre-examination of the CPI and PCL time series data extracted from the repository of World Bank showed that these two series were not only difference stationary series of order one {I(1)} but are also not cointegrated. The results of optimal lag length confirmed that VAR (3) model best fitted the data. The findings showed that Nigeria's crop production index is predictable by Nigeria's permanent cropland and vice versa.
KEYWORDS: Crop Production Index, Permanent Cropland, Unrestrited VAR, Optimal Lag Length, Cointegration, Nigeria

1. INTRODUCTION

It is a well established fact that the quantity of crop production or agricultural produce is majorly determined by the size of cultivated land used in planting crops [2][4][14]. Other factors that may be considered for effective crop production include agro-climatic, edaphic, biotic, socio-economic and crop management factors. Crop production is a branch of agriculture that deals with growing crops for use as food and fibre. These crop produce include, but not limited to grains, cotton, tobacco, vegetables, nuts and plants. One important measure of crop production is the crop production index, which is a measure of crop production for each year relative to the base period i.e. 2004-2006 and includes all crops except fodder crops [21]. Unlike crop production, the availability of cultivable land as defined by permanent cropland is considered as a major factor influencing crop production index. Permanent cropland in this context, refers to a land cultivated with crops that occupy the land for long periods and need not be replanted after each harvest such as cash crops like cocoa, coffee and rubber [21]. In other words, permanent cropland covers all land areas under nut trees, flowering shrubs, vines and fruit trees, with the exception of land areas with trees solely grown for timber. It has been reported that the total global land area is about 13.2 billion ha, with about 12.6% i.e. 1.6 billion ha currently being used for crop cultivation; 35% i.e. 4.6 billion ha being harnessed as woodlands and grasslands; and 28% i.e. 3.7 billion ha as forest [17]. Therefore, understanding the causal relationship between crop production index and permanent cropland is not only crucial to food security and reliable food chains, but also to employment opportunities, tourist attraction and economic growth of any nation.

The use of multivariate time series techniques to model and analyze agricultural variables has become common in recent research works. This is because; available research data on most agricultural variables are often time series data. Another reason for modeling agricultural time series variables using multivariate time series techniques is that, most of these variables usually undergo difference stationary processes before they become stationary. The commonly used multivariate time series technique in the time series econometrics literature are Vector Autoregressive (VAR) model and Vector Error

Correction model (VECM) [5] [11][12]. The former (i.e. VAR) is desirable when time series variables under consideration are difference stationary series of order one {I(1)} and there is no evidence of cointegration between or among the series. But when all the time series variables under consideration are difference stationary series of order one and there is evidence of cointegration between or among the series, the latter (i.e. VECM)is desirable. However, the Autoregressive Distributed Lag (ARDL) model developed by [16] can also be used to analyze difference stationary time series variables of order one {I(1)}depending on the objective of the study. Moreover, regression analysis techniques can be used to model multivariate time series variables if and only if such variables are stationary at levels. Otherwise, the regression analysis approach yields spurious or misleading results [12].

Little efforts have been made to examine the relationship between both crop production index and permanent cropland in developed and developing countries, especially in Nigeria. [8] applied structural equation modeling techniques to analyze the relationship among arable production per capita index, arable production and permanent cropland and forest area. Results from their study showed that arable production per capita index is impacted more by population while the influence of rainfall on the arable production per capita index is weak. To access the determinants of agricultural land expansion in Nigeria [15] used Error Correction Model (ECM) technique to show that cropland growth rates, agricultural production index, livestock population, human population, other land and cereal cropland growth rates have significant impact on agricultural land expansion. In a study, [9] used Toda-Yamamoto techniques to show that cereal yields in Nigeria can be predicted by both cereal production and the size of farmland used for planting cereal crops, and recorded a unidirectional causality relationship among the three variables – cereal production, cereal yield and land under cereal production.

