



Modelling Effects of some Factors that Contribute to Cereals Yields in Nigeria using Toda-Yamamoto Techniques

Garba, M. K.^{1*}, Akanni, S. B.², Yahya, W. B.³, Kareem, K. Y.⁴ and Afolayan, R. B.⁵

1. Department of Statistics, University of Ilorin, Ilorin, Nigeria, garba.mk@unilorin.edu.ng
2. Department of Physical Sciences (Statistics Option), Al-Hikmah University, Ilorin, Nigeria, akannisb2018@gmail.com
3. Department of Statistics, University of Ilorin, Ilorin, Nigeria, wbyahya@unilorin.edu.ng
4. Department of Agric & Biosystems Engineering, University of Ilorin, Ilorin, Nigeria, kareemkola99@gmail.com
5. Department of Statistics, University of Ilorin, Ilorin, Nigeria, afolayan.rb@unilorin.edu.ng

Abstract

This study aimed to examine the direction of causality among Cereals Production (CP), Land used for Cereals Production (LP) and Cereal Yields (CY) in Nigeria for a period of 50 years (1966 to 2016) using techniques of Toda-Yamamoto. The maximum order of integration and optimal lag order of the series confirmed that VAR(2+1) model best fitted the data. Results from the estimated model revealed that two year past values $t-2$ (2014 and 2015) of CP is the major determinant of CP series in current time period t (2016) while one year past value $t-1$ (2015) of CY and two year past values $t-2$ (2014 and 2015) of CP and LP are major determinants of CY series in current time period t (2016). The results of Toda-Yamamoto causality examination showed that CY is Granger-caused by both CP and LP. Based on the sequence of analyses carried out in this study, it was concluded that cereals yields in Nigeria can be predicted by both cereal production and the size of farmland used for planting cereal crops. The study then recommended that adequate plots of land be allocated to farmers interested in cereals production in order to improve yields of cereals and ensure food security in the country.

Keywords: Cereals Production; Order of Integration; Optimal Lag Order; Toda-Yamamoto Causality; Nigeria.

1. Introduction

The concern for sufficient cereals production to meet national and international demand and supply is a global issue that requires urgent solutions. Factors that may be considered for effective cereals production are cereals yields, land used for cereals production, rainfall/irrigation and effective soil management such as fertilizer application, crop rotation practices etc. In the past five decades, the global crop production originated mostly from increased yields and higher cropping intensity (FAO, 2020). Cereals yields, measured as kilograms per hectare (kg/ha) of harvested land, includes wheat, rice, maize, barley, oats, rye, millet, sorghum, buckwheat, and mixed grains. In other words, cereals yield is the weight of cereals production per land area. Therefore, improving cereals yield is crucial to increasing cereals production. According to (McArthura and McCordb, 2017), the key inputs to yield growth are fertilizer, modern variety seeds and water. Cereals Production, in this context, takes an economic approach to estimate amount of cereals produced in metric tonnes. Land under cereals production measured in hectares (ha), is the expanse of land resources used for the production. According to (World Bank Repository, 2018), cereals production has increased by 76%, cereal yield by 58% and cultivated land by 42% in Nigeria the last 50 years. Despite the great potential of Nigeria in cereal production, the frequent occurrence of drought occasioned by erratic rainfall distribution and/or discontinuation of rain during the growing season is the greatest hindrance to increased production and this is more serious in the Northern part of country where most of the cereals are produced (Olaoye, 1999).

When there are more than one time series variables to be analyzed in a study, the best modeling technique to be adopted for such study is the multivariate time series techniques whereby holistic consideration would be



given to the variables. However, the use of a particular multivariate time series model requires that assumptions of such model are satisfied and fulfilled before its usage. For instance, when time series under consideration are differenced once before it became stationary, the series is said to be of order one (i.e. $I(1)$) and the system satisfies the conditions of cointegration relations, then the system will be appropriately specified as a Vector Error Correction Model (VECM) otherwise the Vector Autoregressive (VAR) system developed by Sims in 1980 (see Engle and Granger 1987) would be used. The main goal of forecasting time series variable(s) is to improve predictiveness which consequently increases its realizations. As a result, early crop yield forecasting is important for food safety as well as large-scale food related planning (Wang *et al.* 2014). Moreover, the rate of growth of Nigeria's food production is 2.5% per annum in recent years, while food demand has been growing at the rate of more than 3.5% per annum due to high rate of population growth of 2.83% (Kolawole and Ojo, 2007).

