A quantitative analysis of socio-economic determinants influencing crop drought vulnerability in Sub-Saharan Africa

Author(s):
Kamali, Bahareh; Abbaspour, Karim C.; Wehrli, Bernhard; Yang, Hong

Publication Date:
2019-11

Permanent Link:
https://doi.org/10.3929/ethz-b-000378795

Originally published in:
Sustainability 11(21), http://doi.org/10.3390/su11216135

Rights / License:
Creative Commons Attribution 4.0 International
A Quantitative Analysis of Socio-Economic Determinants Influencing Crop Drought Vulnerability in Sub-Saharan Africa

Bahareh Kamali 1,*, Karim C. Abbaspour 1,*, Bernhard Wehrli 2 and Hong Yang 1,3

1 Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Duebendorf, Switzerland; karim.abbaspour@eawag.ch (K.C.A.); hong.yang@eawag.ch (H.Y.)
2 Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, 8092 Zurich, Switzerland; bernhard.wehrli@eawag.ch
3 Department of Environmental Sciences, University of Basel, 4056 Basel, Switzerland
* Correspondence: bahareh.kamali@eawag.ch

Received: 4 September 2019; Accepted: 31 October 2019; Published: 3 November 2019

Abstract: Drought events have significant impacts on agricultural production in Sub-Saharan Africa (SSA), as agricultural production in most of the countries relies on precipitation. Socio-economic factors have a tremendous influence on whether a farmer or a nation can adapt to these climate stressors. This study aims to examine the extent to which these factors affect maize vulnerability to drought in SSA. To differentiate sensitive regions from resilient ones, we defined a crop drought vulnerability index (CDVI) calculated by comparing recorded yield with expected yield simulated by the Environmental Policy Integrated Climate (EPIC) model during 1990–2012. We then assessed the relationship between CDVI and potential socio-economic variables using regression techniques and identified the influencing variables. The results show that the level of fertilizer use is a highly influential factor on vulnerability. Additionally, countries with higher food production index and better infrastructure are more resilient to drought. The role of the government effectiveness variable was less apparent across the SSA countries due to being generally stationary. Improving adaptations to drought through investing in infrastructure, improving fertilizer distribution, and fostering economic development would contribute to drought resilience.

Keywords: Crop model; EPIC; regression techniques; drought adaptation

1. Introduction

The current world population has already reached 7 billion, and by 2050 it is estimated to increase to 9 billion, with the most significant increase concentrated in South Asia and Sub-Saharan Africa (SSA) [1,2]. A number of studies show that if no drastic actions are taken in consumption patterns, shifting diets, and reducing wastes, meeting the world’s food demand will require doubling food production and this will unavoidably exert extreme pressure on agricultural and natural resources [3]. On the other hand, climate change has exacerbated the situation by increasing drought risks and flood events around the globe with significant impacts on agricultural production. As agriculture in most SSA countries is primarily rainfed, SSA is at the core of this threat. Achieving food security in SSA is an enormous challenge given the currently weak institutions, poor infrastructure, population, unfavorable weather conditions, and high social dependence on declining natural resources [4]. Therefore, to better cope with the consequences of droughts and mitigate their effects, it is essential to incorporate socio-economic information in the analysis of drought vulnerability in SSA.

Drought is a natural disaster, and the degree of its impacts is well recognized in terms of the magnitude of vulnerability. Various definitions have been proposed for vulnerability.
The Intergovernmental Panel on Climate Change (IPCC) defined vulnerability as a function of exposure, sensitivity, and adaptive capacity [5]. The fifth assessment report focused on socio-economic aspects and prioritizing adaptation interventions. Adaptive capacity is the ability of a system to cope with the consequence of drought and represents the potential to implement measures that help reduce potential impacts [6]. Quantification of adaptation capacity is a challenge for two reasons. First, adaptation is the intrinsic property of a system and only manifests itself when the system is exposed to a shock [7]. Second, adaptation is heterogeneous over space and time, and therefore, it is challenging to generalize a set of universal factors that enhance adaptive capacity in different regions over time.