[20] studied the impact of land-take on the land resource base for crop production in the European Union using spatial techniques. Results from the spatial analysis later revealed that increasing land-take due to urbanization threatens availability of fertile soils throughout Europe. [1] used ARIMA (1, 1, 1) model to show that production of cereal crops in Nigeria will continue to increase for the foreseeable future. [19] examined the impacts of nitrogen fertilizer on global crop production. His findings revealed that only 30-40% of applied nitrogen fertilizer is taken up by crops. [10] investigated the effects of biomass crop production on representative southeastern United States farms using Micro-Oriented Agricultural Production System (MOAPS). Their findings showed that erosion was likely to be reduced more by the diversion of cropland to permanent vegetative cover on farms similar to the more highly erodible Major Land Resource Area (MLRA) 134 farms than on the less erodible MLRA 133 farms. To fill the gap in the literature, this study therefore applied the Vector Autoregressive (VAR) modeling techniques to analyze the causal relationships between crop production index and permanent cropland in Nigeria.

2. METHODOLOGY

This work used time series data on crop production index (CPI) and permanent cropland (PCL) in Nigeria spanning 1961 to 2016. The data was extracted from the repository of World Bank via their website *http://data.worldbank.org*.In this work, we used the Vector Autoregressive (VAR) model developed by [18] to analyze the causal relationships between CPI and PCL in Nigeria. In VAR modeling, endogenous variables in the system are usually treated as a function of the lagged values of all of the endogenous variables in the system. Mathematically, the general VAR (p) model is of the form:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} + \dots + A_p y_{t-p} + Bx_t + e_t \qquad (1)$$

where:

$y_t$is a k vector of endogenous variables,x_t is a d vector of exogenous variables,

$A_1 to A_p$ and B are matrices of coefficients to be estimated,e_t is a vector of innovations that may be contemporaneous but is uncorrelated with its own lagged values and uncorrelated with all of the right hand side variables. According to Engle and [7], the VAR (p) model in equation (1) is desirable when two or more time series variables are difference stationary series of order one {I(1)} and are not cointegrated.

2.1 Model Specification for the CPI and PCL Series

The bivariate Vector Autoregressive model of order p {VAR (p)} for crop production index (CPI) and permanent cropland (PCL) series is of the form:

$$\mathrm{CPI}_t = \alpha_{11}\mathrm{CPI}_{t-1} + \alpha_{12}\mathrm{CPI}_{t-2} + \cdots + \alpha_{1p}\mathrm{CPI}_{t-p} + \beta_{11}\mathrm{PCL}_{t-1} + \beta_{12}\mathrm{PCL}_{t-2} + \cdots + \beta_{1p}\mathrm{PCL}_{t-p} + c_{11} + e_{1t} \qquad (2)$$

$$\mathrm{PCL}_t = \alpha_{21}\mathrm{CPI}_{t-1} + \alpha_{22}\mathrm{CPI}_{t-2} + \cdots + \alpha_{2p}\mathrm{CPI}_{t-p} + \beta_{21}\mathrm{PCL}_{t-1} + \beta_{22}\mathrm{PCL}_{t-2} + \dots + \beta_{2p}\mathrm{PCL}_{t-p} + c_{22} + e_{2t} \qquad (3)$$

The Augmented Dickey-Fuller (ADF) and Johansen cointegration tests have been used to determine the order of integration and cointegration status of the CPI and PCL series respectively. For the procedures for conducting the ADF and Johansen cointegration tests, see [6] and [13]. However, the number of lag "p" to be included in the VAR (p) model stated as equations (2) and (3) was determined using selection criteria such as Sequential Modified Likelihood Ratio test statistic (LR), Final Prediction Error (FPE), Akaike Information Criteria (AIC), Schwarz Information Criteria (SIC) and Hannan-Quinn Information Criteria (HQ). Once the VAR (p) model in equations (2) and (3) are estimated, the next thing is to test the connection between VAR and causality. To achieve this, the following hypotheses were fromulated with respect to equations (2) and (3):

Hypothesis for equation (2):
H_{01}: $\beta_{11} = \beta_{12} = \cdots = \beta_{1p} = 0$ (PCL does not Granger cause CPI)

Hypothesis for equation (3):
H_{02}: $\alpha_{21} = \alpha_{22} = \cdots = \alpha_{2p} = 0$ (CPI does not Granger cause PCL)
The decision rule is to reject the null hypothesis if the p-value is less than the 5% level of significance. Otherwise, accept the null hypothesis.
Finally, the CPI and PCL time series variables are further examined to see if they are sensitive to each other or not using impulse response graph.