Several studies have been conducted on cereal crops in the recent times. For instance, Akanni and Adeniyi (2020) forecasted the cereals production of Nigeria using ARIMA(1, 1, 1) model. Results from their study showed that cereals production in Nigeria will continue to increase for the foreseeable future. To access the impacts of cultivated land on cereals production in Nigeria, Akanni *et al.* (2020) employed Vector Autoregressive model of order two {VAR(2)} to show that production of cereal crops in Nigeria is granger-caused (predictable) by the size of farmland used for planting cereal crops. In their study, Fischer and Edmeades (2010) showed that breeding can also help farmers to achieve increased cereal yields when they investigated the breeding and cereal yields progress for several key production regions, namely the United Kingdom and the Yaqui Valley of Mexico for wheat, Japan and Central Luzon in the Philippines for rice, and Iowa and briefly sub-Saharan Africa for maize. Farook and Kannan (2015) adopted the techniques of VAR model to show that climate change variables like maximum and minimum temperature have significant effects on Kharif rice yields while only rainfall has negative impact on the yields of the produce. Furthermore, Badmus and Ariyo (2011) forecasted cultivated areas and maize production in Nigeria using univariate time series approach known as ARIMA model. Maikasawa and Ala (2013) applied trend analysis to study the area and productivity of sorghum in Sokoto State, Nigeria. The results of the trend analysis later revealed a stagnated growth in area and an accelerated growth in productivity. Shabu and Terfa (2011) applied Cobb-Douglas production function and technical efficiency techniques to assess the efficacy of resources used in rice farming enterprise in Kwande Local Government Area of Benue State, Nigeria. This study, therefore, adopted the techniques developed by Toda and Yamamoto in 1995 to model annual time series dataset on cereals production (CP), cereals yields (CY) and land under cereal production (LP) in Nigeria in order to examine the causality relationships that exist among them. The Toda-Yamamoto methodology reportedly has proven superiority over ARIMA and Cobb-Douglas approaches employed by previous studies.

2. Materials and Methods

For this study, an annual time series dataset on cereals production (CP), cereals yields (CY) and land under cereal production (LP) in Nigeria covering 1966 to 2016 was extracted from the repository of World Bank via their website <http://data.worldbank.org>. This study utilized the techniques of Toda-Yamamoto (TY) to analyze the causal effects of cereals production and land size on cereals yields in Nigeria. The approach was developed by Toda and Yamamoto (1995) to overcome the shortcomings of the conventional granger causality methodology developed by Granger (1969). Unlike the conventional granger causality practice which uses an unrestricted VAR model of order p to analyze the dynamic relationships among time series variables, the TY procedure uses an augmented VAR model of order $p+d_{max}$. In order to achieve the goal of the study, the following three basic steps of Toda-Yamamoto procedure were followed:

2.1 Determination of the maximum order of integration

The first step involves the testing of the CP, CY and LP time series to determine the maximum order of integration (d_{max}) of the variables in the system using the Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests. For the ADF, the null hypothesis is that the series are not stationary while for



the KPSS, the null hypothesis is that the series are stationary. For the procedures for conducting ADF and KPSS, see Dickey and Fuller (1979) and Kwiatkowski *et al*, (1992).

2.2 Determination of the optimal lag length (p)

Having determined the order of integration of the series, then determination of the optimal lag length (p) for the series followed using Akaike’s Information Criterion (AIC), Schwarz Information Criterion (SC), Hannan-Quinn (HQ) Information Criterion, Final Prediction Error (FPE) and the Likelihood Ratio (LR) statistics as selection criteria. These selection criteria are used to select the best model out of many possible fitted models. See Akaike (1987), Glover *et al*. (2004) and Gujarati and Porter (2009) for more details. Interestingly, most statistical and econometric software packages have these criteria built into them. Although, not all of these packages report the selection criteria automatically, some of them require the analysts to request for them to be reported in the result output by either selection appropriate checkbox or include the needed code before running the programme.

