

METHODOLOGY

Open Access

# 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

Epule Terence Epule<sup>1\*</sup>, Christopher Robin Bryant<sup>1\*</sup>, Cherine Akkari<sup>1</sup>, Mamadou Adama Sarr<sup>1</sup> and Changhui Peng<sup>2</sup>

## Abstract

This study sets out to verify the key predictors of the dynamics of the arable production per capita index, the arable production and permanent crop land and forest area at a national scale in Cameroon. To achieve this objective, data for twelve time series data variables spanning the period 1961–2000 were collected from Oxford University, the United Nations Development program, the World Bank, FAOSTAT and the World Resource Institute. The data were analysed using structural equation models (SEM) based on the two stage least square approach (2SLS). To optimize the results, variables that showed high correlations were dropped because they will not add any new information into the models. The results show that the arable production per capita index is impacted more by population while the influence of rainfall on the arable production per capita index is weak. Arable production and permanent cropland on its part has as the main predictor arable production per capita index. Forest area is seen to be more vulnerable to trade in forest products and logging than any other variable. The models presented in this study are quite reliable because the p and t values are consistent and overall, these results are consistent with previous studies.

**Keywords:** Arable production; Forest area; Arable production and permanent cropland; Structural equation modelling; Predictors

## Introduction

Assessing the feedbacks and predictors of the arable production per capita index, forest area and arable production and permanent cropland at a national scale in Cameroon is of great pertinence for several reasons. Agriculture contributes about 50% to the GDP in most African countries (FAO and UNIDO 2008). Primarily, in Cameroon, the agricultural sector provides between 55–60% of employment (INS 2009). By 2050, global agricultural production ought to increase by 70–100% if the increasing world population is to be adequately fed (Dubois 2011). However, most developing countries will witness a decline in

agricultural production from about 3% to 1.2% over the period 2006 to 2050 (Bruinsma 2009). Secondly, many of the attempts to increase agricultural production in most developing countries in general and in Cameroon in particular have been based on the expansion of farmland in efforts to increase yields. Rosegrant and Cline (2003) and Rosegrant and Svendsen (1993) have argued that the expansion of farmland or cultivated area is a common method of increasing agricultural production to meet the rising food demand in most developing countries associated with low levels of intensification of agriculture. In most parts of the developing world, such land expansion is often at the expense of large areas of forest. A study by Epule et al. (2011) empirically justifies the conclusion that attempts at increasing arable farming yields in Cameroon have led to arable and permanent cropland being the second most vital cause of

\* Correspondence:

terence.epule.epule@umontreal.ca; chris.r.bryant@umontreal.ca

<sup>1</sup>Département de Géographie, Université de Montréal, Pavillon 520, ch. de la Côte-Sainte-Catherine, Local 332-3, CP 6128, succursale Centre-ville, Montréal, Québec H3C 3J7, Canada

Full list of author information is available at the end of the article

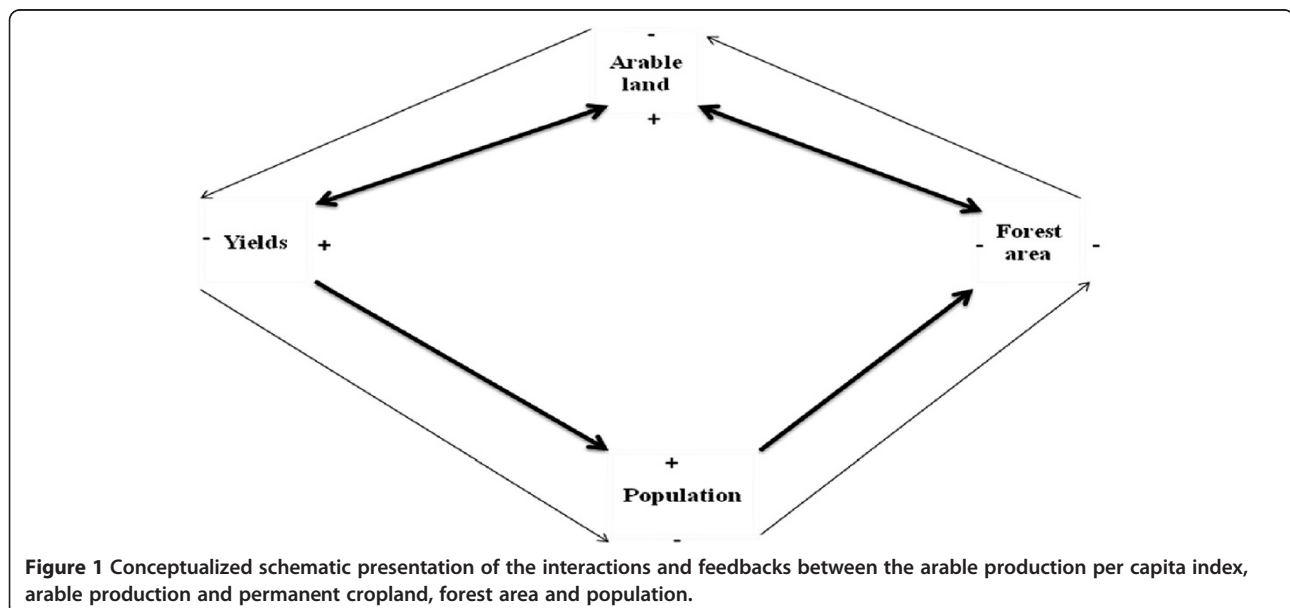
deforestation in Cameroon. Zhao et al. (2006) made similar observations for parts of Asia.

It can therefore be seen that the three variables being verified in this study are intricately related. In spite of this, there are currently no studies on Cameroon that have attempted to verify the key predictors of these variables using structural equation models (SEM). However, the state of research in this area is as follows: Yengoh and Ardo (2014) verified crop yield gaps in Cameroon using biophysical suitability modelling on specific crops; Garcia-Ponce et al. (2012) attributed variations in agricultural productivity in Senegal to government policies; Amujoyegbe et al. (2010) and Kombiok et al. (2013) attribute crop yield declines in Cameroon and Ghana to declining soil fertility; Epule et al. (2011, 2012) used an empirically grounded regression model to assert that population growth is at the centre of forest area decline in Cameroon. While the latter study was able to determine causality, it is weakened because the multiple linear regression approach used only establishes the link between a single dependent variable and several independent variables; in this case, the feedbacks between the variables which mimic real life situations are absent (Petraitis et al. 1996). It is for this reason that this study employs the use of SEM. SEM determines the key predictors among a group of variables and is capable of modelling the feedback between several endogenous and exogenous variables, thus mimicking real life situations or representing the real world more adequately (Petraitis et al. 1996). To the best of our knowledge, this study is the first to adopt this approach in determining the predictors of the endogenous variables under consideration in Cameroon. Thus, this study will verify the predictors of the arable production per capita index, arable production

and permanent cropland and forest area at a national scale in Cameroon. The approach of verifying the predictors of several endogenous variables using SEM is justified by the fact that it enhances prediction by introducing multi-way interactions among the endogenous and exogenous variables.