3. DATA ANALYSIS AND RESULTS

In this section, we present the results of the analyses carried out on the Crop Production Index (PCL) and Permanent Cropland (PCL) series using Eviews 9.0 and the results are presented as follow

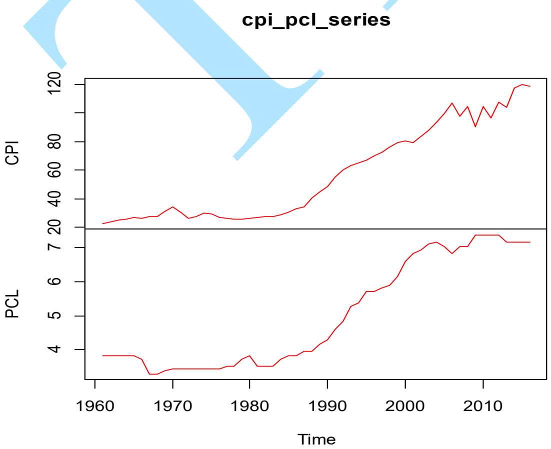

Figure 1: Time plot of Crop Production Index (CPI) series

Figure 1 shows that both CPI and PCL series are trending upward with fluctuations. This means that these two series will have to undergo difference stationary processes to achieve stationarity, see [3].

Table1: Results of Augmented Dickey-Fuller (ADF) Test for CPI and PCL Series

Series	ADF-Statistic	Critical Values	P-Val	Order of Integration
CPI	-3.597842	-2.917650	0.0090	I(1)
PCL	-5.258107	-2.916566	0.0001	I(1)

Table 1 reveals that both CPI and PCL are difference stationary series of order one {I(1)}. This is evident from their respective p-values which are less than the chosen level of significance (α= 0.05). Table 2 shows the result of Johansen cointegration test for the study.

Table 2: Results of Cointegration Tests

Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenval	Trace Statistic	0.05 Critical Value	Prob.**
None	0.104121	8.397061	15.49471	0.4240
At most 1	0.044529	2.459774	3.841466	0.1168
Unrestricted Cointegration Rank Test (Maximum Eigenvalue)				
Hypothesized No. of CE(s)	Eigenval	Max-Eigen Stat	0.05 Critical Value	Prob.**
None	0.104121	5.937286	14.26460	0.6213
At most 1	0.044529	2.459774	3.841466	0.1168

From Table 2, the hypotheses of no cointegration were accepted for both the trace and maximum eigenvalue tests since their respective p-values are greater than the 5% chosen level of significance. Since the CPI and PCL series are difference stationary series of order one {I(1)} and are not cointegrated, VAR (p) model is appropriate for analyzing the series (see [7]).
Then, we determined the number of lags to be included in the VAR model using the lag selection criteria as illustrated in Table 3.

Table 3: Lag Selection Criteria

Lag	LogL	LR	FPE	AIC	SC
0	-268.618	NA	139.3146	10.6125	10.6882
1	-129.179	262.4745	0.6877	5.3011	5.5284
2	-120.065	16.4412	0.5633	5.1006	5.4794
3	-108.094	**20.6553***	**0.4130***	**4.7880***	**5.3183***
4	-107.543	0.9079	0.4747	4.9232	5.6051
5	-104.322	5.0513	0.4925	4.9538	5.7872

* indicates lag order selected by the criterion

From Table 3, result of the selection criteria reveals that VAR of lag 3 is the appropriate model for analyzing the CPI and PCL series. Table 4 shows the VAR (3) model estimates.