Several studies have attempted to incorporate adaptive capacity in the vulnerability assessment using approaches such as aggregated quantitative indicators [8–10] or semi-structured interviews [11–13]. Indicators of adaptive capacity were also grounded in sustainable livelihood theory as a set of different forms of assets to which people have access, such as financial, human, resource, or physical assets [14]. This approach calculates a composite index for each constituent asset. During the last two decades, many studies also attempted to employ the sustainable livelihood concept in their analysis of vulnerability [13,15,16].

The use of indicators is one of the most common ways to define factors influencing vulnerability. Despite the extensive research, improvement is required to identify and choose driving forces and indicators that more accurately determine the relationship between socio-economic factors and vulnerability [17]. Detailed studies along these lines have been recently conducted at the European [7,11,18], Australian [13,16,19], and in few cases at global scales [8,20], but the issue has not been thoroughly addressed in SSA countries. Naumann et al. [21] calculated composite indicators that reflect different aspects of vulnerability and adaptive capacity at the Pan-African level. Their study, however, lacks an empirical and analytical framework that can clearly explain the direction and magnitude of effectiveness (or influence) of individual factors on vulnerability. This limitation impedes policy development to mitigate the vulnerability. Several studies have used regression models to analyze the relationship between the socio-economic variables and crop drought vulnerability in some individual countries or regions in SSA [22,23]. However, these studies are case-specific, and it is challenging to upscale their findings to the whole of SSA.

In assessing the factors influencing vulnerability to drought, most studies have not considered the temporal variations in different countries. Multivariable regression models used for these analyses often take the prevailing situation of individual countries over a specified time period. In reality, vulnerability is time-dependent and varies with the severity of the drought. It is, therefore, important to consider changes in vulnerability and factors that influence it on temporal dimension. As data is increasingly becoming available in developing countries, panel data analysis at the national level provides a more significant basis for modeling the complexity of socio-economic factors influencing vulnerability [24]. Panel analysis is a statistical method to analyze two-dimensional data collected over time for the same entity. However, such an analysis exceeding a single cross-section of time has so far not been conducted in the context of crop drought-vulnerability assessment in SSA.

In the literature, most studies quantifying drought vulnerability are based on some arithmetic aggregation of several representative components. Very few studies have quantified the drought vulnerability by comparing the expected crop yield under certain climate conditions with the recorded yield. The magnitude of the difference between modeled and recorded yield is an indicator of resiliency of a region to drought occurrence. Process-based crop models are practical tools to model how yield changes in response to climate variability or drought. Linking the drought vulnerability defined as the distance between the expected and recorded yield with socio-economic variables may, therefore, enhance the choice of management options.

The aim of this study is to identify those socio-economic variables that contribute to drought-vulnerability of maize production in SSA. To this end we: (1) construct a definition of Crop Drought Vulnerability Index (CDVI) obtained by comparing simulated yield from a crop model with recorded yield at the country level; (2) compare the adaptive capacity variables for economic,
human, resource, infrastructure, and governance categories across countries; (3) apply multivariable panel regression to identify socio-economic variables influencing CDVI for maize; and finally (4) discuss the implications of the results for mitigation of maize vulnerability to drought in SSA.

2. Materials and Methods

2.1. Site Description

SSA is home to over 1 billion people. The average annual precipitation of 795 mm yr\(^{-1}\) is diversely distributed in different regions [25]. Small landholders depend on rainfed agriculture as their primary source of livelihood. The effects of population increase and climate change have exacerbated the risk of hunger [26]. Therefore, reducing crop vulnerability to drought is essential for this region. Maize is the most widely grown crop and staple food in SSA [27] and represents a relevant case for this study. In the last two decades, average maize yields in SSA have increased from around 1.4 to 1.8 t ha\(^{-1}\), with over 40% increase in South Africa, but they are still at the very bottom of globally reported maize yields [28].

2.2. Quantifying Drought Exposure Index

Drought exposure index, used in this study to separate drought from non-drought years, was obtained from the Standardized Precipitation Index (SPI) [29]. To obtain SPI, first, a suitable cumulative probability distribution function (here Gamma distribution) [30,31] is fitted to the precipitation. We chose a two-parameter gamma distribution as the cumulative distribution function [30,31]. SPI is then obtained from applying the inverse normal function with a mean of 0 and a standard deviation of 1 to the cumulative distribution function. After this transformation, 99.7% of data vary between −3 and 3. Negative values are representative of drought situations, whereas the positive values show non-drought cases (Table 1). The five classifications within the [−3, 3] range are defined as wet, near normal, mild drought, moderate drought, and severe drought [30] (Table 1).