*Arable production per capita index in international \$* is used in this study in reference to the amount of food produced in relation to the population (i.e. per head) (FAO 2006). *Arable production and permanent cropland in hectares* is used in reference to the amount of food produced in relation to ecumene land that is viable for such production; in other words, it refers to food production based on the amount of arable land under cultivation (FAO 2006). *Forest area in hectares* refers to the amount of land covered by forest at a given point in time. In this study, *forest area* in hectares is assumed to decline at a rate of about 220 Kha/year (FAO 2006, 2010). The latter is the FAO's estimate for the decade 1990–2000.

Conceptually, it can be observed that when population growth increases the need for more food production becomes apparent. In most developing countries including Cameroon, arable farmers have relied for several centuries on farm land expansion in attempts to increase yields; unfortunately, this has been at the expense of large expanses of forest. The short term repercussions of such expansion are increased yields but in the long run, such yields are often not sufficient to meet population food needs (Figure 1). When arable land is increased, it provides a means for more production in the short term, but in the long term, expansion of arable land alone is not often sufficient to increase yields at proportions that should meet demand. When all the available land



**Figure 1** Conceptualized schematic presentation of the interactions and feedbacks between the arable production per capita index, arable production and permanent cropland, forest area and population.

has been exploited, it becomes difficult to increase production further if agricultural intensification or agroecology methods cannot be used. On Figure 1 the inner thicker arrows represent the short term scenario with yield increase while the outer thinner arrows represent the declines in yields because of the inability to expand land further. The + and – signs show the nature of the effects on the variable.

## Study area and methods

### Study area

Cameroon is located in central Africa. More precisely, the country is located between latitude 2° N and 13° N of the equator. Longitudinally, Cameroon is located between longitude 8° E and 15° E of the prime meridian (Molua 2006). Cameroon has a total surface area of 475,400 km<sup>2</sup> and a population of over 20.3 million people. Agriculture employs about 70–80% of the country's population (Mundex Dataset 2012; Central Intelligence Agency 2012; Carr et al. 2005). Climatically, Cameroon has an equatorial climate in the south with rainfall levels of between 1500 mm–2000 mm per year and with an average annual temperature of about 25°C. In the north, the country has a tropical climate with annual rainfall dropping to as low as 400 mm around the Lake Chad basin region and equally high temperatures of about 28°C (Molua 2006).

### Data collection

To be able to explore the key predictors of the arable production per capita index, the arable and permanent cropland and forest area, twelve time series data variables (Table 1) spanning the period 1961–2000 were collected from various sources. The rainfall data were collected from the climate database of the School of Geography and Environment at Oxford University and the United Nations Development Program (<http://www.geog.ox.ac.uk/research/climate/projects/undp-cp>, United Nations Development Program 2014). The population data were collected from the World Bank, World Development Indicators database ([www.google.com/publicdata](http://www.google.com/publicdata), World Bank 2014). The remaining variables: arable production per capita index, arable production and permanent cropland, cattle stock, CO<sub>2</sub> emissions, fuel wood, forest area, trade in forest products and logging, fertilizers, tractors-import value, tractors-quantity imported were retrieved from the Food and Agricultural Organization's database ([www.faostat.org](http://www.faostat.org), FAO 2014) and cross validated with those from the World Resources Institute's data base ([www.cait2.wri.org](http://www.cait2.wri.org), World Resource Institute 2014).

### Data analyses

Since the aim of this study was to verify the most important predictors of the arable production per capita

**Table 1 Abbreviated and complete names of the twelve variables under study**

Abbreviated name	Complete variable name
ArablePCL	Arable production and permanent cropland (in 'K' Ha)
ArableProd	Arable production per capita index (international \$)
CattleStock	Cattle stock (in 'K' heads)
CO2	Total CO <sub>2</sub> emissions including land use change (in 'K' metric tons)
Fertilizer	Fertilizer consumption (in 'K' metric tons of nutrients)
ForestArea	Forest area (in 'K' Ha)
FuelWood	Fuel wood (in 'K' cubic metres)
POPS	Population
Rainfall	Rainfall (mm)
TractorImport	Agricultural tractors (quantity imported)
TractorVal	Agricultural tractors (import value in 'K' \$)
TradeForest	Trade in forest products/exports/logging (international \$)

index, arable production and permanent cropland and forest area, SEM based on the two stage least square (2SLS) approach was employed. This method can be rationalised by the fact that there are several endogenous variables and exogenous variables whose feedbacks in the system have to be determined in order to identify the actual predictors of the endogenous variables. The analyses were undertaken in the free R statistical software version 2.12.0. Three scenarios with structural simultaneous equations were identified. However, a test of the hypothesis that there are no significant correlations among the variables is required in order to assert that all the variables are suitable for the analysis. The entire procedure of performing the 2SLS method in R has been described by Henningsen and Hamann (2007). From the correlation analysis performed, variables that do not bring in any new information are those with high correlations and for the models to be optimized, such variables were removed from the analysis. In this case, CO<sub>2</sub> emissions were removed because they have a correlation of 0.98 with population. Also, tractors-import values were removed because they have a correlation of 0.83 with tractors-quantity imported. In the paragraphs that follow, we describe the different scenarios, equations and the entire process of computing the predictors. The table that follows (Table 1) shows the abbreviated and complete names of all twelve variables. For a complete list of the time series data, see Additional file 1 section S1. The variables were abbreviated to facilitate handling in R. Furthermore, for the variables such as tractor (quantity imported) and tractors (import value) there were no data for the year 1978. This was dealt with by attributing to these variables the same measurements recorded in 1977 to avoid a gap

and rationalized by the fact that these data points are rising along the series (i.e. 1800 and 19218 K\$).

**Scenario one**

Scenario one was structured to have as endogenous variables the arable production per capita index and forest area (Figure 2). As a result of this, two structural simultaneous equations are derived to determine the key predictors of the two endogenous variables. The equations are:

$$Y_{AP} = \alpha_0 + \alpha_1 X_{FA} + \alpha_2 X_P + \alpha_3 X_R \tag{1}$$

Where  $Y_{AP}$  is the arable production per capita index (endogenous variable);  $\alpha_1 X_{FA}$  is forest area;  $\alpha_2 X_P$  is population;  $\alpha_3 X_R$  is rainfall (exogenous variables).