Table 4: VAR (3) Model Estimates for CPI and PCL series

	Coefficient	Std. Error	t-Statistic	Prob
α_{11}	0.721018	0.125350	5.752022	**<0.0001**
β_{11}	7.053807	3.433719	2.054276	**0.0428**
α_{12}	0.573828	0.132990	4.314832	**<0.0001**
β_{12}	-6.249724	5.339483	-1.170474	0.2448
α_{13}	-0.567749	0.135328	-4.195367	**0.0001**
β_{13}	4.905261	3.549900	1.381803	0.1704
c_{11}	-12.01940	4.458345	-2.695933	**0.0083**
α_{21}	0.007414	0.005242	1.414355	0.1606
β_{21}	1.134206	0.143587	7.899104	**<0.0001**
α_{22}	0.002991	0.005561	0.537782	0.5920
β_{22}	-0.020627	0.223279	-0.092382	0.9266
α_{23}	-0.000565	0.005659	-0.099823	0.9207
β_{23}	-0.316237	0.148445	-2.130332	**0.0358**
c_{22}	0.477897	0.186433	2.563372	**0.0120**

From Table 4, the estimated VAR (3) model are stated as equations (11) and (12)

$$\mathrm{CPI}_t = 0.721018\mathrm{CPI}_{t-1} + 0.573828\mathrm{CPI}_{t-2} - 0.567749\mathrm{CPI}_{t-3} + 7.053807\mathrm{PCL}_{t-1} - 6.249724\mathrm{PCL}_{t-2} + 4.905261\mathrm{PCL}_{t-3} - 12.01940 \quad (11)$$

R^2= 0.989011 Durbin-Watson Stat. = 1.816895

$$\mathrm{PCL}_t = 0.007414\mathrm{CPI}_{t-1} + 0.002991\mathrm{CPI}_{t-2} - 0.000565\mathrm{CPI}_{t-3} + 1.134206\mathrm{PCL}_{t-1} - 0.020627\mathrm{PCL}_{t-2} - 0.316237\mathrm{PCL}_{t-3} + 0.477897 \quad (12)$$

R^2= 0.991803 Durbin-Watson Stat. = 1.974696

Table 5 shows outcome of the VAR Granger causality test for CPI and PCL.

Table 5: Results of VAR Granger Causality Tests for CPI and PCL Series

Dependent var: CPI				
Null hypothesis	**Chi-sq**	**df**	**Prob.**	**Direction of causality**
PCL does not Granger cause CPI	10.02854	3	0.0183	PCL→CPI
Dependent var: PCL				
Null hypothesis	**Chi-sq**	**df**	**Prob.**	**Direction of causality**
CPI does not Granger cause PCL	7.943276	3	0.0472	CPI→PCL

Note: *The p-values are statistically significant at 5% level of significance*

From Table 5, the results of Granger causality tests for CPI and PCL series showed a bidirectional relation that runs from the CPI to PCL and from the PCL to CPI. This means that CPI Granger causes PCL and PCL Granger causes the CPI. Sensitivity responsesof the two variablesusing impulse response graphs are presented in Figure 2.