Table 1. Five categories of the Standardized Precipitation Index (SPI) and Crop Drought Vulnerability index (CDVI). SPI is selected to separate drought and non-drought years [30].

<table>
<thead>
<tr>
<th>SPI Classes</th>
<th>CDVI Classes</th>
<th>SPI or CDVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wet</td>
<td>No drought vulnerability</td>
<td>1.0 and more</td>
</tr>
<tr>
<td>Near normal</td>
<td>Low vulnerability</td>
<td>0.001 to 0.99</td>
</tr>
<tr>
<td>Mild</td>
<td>Mild vulnerability</td>
<td>−0.99 to 0.0</td>
</tr>
<tr>
<td>Moderate</td>
<td>High vulnerability</td>
<td>−1.499 to −1.0</td>
</tr>
<tr>
<td>Severe</td>
<td>Very high vulnerability</td>
<td>−1.50 or less</td>
</tr>
</tbody>
</table>

The analysis was carried out for years 1990–2012. Precipitation was obtained from the WFDEI (WATCH-Forcing-Data-ERA-Interim) meteorological forcing data [32] at 0.5° resolution. The grid level precipitation data were aggregated to the national level using weighted areal average of maize cultivated lands obtained from MIRCA2000 version 1.1 [33]. SPIs were calculated at different time scales. We tested the suitability of 1, 3, 6, 9, and 12-month time scales. For each country, we selected the time scale with the highest correlation coefficient between SPI and simulated maize yield during the growing season.

2.3. Definition of Crop Drought Vulnerability Index (CDVI)

We focused on maize vulnerability to drought occurrences. The definition of maize CDVI follows the study of Bryan, et al., [13], which is the difference between the recorded and modeled yields. In this study, the impact of drought on yield was quantified using the EPIC (Environmental Policy Integrated Climate) crop model. EPIC simulates different processes of farming systems as well as their interactions using data such as weather, soil, landuse, and crop management parameters [34].
EPIC operates on a daily time step and is capable of simulating crop growth under various climatic and environmental conditions and complex management schemes [34].

EPIC is a field-scale model. Following the work done by Kamali et al., [35], its application was extended to larger scales using a Python framework named EPIC+. The framework divides the region of study into several grids based on a specified resolution (here $0.5^\circ$) and executes EPIC on each grid cell. EPIC+ is also equipped with the Sequential Uncertainty Fitting (SUFI-2) [36] algorithm for calibration. More details on the procedure used to simulate and calibrate maize using EPIC+ are found in the work of Kamali, et al., [35].

The yield simulated by EPIC+ reflects the influence of climate variability on crop production and, therefore, represents the expected yield ($Y^{Exp}$). The actual yield ($Y^{Act}$) is obtained from FAO, which is reported at the national level [28]. The recorded yield is influenced by both climate factors and many socio-economic factors. We defined the variable $Y$ as $Y^{Act} - Y^{Exp}$. If $Y^{Act}$ is larger than $Y^{Exp}$ (positive $Y$), then the region is considered to be resilient to changing climate conditions, and so the vulnerability is low. However, if $Y^{Exp}$ is larger than $Y^{Act}$ (negative $Y$), then the maize production is sensitive to climate variability, and consequently the vulnerability is high. Finally, if $Y^{Act}$ is equal to $Y^{Exp}$ (Zero $Y$), then the region has effective adaptation. For better spatial and temporal comparisons of $Y$ across SSA, we normalized it using $Z$-score transformation calculated as [37]:

$$CDVI_t = \frac{Y_t - \text{Mean}(Y)}{\text{STD}(Y)}$$ (1)

where $Y_t$ is the yield residuals for year $t$ i.e., $Y_t = Y^{Act}_t - Y^{Exp}_t$; Mean$(Y)$ and STD$(Y)$ are the mean and standard deviation of the yield residuals $(Y)$, respectively. Similarly, and using this transformation, CDVI can also be classified into five levels of vulnerability severity (Table 1).