2.3 Testing for causality

This was done by using the Modified Wald’s statistical procedure to determine the VAR (k). The optimal lag length is equal to $k = (p+d_{max})$. The Modified Wald’s test has an asymptotic chi-square distribution with p degrees of freedom in the limit when a VAR($p+d_{max}$) is estimated.

In order to test for TY causality among the CP, CY and LP series, the following trivariate VAR(k) model is formulated as follows:

$$CP_t = \alpha_{10} + \sum_{i=1}^k \phi_{1i} CP_{t-i} + \sum_{i=1}^k \beta_{1i} CY_{t-i} + \sum_{i=1}^k \delta_{1i} LP_{t-i} + e_{1t} \quad (1)$$

$$CY_t = \alpha_{20} + \sum_{i=1}^k \phi_{2i} CP_{t-i} + \sum_{i=1}^k \beta_{2i} CY_{t-i} + \sum_{i=1}^k \delta_{2i} LP_{t-i} + e_{2t} \quad (2)$$

$$LP_t = \alpha_{30} + \sum_{i=1}^k \phi_{3i} CP_{t-i} + \sum_{i=1}^k \beta_{3i} CY_{t-i} + \sum_{i=1}^k \delta_{3i} LP_{t-i} + e_{3t} \quad (3)$$

Where: CP_t denotes Cereals Production at time t , CY_t denotes Cereal Yield at time t , LP_t denotes Land under Cereal Production at time t , e_{1t} , e_{2t} and e_{3t} are uncorrelated error terms in the equations.

Consequently, the following hypotheses were formulated with respect to equations (1) to (3):

Hypotheses for equation (1)

$$H_{01}: \beta_{11} = \beta_{12} = \beta_{13} = \dots = \beta_{1k} = 0 \text{ (CY does not Granger cause CP)}$$

$$H_{02}: \delta_{11} = \delta_{12} = \delta_{13} = \dots = \delta_{1k} = 0 \text{ (LP does not Granger cause CP)}$$

Hypotheses for equation (2)

$$H_{03}: \phi_{21} = \phi_{22} = \phi_{23} = \dots = \phi_{2k} = 0 \text{ (CP does not Granger cause CY)}$$

$$H_{04}: \delta_{21} = \delta_{22} = \delta_{23} = \dots = \delta_{2k} = 0 \text{ (LP does not Granger cause CY)}$$

Hypotheses for equation (3)

$$H_{05}: \phi_{31} = \phi_{32} = \phi_{33} = \dots = \phi_{3k} = 0 \text{ (CP does not Granger cause LP)}$$

$$H_{06}: \beta_{31} = \beta_{32} = \beta_{33} = \dots = \beta_{3k} = 0 \text{ (CY does not Granger cause LP)}$$

The decision rule is to reject the null hypothesis if the p-value is less than the 5% significance level. Otherwise, the null hypothesis will not be rejected.

2.4 Model Specification for CP, CY and LP Series

The trivariate Augmented Vector Autoregressive model of order two plus one {VAR (2+1)} for cereals production (CP), cereals yield (CY) and land under cereal production (LP) series is given by:

$$CP_t = \phi_{11}CP_{t-1} + \phi_{12}CP_{t-2} + \beta_{11}CY_{t-1} + \beta_{12}CY_{t-2} + \delta_{11}LP_{t-1} + \delta_{12}LP_{t-2} + \alpha_{10} + \phi_{13}CP_{t-3} + \beta_{13}CY_{t-3} + \delta_{13}LP_{t-3} \tag{4}$$

$$CY_t = \phi_{21}CP_{t-1} + \phi_{22}CP_{t-2} + \beta_{21}CY_{t-1} + \beta_{22}CY_{t-2} + \delta_{21}LP_{t-1} + \delta_{22}LP_{t-2} + \alpha_{20} + \phi_{23}CP_{t-3} + \beta_{23}CY_{t-3} + \delta_{23}LP_{t-3} \tag{5}$$

$$LP_t = \phi_{31}CP_{t-1} + \phi_{32}CP_{t-2} + \beta_{31}CY_{t-1} + \beta_{32}CY_{t-2} + \delta_{31}LP_{t-1} + \delta_{32}LP_{t-2} + \alpha_{30} + \phi_{33}CP_{t-3} + \beta_{33}CY_{t-3} + \delta_{33}LP_{t-3} \tag{6}$$