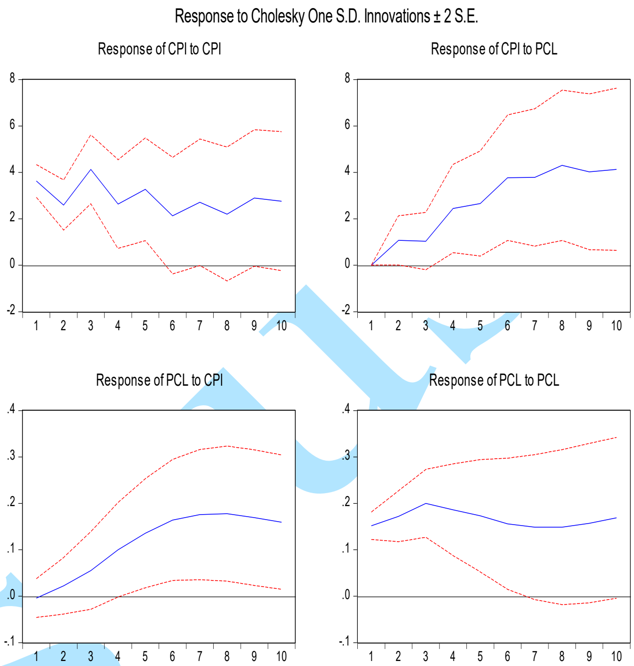

Figure 2: Impulse Response Graphs

From Figure 2, it can be deduced that crop production index (CPI) is sensitive to permanent cropland (PCL) evident from the graph labelled "response of CPI to PCL" at the top right of Figure 2. It can also be deduced that PCL is sensitive to CPI evident from the graph labelled "response of PCL to CPI" at the bottom left of Figure 2. This further buttressed the findings on Granger causality earlier discussed.

5. SUMMARY OF FINDINGS

This work aimed at studying the relationship between crop production index (CPI) and permanent cropland (PCL) in Nigeria using time series data from 1961 to 2016. Unrestricted Vector Autoregressive (VAR) model was used as the appropriate technique in modeling these two time series variables, because all the assumptions of VAR (p) model have been carefully considered and have been satisfied in this study. The order of integration of the individual series was determined using the Augmented Dickey-Fuller (ADF) test. Results reported by the time plot presented as Figure 1 and results of the ADF tests reported in Table 1 revealed that, these series become stationary after their first differences {i.e. I(1)}. By implication, the first assumption of VAR (p) model was satisfied. Moreover, results of Johansen cointegration tests presented in Table 2 showed that the hypotheses of no cointegration were accepted for both the trace and maximum eigenvalue tests; which implied that VAR (p) model was the appropriate model for the series. Then, the result of the selection criteria reported in Table 3 revealed that VAR of lag 3 {VAR (3)} is the appropriate model for analyzing the CPI and PCL series. Post-estimation examination of the fitted model stated as equations (11) and (12) showed that the estimated bi-variate VAR (3) model are non-spurious. This is evident from the values of the Durbin-Watson statistic which are greater than the values of the coefficient of determination (R^2) in both equations. Hence, results of the Granger causality tests reported in Table 5 revealed that a bidirectional relation runs from crop production index (CPI) to permanent cropland (PCL) and from permanent cropland (PCL) to crop production index (CPI). Finally, the impacts of stochastic innovations to CPI on PCL and to PCL on CPI were explored using the impulse response function graph. It is evident from this graph that CPI is sensitive to PCL and PCL is sensitive to CPI based on the data used for this work.

In practice, the statistical findings of the study offer a strong standpoint to unequivocally submit that PCL and CPI are both mutually predictable. This implies that availability of adequate PCL in a country can guarantee a favorable CPI. Although sufficient PCL may also create a decline in biodiversity tendencies, as creation of artificial ecosystems may be greatly hampered due to agricultural intensification. However, adequate PCL or CPI offers a compensatory balance in terms of improved and sustainable agroecosystems, food availability, raw materials production, employment creation and self-empowerment, thereby projecting such a nation's gross domestic product (GDP) to a reasonable height. Biodiversity degradation may be addressed when modern agricultural practices and technologies are embraced. Such practices aimed at reducing land use intensification solely due to crop cultivation may include aquaponics, hydroponics, precision agriculture, vertical farming, greenhouse farming and application of artificial intelligence in agriculture e.g. industry tools 4.0.

The interdependence of PCL on CPI and vice versa can also play vital roles in forecasting studies when there is sufficient data on either of the two variables. With the findings in this study, agricultural policies can be formed and implemented easily with lesser uncertainties.

6. CONCLUSION AND RECOMMENDATION

In this work, we reported the causal relationships between Crop Production Index (CPI) and Permanent Cropland (PCL) in Nigeria from 1961 to 2016 using the unrestricted Vector Autoregressive (VAR) model. Based on the outcomes of the analysis and summary of findings, it can be concluded that a bidirectional relationship exists between the CPI and PCL running from CPI to PCL and PCL to CPI. In other words, Nigeria's crop production index is predictable by Nigeria's permanent cropland and Nigeria's permanent cropland is also predictable by Nigeria's crop production index.The findings of this work can be used to further improve both crop production and permanent cropland in Nigeria through appropriate policy formulation.