For the computation of these variables in the R interface, the procedure below is used; however, for the detailed codes used in R, see Additional file 1 section S2:

The 2SLS estimated parameters are as follows:

Model Formula: ArableProd ~ ForestArea + Population + Rainfall + Instruments: ~Rainfall + Population + FuelWood + Tradeforest

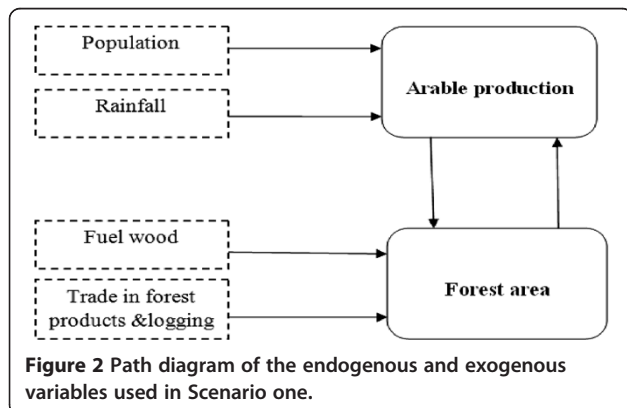
$$Y_{FA} = \beta_0 + \beta_1 X_{AP} + \beta_2 X_{FW} + \beta_3 X_{TF} \tag{2}$$

Where  $Y_{FA}$  is forest area (endogenous variable);  $\beta_1 X_{AP}$  is arable production;  $\beta_2 X_{FW}$  is fuel wood;  $\beta_3 X_{TF}$  is trade in forest products and logging (exogenous variables).

For the computation of these variables in the R interface, the procedure below is used; however, for the detailed codes used in R, see Additional file 1 section S2:

The 2SLS estimated parameters are as follows:

Model Formula: ForestArea ~ ArableProd + FuelWood + Tradeforest + Instruments: ~Rainfall + population + FuelWood + Tradeforest



**Scenario two**

Scenario two has been structured to have three key endogenous variables which are the arable production per capita index and forest area and arable production and permanent cropland (Figure 3). As a result of this, three structural simultaneous equations are derived to determine the key predictors of the three endogenous variables. The equations are:

$$Y_{AP} = \alpha_0 + \alpha_1 X_{FA} + \alpha_2 X_{A\_pcl} + \alpha_3 X_P + \alpha_4 X_R \tag{3}$$

Where:  $Y_{AP}$  is the arable production per capita index (endogenous variable);  $\alpha_1 X_{FA}$  is forest area;  $\alpha_4 X_R$  is rainfall;  $\alpha_2 X_{A\_pcl}$  is arable and permanent cropland;  $\alpha_3 X_P$  is population (exogenous variables).

For the computation of these variables in the R interface, the procedure below is used; however, for the detailed codes used in R, see Additional file 1 section S2:

The 2SLS estimated parameters are as follows:

Model Formula: ArableProd ~ ForestArea + ArablePCL + Population + Rainfall + Instruments: ~Rainfall + CattleStock + FuelWood + Tradeforest + Fertilizer + TractorImport

$$Y_{FA} = \beta_0 + \beta_1 X_{AP} + \beta_2 X_{A\_pcl} + \beta_3 X_{FW} + \beta_4 X_{TF} \tag{4}$$

Where:  $Y_{FA}$  is forest area (endogenous variable);  $\beta_1 X_{AP}$  is the arable production per capita;  $\beta_2 X_{A\_pcl}$  is arable production and permanent cropland;  $\beta_3 X_{FW}$  is fuel wood;  $\beta_4 X_{TF}$  is trade in forest products and logging (exogenous variables).

For the computation of these variables in the R interface, the procedure below is used; however, for the detailed codes used in R, see Additional file 1 section S2:

The 2SLS estimated parameters are as follows:

Model Formula: ForestArea ~ ArableProd + ArablePCL + FuelWood + Tradeforest + Instruments: ~Rainfall + Population + FuelWood + Tradeforest + Fertilizer + TractorImport

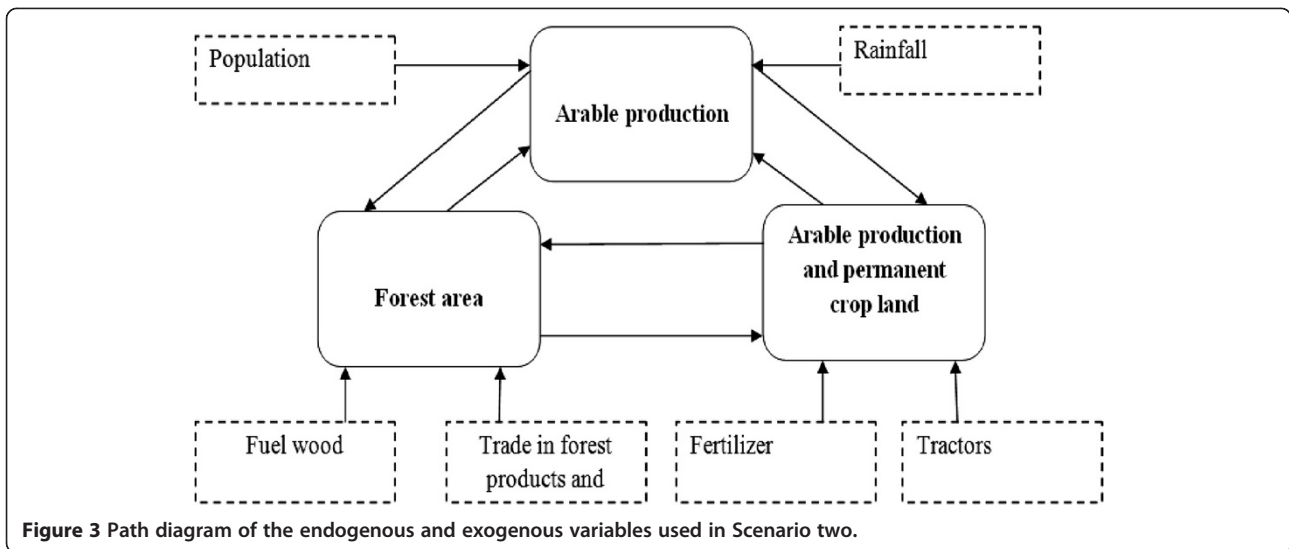
$$Y_{A\_pcl} = \gamma^0 + \gamma_1 X_{FA} + \gamma_2 X_{AP} + \gamma_3 X_F + \gamma_4 X_T \tag{5}$$

Where:  $Y_{A\_pcl}$  is arable production and permanent cropland (endogenous variable);  $\gamma_1 X_{FA}$  is forest area;  $\gamma_2 X_{AP}$  is the arable production per capita index;  $\gamma_3 X_F$  is fertilizers;  $\gamma_4 X_T$  is tractors (import value) (exogenous variables).