3. Results

The analyses of the time series data for CP, CY and LP were carried out using Eviews 9.0 and the results are presented as follows. Figure 1 shows that CP, CY and LP series exhibit upward trends albeit with fluctuations. This implies that the three series are not stationary at levels. Differencing these series once or twice will make them stationary as suggested by Box and Jenkins (1976). Furthermore, unit root tests such as ADF and KPSS were used to confirm the revelations by the time plots and the results are presented in Table 1.

To properly determine the order of integration of the series, each of the series was subjected to unit root analysis using both Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests. Results from the ADF test in Table 1 showed that both CP and LP series are differenced stationary series of order one I(1) while CY series is differenced stationary series of order two I(2). Also from the same Table 1, the KPSS test showed that CP, CY and LP series are all differenced stationary series of order one I(1). Based on the superiority of KPSS test to ADF test, the maximum order of integration in the system was taken to be one. This implied that the first condition of TY causality test was satisfied. Then the VAR optimal lag selection criteria were carried out from lag 1 to lag 4 for the CP, CY and LP series. Results of the optimal lag length in Table 2 showed that the least value of the selection criteria LR, FPE, AIC and HQ occurred at lag 2 indicating that VAR(2) was the appropriate model for analyzing the CP, CY and LP series.



Figure 1: Time plots of Cereals Production (CP), Cereals Yields (CY) and Land under Cereals Production (LP)

Table 1: Results of ADF and KPSS Tests for CP, CY and LP Series

ADF				
Variable	ADF-statistic	Critical values	P-val	Order of integration
CP	-4.186495	-2.923780	0.0018	I(1)
CY	-7.218603	-2.931404	<0.0001	I(2)
LP	-8.151617	-2.922449	<0.0001	I(1)
KPSS				
Variable	KPSS-statistic	Critical values	Max. order of integration (dmax)	
CP	0.098989	0.347000	I(1)	
CY	0.182924	0.347000	I(1)	
LP	0.099023	0.347000	I(1)	

Table 2: VAR Lag Order Selection Criteria for CP, CY and LP Series

Lag	LR	FPE	AIC	SC	HQ
0	NA	1.24e+30	77.80485	77.92294	77.84929
1	174.8744	3.11e+28	74.12098	74.59336*	74.29874
2	23.30203*	2.57e+28*	73.92141*	74.74807	74.23249*
3	12.46763	2.72e+28	73.96743	75.14837	74.41182
4	11.15824	2.95e+28	74.02222	75.55745	74.59994

Table 3: Estimates of Augmented VAR (2+1) Model

	Coef.	Std. Error	t-Statistic	P-val
ϕ_{11}	0.157880	0.348637	0.452850	0.6515
ϕ_{12}	1.041430	0.374932	2.777649	0.0064
β_{11}	4603.711	4305.602	1.069237	0.2872
β_{12}	-5571.406	5363.556	-1.038752	0.3011
δ_{11}	0.873829	0.434037	2.013259	0.0464
δ_{12}	-0.860752	0.517349	-1.663775	0.0989
α_{10}	-5021145.	5475161.	-0.917077	0.3610
ϕ_{13}	-0.615900	0.371565	-1.657582	0.1002
β_{13}	6130.590	4540.887	1.350086	0.1797
δ_{13}	0.439547	0.513024	0.856776	0.3934
ϕ_{21}	-2.49E-05	2.24E-05	-1.110002	0.2693
ϕ_{22}	6.49E-05	2.41E-05	2.696431	0.0081
β_{21}	1.040467	0.276589	3.761787	0.0003
β_{22}	-0.528883	0.344551	-1.534995	0.1276
δ_{21}	5.54E-05	2.79E-05	1.986990	0.0493
δ_{22}	-8.72E-05	3.32E-05	-2.624427	0.0099
α_{20}	1069.105	351.7202	3.039646	0.0029
ϕ_{23}	2.07E-05	2.39E-05	0.867431	0.3875
β_{23}	-0.375941	0.291703	-1.288780	0.2001
δ_{23}	-4.40E-05	3.30E-05	-1.335068	0.1845
ϕ_{31}	-0.058785	0.289569	-0.203010	0.8395
ϕ_{32}	0.056140	0.311409	0.180278	0.8573
β_{31}	-936.6434	3576.126	-0.261916	0.7939
β_{32}	2097.439	4454.836	0.470823	0.6387
δ_{31}	0.728688	0.360500	2.021323	0.0456
δ_{32}	0.290919	0.429697	0.677032	0.4998
α_{30}	-7019479.	4547532.	-1.543580	0.1255
ϕ_{33}	-0.451358	0.308613	-1.462537	0.1463
β_{33}	5706.431	3771.547	1.513021	0.1330
δ_{33}	0.461958	0.426105	1.084140	0.2806