### 2.4. Selecting Socio-Economic Variables Relating to CDVI

To select the variables influencing CDVI, we used the following step-wise approach:

**Step (1):** Candidate socio-economic variables were collected as recommended by Brooks, et al., [38] and Naumann, et al., [21]. The two studies together introduced over 50 socio-economic variables that might be important for climate vulnerability assessments at the national scale. We selected 17 drought-relevant variables with more than 50% data availability for the study period (1990–2012) (Table 2). For the ease of referencing, we classified the variables into five categories as economic, human, resource, infrastructure, and governance. To fill the missing values of the variables, we used the spline interpolation procedure applied by Simelton, et al., [9]. For variables such as fertilizer use, where temporal data was missing for a certain period, we used the average of the years with available data.

**Step (2):** The variables were expressed in a variety of ranges or scales. In this step, they were transformed into a uniform dimension. We use the $Z$-score transformation as the normalization technique calculated as the ratio of the residual of the variable and its mean divided by the standard deviation as [39].

$$N_{\text{variable}}_{t,i} = \frac{\text{variable}_{t,i} - \text{Mean}(\text{variable}_{i})}{\text{STD}(\text{variable}_{i})}$$ (2)

where Mean$(\text{variable}_{i})$ and STD$(\text{variable}_{i})$ are the mean and standard deviation of variable $i$ ($i = 1, 2, \ldots, 17$) in year $t$. Equation (2) converts all variables to a common scale with an average of 0 and standard deviation of 1. It is also consistent with the standardized scale of CDVI. For those variables such as “Interest payment” (Table 2), which are negatively correlated to CDVI, the inverse of values were used (i.e., $1/\text{variable}$) in normalization. We did not transform variables of the governance category as these variables were obtained from a combination of different criteria that were already normalized during the procedure of quantifying these variables.
Table 2. Potential socio-economic variables influencing drought vulnerability and their definitions (variables are defined per year).

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>Definition</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic</td>
<td>GDP/capita</td>
<td>Gross domestic product</td>
<td>US$/capita</td>
</tr>
<tr>
<td></td>
<td>Interest payment</td>
<td>Interest payments on external debt</td>
<td>% GNI</td>
</tr>
<tr>
<td></td>
<td>GNI</td>
<td>Gross national income</td>
<td>US$/capita</td>
</tr>
<tr>
<td></td>
<td>Agriculture GDP</td>
<td>Agriculture GDP</td>
<td>% GDP in total GDP</td>
</tr>
<tr>
<td>Human</td>
<td>HDI</td>
<td>Human development index</td>
<td>ratio ranging from 0 to 1</td>
</tr>
<tr>
<td></td>
<td>Health expenditure</td>
<td>Health expenditure per capita</td>
<td>US$/capita</td>
</tr>
<tr>
<td></td>
<td>Maternal mortality</td>
<td>Maternal mortality ratio</td>
<td>per 100,000 live births</td>
</tr>
<tr>
<td></td>
<td>Calorie intake</td>
<td>Calorie intake per capita</td>
<td>calorie</td>
</tr>
<tr>
<td></td>
<td>Food production index</td>
<td>Food crops that are edible and contain nutrients excluding coffee and tea. (average of 2004–2006 equals 100)</td>
<td>ratio of each year to the base period (2004–2006)</td>
</tr>
<tr>
<td>Resource</td>
<td>Agricultural area (ha/capita)</td>
<td>per capita land area that is either arable, under permanent crops, or under permanent pastures</td>
<td>ha/capita</td>
</tr>
<tr>
<td></td>
<td>Fertilizer use (t/ha)</td>
<td>Nitrogen fertilizer use</td>
<td>tons/ha</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Water access (%)</td>
<td>% of population with access to improved drinking water source</td>
<td>percentage</td>
</tr>
<tr>
<td></td>
<td>Electricity access (%)</td>
<td>% of rural population with access to electricity</td>
<td>percentage</td>
</tr>
<tr>
<td>Governance</td>
<td>Control of corruption</td>
<td>The extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as “capture” of the state by elites and private interests</td>
<td>normalized values ranging from −2.5 to 2.5</td>
</tr>
<tr>
<td></td>
<td>Government effectiveness</td>
<td>The quality of public and civil service, policy formulation and implementation, the degree of its independence to political pressures, the credibility of the government’s commitment to policies</td>
<td>normalized values ranging from −2.5 to 2.5</td>
</tr>
<tr>
<td></td>
<td>Political stability</td>
<td>The likelihood that the government will be destabilized or overthrown by unconstitutional or violent means</td>
<td>normalized values ranging from −2.5 to 2.5</td>
</tr>
<tr>
<td></td>
<td>Voice &amp; accountability</td>
<td>The extent to which a country’s citizens are able to participate in selecting their government, freedom of expression, freedom of association, and a free media</td>
<td>normalized values ranging from −2.5 to 2.5</td>
</tr>
</tbody>
</table>