For the computation of these variables in the R interface, the procedure below is used; however, for the detailed codes used in R, see Additional file 1 section S2:

The 2SLS estimated parameters are as follows:

Model Formula: ArablePCL ~ ArableProd + ForestArea + Fertilizer + TractorImport + Instruments: ~Rainfall + CattleStock + FuelWood + Tradeforest + Fertilizer + TractorImport



**Figure 3** Path diagram of the endogenous and exogenous variables used in Scenario two.

**Scenario three**

Scenario three has been designed to have three key endogenous variables which are the arable production per capita index, cattle stock and arable production and permanent cropland (Figure 4). As a result of this, three structural simultaneous equations are derived to determine the key predictors of the three endogenous variables. The equations are:

$$Y_{AP} = \alpha_0 + \alpha_1 X_{CS} + \alpha_2 X_{A\_pcl} \tag{6}$$

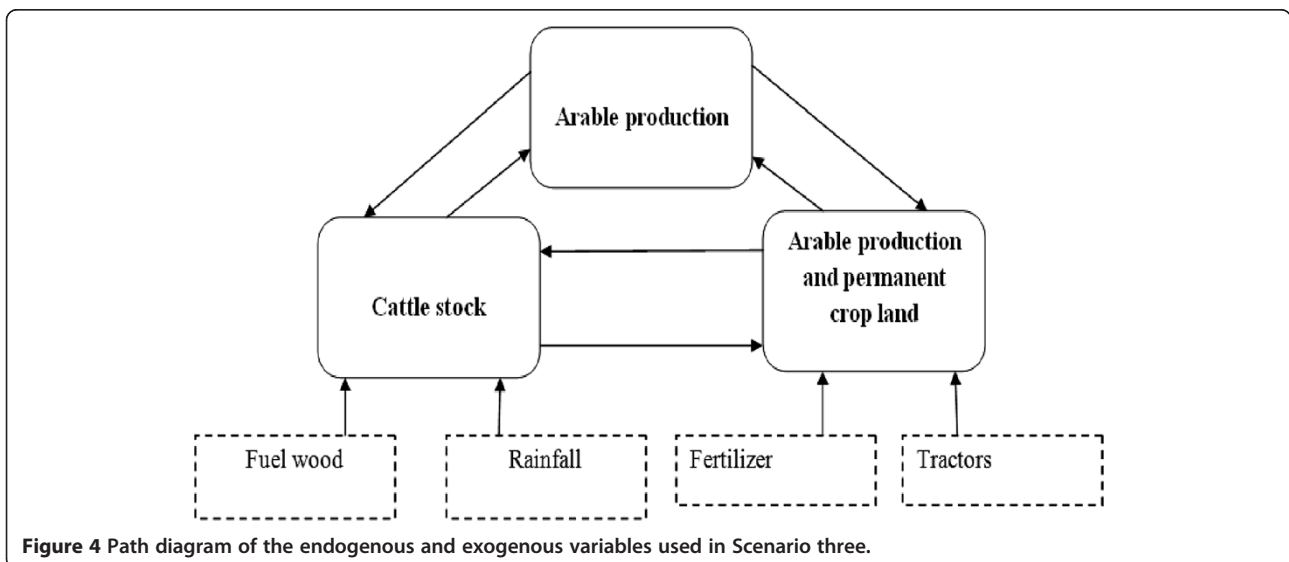
Where:  $Y_{AP}$  is the arable production per capita index;  $\alpha_1 X_{CS}$  is cattle stock;  $\alpha_2 X_{A\_pcl}$  is arable production and permanent cropland.

For the computation of these variables in the R interface, the procedure below is used; however, for the detailed codes used in R, see Additional file 1 section S2:

The 2SLS estimated parameters are as follows:  
 Model Formula:  $ArableProd \sim ArablePCL + CattleStock + Instruments: \sim Rainfall + FuelWood + Fertilizer + TractorImport$

$$Y_{CS} = \beta_0 + \beta_1 X_{AP} + \beta_2 X_{A\_pcl} + \beta_3 X_{FW} + \beta_4 X_{RF} \tag{7}$$

Where:  $Y_{CS}$  is cattle stock;  $\beta_1 X_{AP}$  is the arable production per capita index;  $\beta_3 X_{FW}$  is fuel wood;  $\beta_4 X_{RF}$  rainfall is rainfall;  $\beta_2 X_{A\_pcl}$  is arable production and permanent cropland.



**Figure 4** Path diagram of the endogenous and exogenous variables used in Scenario three.



For the computation of these variables in the R interface, the procedure below is used, however, for the detailed codes used in R, see Additional file 1 section S2:

The 2SLS estimated parameters are as follows:

Model Formula: CattleStock ~ ArableProd + ArablePCL + FuelWood + Rainfall + Instruments: ~Rainfall + FuelWood + Fertilizer + TractorImport

$$Y_{A\_pcl} = \gamma^0 + \gamma_1 X_{CS} + \gamma_2 X_{AP} + \gamma_3 X_F + \gamma_4 X_T. \quad (8)$$

Where:  $Y_{A\_pcl}$  is the arable production and permanent cropland (endogenous variable);  $\gamma_2 X_{AP}$  is the arable production per capita;  $\gamma_1 X_{CS}$  is cattle stock;  $\gamma_3 X_F$  is fertilizers;  $\gamma_4 X_T$  is tractors.

For the computation of these variables in the R interface, the procedure below is used; however, for the detailed codes used in R, see Additional file 1 section S2:

2SLS Estimates

Model Formula: ArablePCL ~ ArableProd + CattleStock + Fertilizer + TractorImport + Instruments: ~Rainfall + FuelWood + Fertilizer + TractorImport

Once the above scenarios were set up and the data were analysed, a normal probability test was used to show whether process data exhibit the standard normal bell curve or the Gaussian distribution.

## Results and discussion

### Scenario one

For the results on the most important variable impacting the first endogenous variable in Scenario one (arable production:  $Y_{AP}$ ), represented by Equation (1), the SEM shows that population growth ( $\alpha_2 X_P$ ) is the most significant determinant of ( $Y_{AP}$ ). This is supported by the fact that  $\alpha_2 X_P$  has the highest t value and the lowest p value (<0.05) when compared to those of rainfall ( $\alpha_3 X_R$ ) and forest area ( $\alpha_1 X_{FA}$ ). In addition, the coefficient of  $\alpha_2 X_P$  shows that a unit of change in  $\alpha_2 X_P$  will produce 6.80 units of change in  $Y_{AP}$ . Overall, we can say that when  $\alpha_2 X_P$  increases, arable production reduces mainly due to the inelastic nature of arable production in the short run and the increase in pressure of population growth on land.  $\alpha_1 X_{FA}$  is the second most significant variable and  $\alpha_3 X_R$  is the least most significant variable (Table 2). By implication, it can be judged that the  $\alpha_1 X_{FA}$  decline seen mainly through deforestation will have negative

**Table 2 SEM outputs for Equation (1), Scenario one**

Exogenous variables	Coefficients	Standard Error	t-values	p-value	Rank of t-values
$\alpha_1 X_{FA}$	-0.0065	0.00282	-2.32	0.02	2
$\alpha_3 X_R$	0.15	0.13	1.10	0.27	3
$\alpha_2 X_P$	-6.80	2.20	-3.06	0.004*	1

$\alpha_1 X_{FA}$  forest area,  $\alpha_2 X_P$  population,  $\alpha_3 X_R$  rainfall; \*most important predictor.