Table 4: Results of Toda-Yamamoto Causality Tests for CP, CY and LP series

Null hypothesis				
Dependent var: CP	Chi-sq	df	P-val	Direction of causality
CY does not Granger cause CP	1.406908	2	0.4949	No causation
LP does not Granger cause CP	4.790805	2	0.0911	No causation
Dependent var: CY	Chi-sq	df	P-val	Direction of causality
CP Granger causes CY	7.273464	2	0.0263	CP → CY
LP Granger causes CY	7.727701	2	0.0210	LP → CY
Dependent var: LP	Chi-sq	df	P-val	Direction of causality
CP does not Granger cause LP	0.053125	2	0.9738	No causation
CY does not Granger cause LP	0.221870	2	0.8950	No causation

The results presented in Table 3 which is the estimated VAR(2+1) model are stated as equations (7), (8) and (9)

$$CP_t = 0.157880CP_{t-1} + 1.041430CP_{t-2} + 4603.711CY_{t-1} - 5571.406CY_{t-2} + 0.873829LP_{t-1} - 0.860752LP_{t-2} - 5021145 - 0.615900CP_{t-3} + 6130.590CY_{t-3} + 0.439547LP_{t-3} \quad (7)$$

$$CY_t = -2.49E - 05CP_{t-1} + 6.49E - 05CP_{t-2} + 1.040467CY_{t-1} - 0.528883CY_{t-2} + 5.54E - 05LP_{t-1} - 8.72E - 05LP_{t-2} + 1069.105 + 2.07E - 05CP_{t-3} - 0.375941CY_{t-3} + -4.40E - 05LP_{t-3} \quad (8)$$

$$LP_t = -0.058785CP_{t-1} + 0.056140CP_{t-2} - 936.6434CY_{t-1} + 2097.439CY_{t-2} + 0.728688LP_{t-1} + 0.290919LP_{t-2} - 7019479 - 0.451358CP_{t-3} + 5706.431CY_{t-3} + 0.461958LP_{t-3} \quad (9)$$

The Table 4 has the summary of the results of Toda-Yamamoto causality tests for the CP, CY and LP series. The Table 4 has the summary of the results of Toda-Yamamoto causality tests for the CP, CY and LP series. It can be deduced that neither CY nor LP granger causes CP. Similarly, neither CP nor CY granger causes LP. However, each of CP and LP granger causes CY. The implications of these results are further stated in the subsequent section.

4. Discussion

In this work, we investigated the effects of cereal production (CP) and land under cereal production (LP) on cereal yield (CY) in Nigeria using time series data from 1966 to 2016. The Toda-Yamamoto (TY) causality techniques were used to analyze the time series variables. The three basic steps of TY causality test were carefully considered and were fulfilled accordingly in this work.

Based on the results presented in Table 3, only the lag of CP series in (t-2) period has a significant effects on the CP series in current time period t for equation (1) while for equation (2), CY series in (t-1) period, CP and LP series in (t-2) periods have significant effects on CY series in current time period t. However, none of the lagged values of CP, CY and LP series have significant effects on LP series in current time period t.

Finally, results of the Toda-Yamamoto causality test in Table 4 showed that Cereals Yields (CY) is not only granger-caused by Cereals Production (CP) but also by Land under Cereals Production (LP). However, CY does not granger cause CP and LP does not granger cause CP. Also, CP does not granger cause LP and CY does not granger cause LP.