REFERENCES

[1] **S. B. Akanni, O.I. Adeniyi** - On the forecast of cereals production in Nigeria. *Federal University Wukari Trends in Science and Technology Journal (FTSTJ)*. Vol. 5 No. 2, pp. 331-335, 2020.

[2] **S.B. Akanni, M. K. Garba, A. W. Banjoko and R. B. Afolayan** -"Econometric analysis of the effects of land size on cereals production in Nigeria." *Islamic University Multidisciplinary Journal*, Vol.7, No. 1, 252-258, 2020.

[3] **G.E.P. Box, G.M. Jenkins** - "Time series analysis: Forecasting and control San Francisco." Calif: Holden-Day, 1976.

[4] **P. Ciaian, F. Guri., M. Rajcaniova, D. Drabik and S.G., y Paloma** - Land fragmentation and production diversification: A case study from rural Albania. *Land use policy*, *76*, 589-599, 2018

[5] **J.H. Cochrane**. Time series for macroeconomics and finance. *Manuscript, University of Chicago*, 1-136, 2005

[6] **D.A. Dickey, WA. Fuller** - Distribution of the Estimators for AutoregressiveTime Series with a Unit Root. *Journal of the American Statistical Association*, 74, pp 427–431, 1979

[7] **R.F. Engle, C.W.J Granger** - Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica* 55:251-76, 1987

[8] **E. T., Epule, C. R., Bryant, C. Akkari, M. A., Sarr and C., Peng**. An assessment of the predictors of the dynamics in arable production per capita index, arable production and permanent cropland and forest area based on structural equation models. *Springer Plus*, *3*(1), 597, 2014.

[9] **M.K. Garba, S.B. Akanni, W.B. Yahya, K.Y. Kareem, R.B. Afolayan** - Modelling Effects of some Factors that Contribute to Cereals Yields in Nigeria using Toda-Yamamoto Techniques. Sule Lamido University Journal of Science and Technology (SLUJST) Vol. 1 No., pp. 50-56, 2020.

[10] **W. R. Goodman, S. A. Coady and B.C. English** - Biomass crop production on representative southeastern US farms: Farm profitability, erosion control and the 1985 Food Security Act. *Bioresource Technology*, *36*(3), 207-214, 1991.

[11] **W.H. Greene** - *Econometric analysis*. Pearson Education India, 2003.

[12] **D.N. Gujarati, D.C. Porter** - Basic econometrics, 5th edition. McGraw-Hill/Irwin, New York, 2009.

[13] **S. Johansen** - Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," Econometrica, Vol. 59, No. 6, pp. 1551–1580, 1991.

[14] **H. Lu, H. Xie, Y. He, Z. Wu and X. Zhang** - Assessing the impacts of land fragmentation and plot size on yields and costs: A translog production model and cost function approach. *Agricultural Systems*, *161*, 81-88, 2018.

[15] **S. Oyekale** - Determinants of agricultural land expansion in Nigeria: Application of Error Correction Modeling (ECM). *Journal of Central European Agriculture*, 2007.

[16] **M.H. Pesaran, Y. Shin and R.J. Smith** - Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, *16*(3), 289-326, 2001.

[17] **K. Radoslava** - Agriculture and Its Impact on Land-Use, Environment, and Ecosystem Services, Landscape Ecology - The Influences of Land Use and Anthropogenic Impacts of Landscape Creation, *Amjad Almusaed, IntechOpen*, doi: 10.5772/63719, 2016.

[18] **C.A. Sims** - Macroeconomics and Reality. Econometrica. 48, pp.1-48, 1980.

[19] **V. Smil** - Nitrogen in crop production: An account of global flows. *Global biogeochemical cycles*, *13*(2), 647-662, 1999

[20] **G. Tóth** - Impact of land-take on the land resource base for crop production in the European Union. *Science of the Total Environment*, *435*, 202-214, 2012.

[21] **World Bank Repository** - Data on Food Production Index. Retrieved from http:// data.worldbank.org, 2018