**Table 3 SEM outputs for Equation (2), Scenario one**

Exogenous variables	Coefficients	Standard Error	t-value	p-value	Rank of t-value
$\beta_1 X_{AP}$	55.94	29.65	1.88	0.06	2
$\beta_2 X_{FW}$	-0.008	0.64	-0.01	0.98	3
$\beta_3 X_{TF}$	-0.01	0.0038	-3.55	0.0010*	1

$\beta_1 X_{AP}$  arable production,  $\beta_2 X_{FW}$  fuel wood,  $\beta_3 X_{TF}$  trade in forest products and logging;

\*most important predictor.

repercussions on  $Y_{AP}$  because of increased soil erosion and reduced soil organic carbon and organic nitrogen. The fact that the influence of  $\alpha_3 X_R$  is least confirms the hypothesis that the influence of  $\alpha_3 X_R$  in the food sovereignty scheme of Cameroon and most sub-Saharan African countries is weak in spite of rising rainfall in most of sub-Saharan Africa.

Considering the second endogenous variable in scenario one ( $Y_{FA}$ ), represented by Equation (2) of the SEM, it is observed that trade in forest products and logging ( $\beta_3 X_{TF}$ ) is the most significant determinant of forest area ( $Y_{FA}$ ). This is supported by the fact that  $\beta_3 X_{TF}$  has the highest t value and the lowest p value (<0.05) when compared to those of arable production ( $\beta_1 X_{AP}$ ) and fuel wood ( $\beta_2 X_{FW}$ ). The only anomaly is that instead of  $\beta_3 X_{TF}$   $\beta_1 X_{AP}$  has the largest coefficient. This however does not imply  $\beta_1 X_{AP}$  is more significant because when the coefficients are divided by the standard errors, the t and p values remain the critical determinants of the most important variables because they consider both the coefficients and the standard deviation. In the latter situation,  $\beta_3 X_{TF}$  remains the most important determinant of  $Y_{FA}$ . Overall, we can say that when  $\beta_3 X_{TF}$  increases,  $Y_{FA}$  decreases (Table 3).

### Scenario two

For the results regarding the most important variable impacting the first endogenous variable in Scenario two (arable production:  $Y_{AP}$ ), represented by Equation (3), the SEM shows that population growth ( $\alpha_3 X_P$ ) is the most significant determinant of ( $Y_{AP}$ ). This is supported by the fact that  $\alpha_3 X_P$  has the highest t value, the highest

**Table 4 SEM outputs for Equation (3), Scenario two**

Exogenous variables	Coefficients	Standard Error	t-value	p-value	Rank of t-value
$\alpha_1 X_{FA}$	-0.03	0.0043	-7.59	0.05	2
$\alpha_4 X_R$	-0.09	0.10	-0.95	0.82	4
$\alpha_2 X_{A\_pcl}$	-0.04	0.006	-6.75	0.34	3
$\alpha_3 X_P$	-2.10	2.62	-8.16	0.02*	1

$\alpha_1 X_{FA}$  forest area,  $\alpha_4 X_R$  rainfall,  $\alpha_2 X_{A\_pcl}$  arable and permanent cropland,  $\alpha_3 X_P$  population;

\*most important predictor.

**Table 5 SEM outputs for Equation (4), Scenario two**

Exogenous variables	Coefficients	Standard Error	t-value	p-value	Rank of t-value
$\beta_1 X_{AP}$	-21.85	4.26	-5.11	1.12	3
$\beta_2 X_{A\_pcl}$	-3.10	0.06	-45.16	0.15*	1
$\beta_3 X_{FW}$	-1.43	0.09	-15.85	1.49	2
$\beta_4 X_{TF}$	0.0008	0.0006	1.44	1.27	4

$\beta_1 X_{AP}$  arable production,  $\beta_2 X_{A\_pcl}$ : arable production and permanent cropland,  $\beta_3 X_{FW}$  fuel wood,  $\beta_4 X_{TF}$  trade in forest products and logging; \*most important predictor.

coefficient and the smallest p value when compared to those of rainfall ( $\alpha_4 X_R$ ), forest area ( $\alpha_1 X_{FA}$ ) and arable and permanent cropland ( $\alpha_2 X_{A\_pcl}$ ). In addition, the coefficient of  $\alpha_3 X_P$  shows that a unit of change in  $\alpha_3 X_P$  will produce 2.10 units of change in  $Y_{AP}$ . Overall, we can say that when  $\alpha_3 X_P$  increases,  $Y_{AP}$  reduces mainly due to the inelastic nature of  $Y_{AP}$  in the short run and the increased pressure of population growth on land.  $\alpha_1 X_{FA}$  is the second most significant variable and  $\alpha_2 X_{A\_pcl}$  is the third one while  $\alpha_4 X_R$  is the least important (Table 4). The fact that the influence of  $\alpha_4 X_R$  is least tends to confirm further the hypothesis that the influence of  $\alpha_4 X_R$  in the food sovereignty scheme of Cameroon and most sub-Saharan African countries is weak.

Considering the second endogenous variable in Scenario two ( $Y_{FA}$ ), represented by Equation (4) of the SEM, it is observed that arable production and permanent cropland ( $\beta_2 X_{A\_pcl}$ ) are the most significant determinants of forest area ( $Y_{FA}$ ). This is supported by the fact that  $\beta_2 X_{A\_pcl}$  has the highest t value when compared to the other variables. The irregularities observed in the coefficients create anomalies and determines that the t values are the only reliable determinants of causality because the t value considers the coefficients and the standard deviation. Overall, we can say that when  $\beta_2 X_{A\_pcl}$  increases,  $Y_{FA}$  decreases as the expansion of farm lands is often at the expense of forest area (Table 5).

In the case of Equation (5) of scenario two, we observe that the most significant determinant of arable production and permanent cropland ( $Y_{A\_pcl}$ ) is the arable production per capita index ( $\gamma_2 X_{AP}$ ). This is supported by the fact that  $\gamma_2 X_{AP}$  has the highest t value and the lowest p value. Generally, we expect that an increase in  $\gamma_2 X_{AP}$  will trigger a

**Table 6 SEM outputs for Equation (5), Scenario two**

Exogenous variables	Coefficients	Standard Error	t-value	p-value	Rank of t-value
$\gamma_1 X_{FA}$	-0.18	0.01	-2.19	1.32	3
$\gamma_2 X_{AP}$	-6.63	3.03	-12.84	0.03*	1
$\gamma_3 X_F$	4.10	2.59	1.57	8.28	4
$\gamma_4 X_T$	0.02	0.003	6.58	0.12	2

$\gamma_1 X_{FA}$  forest area,  $\gamma_2 X_{AP}$  arable production,  $\gamma_3 X_F$  fertilizers,  $\gamma_4 X_T$ : tractors; \*most important predictor.