5. Conclusion and Recommendation

As a result of the sequence of statistical analyses carried out on the data on cereals production (CP), cereal yield (CY) and land under cereals production (LP) in Nigeria from 1966 to 2016, it can be concluded that a



unidirectional causality relationship exists among the three variables running from CP to CY and LP to CY. This implies that cereals yields in Nigeria can be predicted by both cereal production and the size of farmland used for planting cereal crops.

Therefore, the study suggests that adequate plots of land should be allocated to state farmers solely interested in cereals production in order to improve yields of cereals and ensure food security in the country.

References

- Akanni, S. B. and Adeniyi, O. I (2020). On the forecast of cereals production in Nigeria. *Federal University Wukari Trends in Science and Technology Journal, FSTJ-0216*
- Akanni, S. B., Garba, M. K., Banjoko, A.W. and Afolayan, R.B. (2020). Econometric analysis of the effects of land size on cereals production in Nigeria. *Islamic University Multidisciplinary Journal, vol. 7 (1), pp.252-258.*
- Badmus, M. A. and Ariyo, O. S. (2011). Forecasting cultivated areas and production of maize in Nigeria using ARIMA model. *Asian Journal of Agricultural Sciences 3(3): pp.171-176*
- Box, G. E. and Jenkins, G. M. (1976). Time series analysis, control, and forecasting. *San Francisco, CA: Holden Day, 3226(3228), 10.*
- Dickey, D. A. and Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association, 74, pp 427–431.*
- Engle, R. F. and Granger, C. W. J. (1987). Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica 55:251-76.*
- Food and Agriculture Organization (2020). *Fact files about World Crop Production*. Retrieved from www.fao.org
- Farook, A. J and Kannan, K. S. (2015). Climate Change Impact on Rice Yield in India – Vector Autoregression Approach. *Sri Lankan Journal of Applied Statistics, Vol (16-3), pp. 161-178*
- Fischer, R. A. and Edmeades, G. O. (2010). Breeding and Cereal Yield Progress. *Crop Science, Vol. 50: S-85–S-98. Retrieved from www.crops.org*
- Glover, S., Rosenbaum, D. A., Graham, J., and Dixon, P. (2004). Grasping the meaning of words. *Experimental Brain Research, 154(1), 103-108.*
- Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods *Econometrica, Vol. 37, 424-438*
- Gujarati, D. N. and Porter, D. C. (2009). *Basic econometrics, 5th edition*. McGraw-Hill/Irwin, New York.
- Kolawale, O. and Ojo, S. O. (2007). Economic efficiency of small scale food crop production in Nigeria. *Journal For Social Sciences, 14(2): 123-130.*
- Kwiatkowski, D. P., Phillips, C. B., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationary against the alternative of a unit root,” *Journal of Econometrics, 54, 159-178.*
- Maikasuwa, M. A. and Ala, A. L. (2013). Trend analysis of area and productivity of sorghum in Sokoto State, Nigeria, 1993-2012. *European Scientific Journal. Vol.9, No.16. pp 69-75*
- McArthur, J. W. and McCord, G. C. (2017). Fertilizing growth: Agricultural inputs and their effects in economic development. *Journal of Development Economics 127 (2017), pp. 133–152*
- Olaoye, O. (1999). Developing drought tolerant varieties for the savannah agro-ecologies of Nigeria in 25th year commemorative publication of Genetic Society of Nigeria, pp. 173-182.
- Shabu, T. and Terfa, D. (2011). Efficiency of resource use in rice farming enterprise in Kwande Local Government Area of Benue State, Nigeria. *International Journal of Humanities and Social Science Vol. 1 No. 3, pp. 215-220*
- Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica. 48, pp.1-48.*
- Toda, H. Y. and Yamamoto, T. (1995). Statistical inference in Vector Autoregressions with possibly integrated process. *Journal of Econometrics 66:225–250.*
- Wang, M., Tao, F. and Shi, W. (2014). Corn yield forecasting in northeast china using remotely sensed spectral indices and crop phenology metrics. *Journal of Integrative Agriculture, Vol. 3(7): pp. 1538-1545*
- World Bank Repository (2018). *Data on cereal production and land under cereal production*. Retrieved from <http://data.worldbank.org>