Sources: The variables in the first four categories were obtained from the World Bank (http://www.worldbank.org). The variables for the “governance” category were obtained from Kaufmann et al., [40] and can be downloaded from www.govindicators.org.

Step (3): The influence of ‘multicollinearity’ is an essential issue in regression models which can seriously distort the interpretation of a model [41]. This is related to situations when predictors are correlated with each other. Therefore, some factors will be redundant. To avoid the redundancy of variables, a bivariate correlation matrix was constructed between Nvariables of each category. This helped us to evaluate the strength and direction of the linear relationships between the variables [39]. The statistically most significant variables were selected, and the rest were removed from further assessment.
Step (4): The selected variables of each country were averaged to obtain an aggregated value for the economic, human, resource, infrastructure, and governance categories. This facilitates the spatial comparison of the selected aspects and provided insights into the adaptive capacity of different countries.

Step (5): Statistical regression models were constructed to analyze the role of selected variables in Step 3 for characterizing CDVI (see next Section 2.5 for details).

2.5. Linking CDVI and Socio-Economic Variables Using Regression Models

Two model formulations were constructed, and their results were compared to get some insights into the influence of the temporal dimension on explanatory variables. In both models, the socio-economic variables selected in Step 3 were considered as explanatory variables and CDVI as a dependent variable. The first model (Model I) did not take into account temporal variability of explanatory and dependent variables. Only their averages were included. The aim of this model was to check if the temporally aggregated variables (during drought years) can explain the relationship between socio-economic conditions and maize vulnerability. The model I is expressed as:

\[
\text{Model I: } \text{CDVI}_i = \alpha_0 + \beta \bar{x}_i + \epsilon_i
\]

where \(\text{CDVI}_i\) and \(\bar{x}\) stand for average CDVI and selected socio-economic variables for country \(i\) during drought years respectively, \(\alpha_0\) is a constant intercept term, \(\beta\) is a vector of coefficient for each explanatory variable, and \(\epsilon_i\) is the error term. The model is fitted with all socio-economic variables and simplified by comparing versions with and without a particular explanatory term using the Likelihood Ratio tests. In each step, we remove insignificant terms until all remaining factors are significant at 1%, 5%, or 10% levels.

Model II takes into account the temporal dimension of socio-economic factors using the panel regression method. A Linear Fixed-effect model is designed to analyze the impact of variables that vary over time. The formulation for Model II is expressed as:

\[
\text{Model II: } \text{CDVI}_{z,t} = \alpha_i + \beta x_{i,t} + \epsilon_{i,t}
\]

where \(\text{CDVI}_{z,t}\) is the dependent variable at time \(t\) for country \(z\), and \(x\) stands for selected socio-economic variables as the dependent variables. In this model, each country has its characteristics that may or may not influence the predictor. The \(\alpha_i\) values take into account the individual effects of country \(z\) that are not explained by the dependent variables included in Model II. Explanatory variables were added one by one in a forward stepwise manner to check their significance. The model was then selected by comparing versions with and without a particular explanatory term using Likelihood Ratio tests. In each step, we selected a model in which all factors are significant at 1%, 5%, or 10% levels.

MATLAB ver2018 was used to implement the statistical regression models.