**Table 7 SEM outputs for Equation (6), Scenario three**

Exogenous variables	Coefficients	Standard Error	t-value	p-value	Ranks of t-value
$\alpha_1 X_{CS}$	-0.006	0.002	-3.04	0.004*	1
$\alpha_2 X_{A\_pcl}$	0.005	0.004	1.09	0.279	2

$\alpha_1 X_{CS}$  cattle stock,  $\alpha_2 X_{A\_pcl}$  arable production and permanent crop land, \*most important predictor.

decline in  $Y_{A\_pcl}$ . In the same way, when  $\gamma_1 X_{FA}$  increases,  $Y_{A\_pcl}$  will decrease (Table 6).

**Scenario three**

As concerns Equation (6) in Scenario three, we observe that the most important determinant of arable production ( $Y_{AP}$ ) is cattle stock ( $\alpha_1 X_{CS}$ ). This is seen as in this SEM equation, the latter has the highest t value and the lowest p value which is equally also less than 0.05. The overall implication of this equation is that  $\alpha_1 X_{CS}$  reduces  $Y_{AP}$  because an increase in cattle rearing requires more land and often leads to a reduction in arable farmland since cattle rearing is often land dependent (Table 7). This is this case in the northern regions of Cameroon and parts of the North West and Western Highlands which constitute the cattle rearing hub of the country.

In the case of Equation (7), Scenario three, we observe that arable production and permanent cropland ( $\beta_2 X_{A\_pcl}$ ) is the most influential variable affecting cattle stock ( $Y_{CS}$ ). It can be observed therefore that an increase in  $\beta_2 X_{A\_pcl}$  will trigger an increase in  $Y_{CS}$  since more land will be established for crop production and animal rearing. Fuel wood is ( $\beta_3 X_{FW}$ ) seen as the second most important variable here and its influence is seen as it enhances  $Y_{CS}$  because when trees are cut to produce fire wood, more land becomes available for cattle rearing (Table 8).

In the case of Equation (8), Scenario three, tractors ( $\gamma_4 X_T$ ) are seen as the most important variable affecting arable production and permanent cropland  $Y_{A\_pcl}$ . This is evident from the very low p value and the high t value. It can be suggested that the more tractors the greater the  $Y_{A\_pcl}$ . Also, arable production per capita index ( $\gamma_2 X_{AP}$ ) is the second most important variable (Table 9).

It can be summarized from these scenarios that population is the most important predictor of the arable production per capita index. The second and third predictors in

**Table 8 SEM outputs for Equation (7), Scenario three**

Exogenous variables	Coefficients	Standard Error	t-value	p-value	Rank of t-value
$\beta_1 X_{AP}$	-9.00	6.34	-1.41	0.99	3
$\beta_3 X_{FW}$	0.49	0.06	7.67	0.16	2
$\beta_4 X_{RF}$	-0.00069	6.04	-0.0001	5.23	4
$\beta_2 X_{A\_pcl}$	1.42	0.10	13.74	0.04*	1

$\beta_1 X_{AP}$  arable production,  $\beta_3 X_{FW}$  fuel wood,  $\beta_4 X_{RF}$  rainfall,  $\beta_2 X_{A\_pcl}$  arable production and permanent cropland; \*most important predictor.

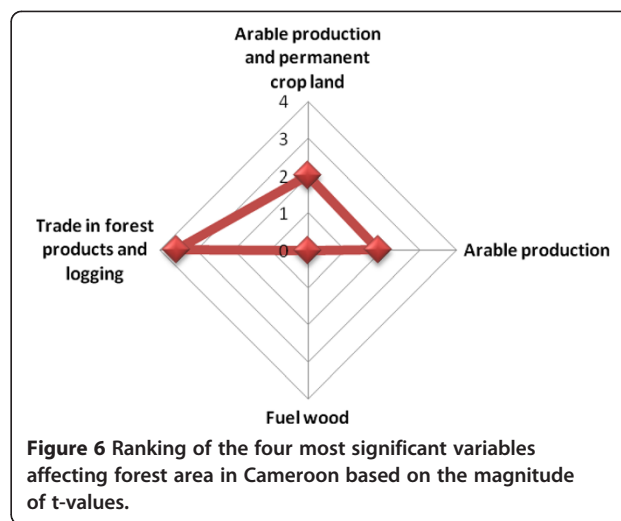
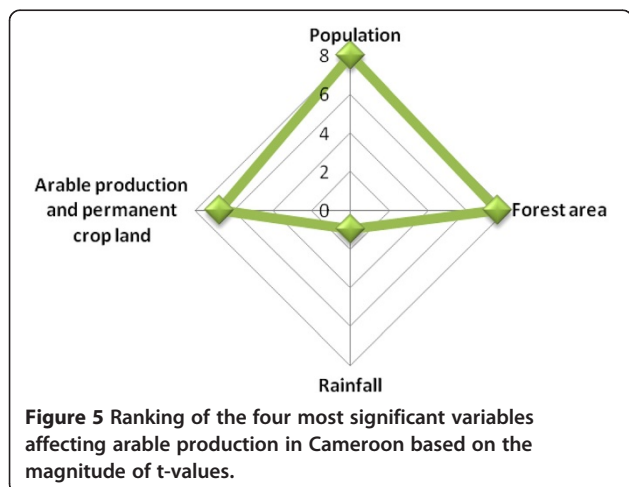
**Table 9 SEM outputs for Equation (8), Scenario three**

Exogenous variables	Coefficients	Standard Error	t-value	p-value	Rank of t-value
$\gamma_2 X_{AP}$	-59.27	1.30	-0.45	0.65	2
$\gamma_1 X_{CS}$	0.04	8.57	0.04	0.96	4
$\gamma_3 X_F$	12.52	3.03	0.41	0.68	3
$\gamma_4 X_T$	0.16	1.12	1.46	0.15*	1

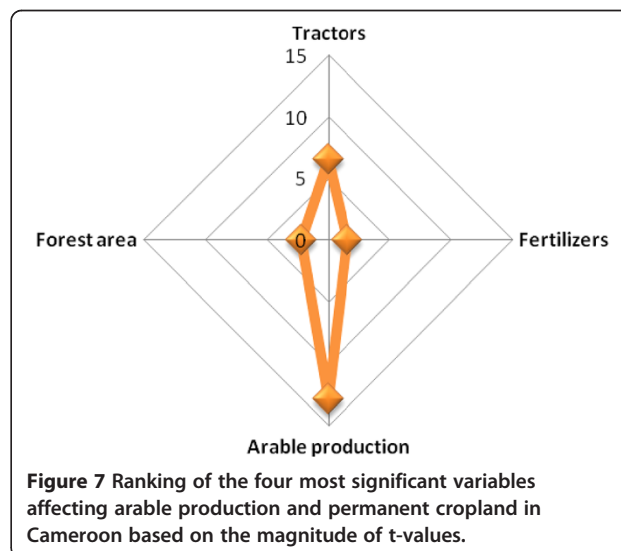
$\gamma_2 X_{AP}$  arable production,  $\gamma_1 X_{CS}$  cattle stock,  $\gamma_3 X_F$  fertilizers  $\gamma_4 X_T$  tractors; \*most important predictor.

order of importance are forest area and arable and permanent crop land. Rainfall, however is seen as the weakest of all the variables under study (Figure 5). These results are consistent with several previous findings. For example, Alexandratos (2005, 2008) argues that rapid population growth impacts environmental resources because of the pressure of population on finite resources. This is similar to the reducing effect that population growth in Cameroon has on arable production. Borlaug (1999) also affirms the influence of a galloping world population as a major constraint on world food yields. There is also existing literature that is consistent with the influence of rainfall on food production in Cameroon and most of sub Saharan Africa is becoming weaker. For example, Olsson and Mryka (2008), Eklundh and Olsson (2003) and Hulme (2001) argue that the current trends in decline in food production in most of the Sahel and sub Saharan Africa cannot be attributed to rainfall which tends to be increasing but to the dynamics of various human land use processes. This constitutes a response to the question: If rainfall is increasing in most of the Sahel since the 1990s why does the region still face acute problems of food security?

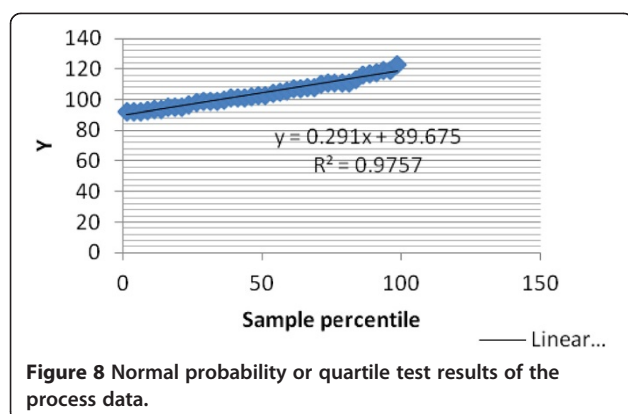
With regards to the most important predictors of forest area dynamics, trade in forest products and logging concessions are the most important ones. Arable production and permanent cropland are second and third



respectively (Figure 6). It has been argued that the increased profits obtained by timber companies and extractors of other forest products are key constraints on forest area decline in many parts of the world (Mertens and Lambin 1997; Carr et al. 2005; Vanclay 1993; Houghton 1991; Zhao et al. 2006). However, Angelsen and Kaimowitz (1999) argue that even when trade in forest products and logging are said to be dominant causes of forest area decline, population pressure is always at the centre of increased trade in forest products and logging. In the case of arable production and permanent cropland, it would be useful to consider how cropland expansion and cattle ranching have been able to produce forest area decline. In as much as this is true, the connection in all these studies is that as population grows, the need to feed more mouths increases and this often means more forest clearance.







Other factors that are considered important are the arable and permanent cropland as well as the arable production per capita index which are the second and third most valuable factors that explain deforestation in Cameroon. This result is highly consistent with other studies elsewhere. A study carried out in central Argentina reports that agricultural expansion for the purpose of cultivating Soya beans in particular is the main proximate cause of forest loss (Zak et al. 2008); this a view supported by Geist and Lambin (2002), when they report that about 96% of deforestation in Africa, Asia and Latin America is caused by agricultural expansion. As such, agricultural expansion remains a significant cause of deforestation and at times, it is related to population pressure. In support of this, it has been stated that one of the principal causes of deforestation in Panama has been the expansion of agricultural frontiers through extensive shifting cultivation systems.

Finally, the most important predictors of arable production and permanent cropland in order of importance are the arable production per capita index, tractors, forest area and fertilizers (Figure 7). Agricultural expansion is also argued to be a key driver of arable and permanent cropland. This is seen as the incentives to cut more trees and establish more farmland are driven by arable production per capita. This is consistent with various studies (Zak et al. 2008; Geist and Lambin 2002; Angelsen and Kaimowitz 1999).

To further verify the reliability of the data used in this study, a normal probability or quartile test was performed. This test is used to show whether process data exhibit the standard normal bell curve or Gaussian distribution or to tell how well the plotted points fit the normal line (Figure 8). If they fit well then, it can be assumed that the processed data are normally distributed. In our case, all the points fit the curve line adequately and an  $R^2$  of about 97% is obtained indicating a very high level of reliability.

## Conclusions

This study has demonstrated that the arable production per capita index is impacted more by population while

the influence of rainfall on arable production is weak. The fact that the influence of rainfall is weak goes to show that the declining trends in production are more associated with human related land use activities. Furthermore, arable production and permanent cropland has as its main predictor arable production per capita. Forest area is seen to be vulnerable to trade in forest products and logging rather than to any other variable. This implies that, the loss in Cameroon's forests are more linked to the commercialization of the forest than on the use of the forest for basic subsistence such as fuel wood collection by the local populations.

SEM models have been described as models that capture the feedbacks between several variables while determining the predictors of specific variables. While the aspect of introducing several endogenous and exogenous variables mimics reality, it in fact creates complexity that other models do not. As such, it is somehow difficult to use SEM to determine the predictors of a single endogenous variable because SEMs are structured to create diversity and consider several endogenous variables and instruments. As a result of this, this study observes that multiple linear regression models could be used when the objective is to create less diversity and to identify the predictors of only one dependent variable. Furthermore, it could be of pertinence if the vulnerability of specific crops to variables such as rainfall, population, fertilizers and machinery at the centre of interest. This would provide crop specific vulnerability information needed to inform policy. A meta analysis of the relative contributions of organic and conventional fertilizers on crop yields in Cameroon and Africa could also be valid alternatives for further research.

## Additional file

**Additional file 1:** Synthesis of the time series data of the twelve variables under used in this study.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contribution

E.T.E. acquired, cleaned and analysed the data; wrote and edited the manuscript. C.R.B. edited the manuscript and supervised the entire study. C. A. and M.A.S. had discussions with E.T.E. on the issues the paper handles and they also proof read the paper. C.P. assisted in providing information on the modelling approach. All authors read and approved the final manuscript.

### Acknowledgements

We are thankful to the Fonds de recherche du Québec -Société et Culture for funding this study through post doctoral grant number (2015-B3-180319). We also thank the editor and three anonymous reviewers for their comments and suggestions.

### Author details

<sup>1</sup>Département de Géographie, Université de Montréal, Pavillon 520, ch. de la Côte-Sainte-Catherine, Local 332-3, CP 6128, succursale Centre-ville, Montréal, Québec H3C 3J7, Canada. <sup>2</sup>Institute of Environmental Sciences, University of

Quebec at Montreal (UQAM), CP 8888, succursale Centre-Ville, Montréal, QC H3C 3P8, Canada.

Received: 5 August 2014 Accepted: 4 October 2014  
Published: 11 October 2014

## References

- Alexandratos N (2005) Countries with rapid population growth and resource constraints: issues of food, agriculture, and development. *Popul Dev Rev* 31(2):237–258
- Alexandratos N (2008) Food price surges: possible causes, past experiences, relevance for exploring long-term prospects. *Popul Dev Rev* 34(4):663–697
- Amujoyegbe B, Opabode A, Olayinka A (2010) Effect of organic and inorganic fertilizer on yield and chlorophyll content of maize (*Zea mays* L.) and sorghum (*Sorghum bicolor* L.). *Afr J Biotechnol* 6:1869–1873
- Angelsen A, Kaimowitz D (1999) Re-thinking the causes of deforestation: lessons from economic models. *World Bank Res Obs* 14:73–98
- Borlaug N (1999) Feeding a World of 10 Billion People: the Miracle Ahead. Lecture presented at De Montfort University, Leicester, UK
- Bruinsma J (2009) The resource outlook to 2050. In: Expert meeting on how to feed the world. Rome, Italy. 24–26 June 2009. FAO, Rome
- Carr D, Suter L, Barbieri A (2005) Population dynamics and tropical deforestation: state of the debate and conceptual challenges. *Popul Environ* 27:90–113
- Central Intelligence Agency (2012) Population Growth Rate., Available online at: www.cia.gov
- Dubois O (2011) The state of the world's land and water resources for food and agriculture: managing systems at risk. FAOSTAT, Rome, Italy
- Eklundh L, Olsson L (2003) Vegetation index trends for the African Sahel 1982–1999. *Geophys Res Lett* 30(8):1430, doi:10.1029/2002GL016772
- Epule TE, Changhui P, Laurent L, Zhi C (2011) Forest loss triggers in Cameroon: a quantitative assessment using multiple linear regression approach. *J Geogr Geol* 3(1):30–40
- Epule TE, Changhui P, Laurent L, Zhi C, Nguh BS (2012) The environmental quadruple: forest area, rainfall, CO<sub>2</sub> emissions and arable production interactions in Cameroon. *Br J Environ Clim Change* 2(1):12–27
- FAO (2006) World agriculture: towards 2030/2050 – interim report. Rome, Online available at: http://www.fao.org/fileadmin/user\_upload/esag/docs/Interim\_report\_AT2050web.pdf
- FAO (2010) Global Forest Resources Assessment (GFRA). Main report. FAO Forestry paper. Rome., ISBN 978-92-5-106654-6. Online available at: http://www.fao.org/forestry/fra/fra2010/en/
- FAO (2014) FAOSTAT., Online available online at: http://faostat.org
- FAO, UNIDO (2008) Agricultural mechanization in Africa: time for action. Planning investment for enhanced agricultural productivity report of an expert group meeting in January 2008, Vienna Austria. FAO, Rome., Online available at: http://www.unido.org/fileadmin/user\_media/Publications/Pub\_free/agricultural\_mechanization\_in\_Africa.pdf
- García-Ponce E, Gomez-Macpherson H, Diallo O, Djibril M, Baba C, Porcel O, Mathieu B, Comas J et al (2012) Contribution of sorghum to productivity of small-holder irrigation schemes: on-farm research in the Senegal River Valley, Mauritania. *Agric Syst* 115:72–82
- Geist HJ, Lambin EF (2002) Proximate causes and underlying driving forces of tropical deforestation. *Bio Science* 52:143–150
- Henningsen A, Hamann JD (2007) Systemfit: a package for estimating systems of simultaneous equations in R. *J Stat Softw* 23(4):10–24, http://www.jstatsoft.org/v23/i04/paper
- Houghton RA (1991) Tropical deforestation and atmospheric carbon dioxide. *Climate Change* 19:99–118
- Hulme M (2001) Climatic perspective on Sahelian desiccation: 1973–1998. *Glob Environ Chang* 11:19–29
- INS (2009) Annuaire Statistique du Cameroun 2004. Institut National de la Statistique. Ministère de l'Economie et des Finances, Yaounde'
- Kombiok J, Buah S, Dzomeku L, Abdulai H (2013) Sources of pod yield losses in groundnut in the Northern Savanna zone of Ghana. *West African Journal of Applied Ecology* 20:53–63
- Mertens B, Lambin E (1997) Spatial modeling of deforestation in southern Cameroon: spatial disaggregation of diverse deforestation processes. *Appl Geogr* 17(2):143–162
- Molua E (2006) Climate trends in Cameroon: implications for agricultural management. *Clim Res* 30:255–262
- Mundex Dataset (2012) Cameroon Population Growth Rate., Online available at: http://www.indexmundi.com
- Olsson L, Mryka H (2008) Greening of the Sahel. In: The Encyclopedia of Earth., Online available at: http://www.eoearth.org/article/Greening\_of\_the\_Sahel
- Petraitis PS, Dunham AE, Niewiarowski PH (1996) Inferring multiple causality: the limitations of path analysis. *Funct Ecol* 10:421–431
- Rosegrant MW, Cline SA (2003) Global food security: challenges and policies. *Science* 302:1917–1919
- Rosegrant MW, Svendsen M (1993) Asian food production in 1990s: irrigation investment and management in Asia. *Food Policy* 18:13–32
- United Nations Development Program (2014) Rainfall data country profiles. Department of Geography, Oxford University., Online available online at: http://www.geog.ox.ac.uk/research/climate/projects/undp-cp/
- Vanclay J (1993) Saving the tropical forest: needs and prognosis. *Ambio* 22(4):225–231
- World Bank (2014) World development indicators. World Bank., Available online at: www.google.com/publicdata
- World Resource Institute (2014) Climate data explorer., Online available at: www.cait2.wri.org
- Yengoh GT, Ardo J (2014) Crop yield gaps in Cameroon. *Ambio* 43:175–190, doi:10.1007/s13280-013-0428-0
- Zak RM, Cabido DC, Diaz S (2008) What drives accelerated land cover change in central Argentina? Synergistic consequences of climatic, socioeconomic, and technological factors. *Environ Manag* 42:181–189
- Zhao S, Peng C, Jiang DT, Lei X, Zhou X (2006) Land use change in Asia and ecological consequences. *Ecol Res* doi:10.1007/s11284-006-0048-2

doi:10.1186/2193-1801-3-597

**Cite this article as:** Epule et al.: 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. *SpringerPlus* 2014 **3**:597.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)