3. Results

3.1. Temporal and Spatial Patterns of CDVI

The SPI-based spatial and temporal distribution of yearly drought exposure during 1990 and 2012 showed many drought events in all SSA countries with higher severity from 1990 to 1999 as compared to the period 2000–2012 (Figure 1). Between 1990 and 1995, almost all countries experienced drought events. The most severe intensity in Southern Africa occurred in 1992, in Central Africa during 1994–1995, and in Western Africa in 1990. From 1996 to 2012, less severe drought events were recorded. The severe droughts were only observed in countries of Western Africa in 2002 and Eastern Africa in 2004. Other countries mostly experienced mild to moderate droughts (Figure 1).
Figure 1. Country-level spatial distribution of yearly Standardized Precipitation Index (SPI) during 1990–2012; the classification is based on Table 1 definition.

CDVI calculated based on the difference between actual and expected yield identified sensitive and resilient countries during 1990 and 2012 (Figure 2). The expected yield was obtained from the crop model, only reflecting climate-induced stresses. By contrast, the recorded yield showed the influence of actual climate as well as socio-economic conditions. Countries with high vulnerability had negative CDVI, whereas resilient ones exhibited positive values. The yearly national level CDVI showed many severe to moderate intensities of vulnerability during 1990–2012.

Figure 2. Country-level spatial distribution of yearly CDVI (Crop Drought Vulnerability Index) calculated based on the difference between the actual and expected yields; the classification is based on Table 1 definitions.
During 1990–2012, CDVI showed similar temporal trends as SPIs, as both indices showed more severe drought and higher vulnerability during 1990–1999 as compared to 2000–2012. However, the two indices displayed different pictures in terms of annual severity, revealing that different countries had varying resiliency during drought events. The Southern African countries were less affected by extreme CDVI, especially after 2000, indicating higher resiliency. For example, the moderate drought in 2012 in South Africa (Figure 1) did not cause the same level of vulnerability (Figure 2). The 2011 drought, which occurred in most SSA countries (Figure 1), had different vulnerabilities in different countries (Figure 2). An opposite situation occurred in 1999 when a moderate vulnerability in Western Africa was seen despite a non-drought situation.

Overall, based on SPI classification, Southern African countries, as well as Kenya, Tanzania, and Ethiopia from Eastern Africa, Mali, Niger, Nigeria, Botswana, Chad, and Central African Republic were more prone to severe droughts whereas Central African countries and Madagascar experienced less severe droughts due to the occurrence of higher rainfall (Figure 3a). However, CDVI showed a different picture as South Africa, Botswana, Mozambique, Nigeria, and Cameroon were less vulnerable (Figure 3b). The severity of vulnerability was lower than the severity of drought in countries such as Kenya, Tanzania, and Mali.

![Figure 3. Country-level comparison of (a) average SPI and (b) average CDVI during 1990 to 2012.](image)

### 3.2. Socio-Economic Factors Influencing CDVI

The bivariate correlation coefficient, i.e., the ‘r’ values between normalized variables were calculated to identify variables that are significantly correlated to each other (Step 3). Overall, nine socio-economic variables were selected (Table 3). In the economic category, the highest correlation coefficient was found between “GDP/capita” and “Interest payment” (r = 0.64), between “GDP/capita” and “GNI” (r = 0.83), and between “Interest payment” and “GNI” (r = 0.59) (Figure S1 in supplementary materials). This shows that these three variables have similar effects. Therefore, we selected “GDP/capita” and “Agricultural GDP (%GDP)” as the final variables from the economic category.

The human category showed the highest correlation coefficient (Figure S2 in supplementary materials). For example, we found r = 0.84 between “HDI” and “Health expenditure” and r = 0.73 between “Calorie intake” and “Food production index” (Figure S2 in supplementary materials). “HDI” was selected as the most representative variable, as it is a composite statistic of life expectancy, education, and per capita income and encompasses a more general aspect of human development. “Food production index” was selected due to the significant correlation coefficient with other explanatory variables from the human category. It was also an indicator of human nutrition status and a representative of human health aspects (definition in Table 2). While “HDI” and “Food production index” were also correlated, we selected both at this step and decided to check the suitability of one or
both for the regression model. We retained both variables in the resource and infrastructure categories (Water resource access (%) and Electricity access (%)) due to their relevance for drought vulnerability. In the governance category, the correlation coefficient between variables was larger than 0.75 (Figure S3 in supplementary materials). We only kept “Government effectiveness” as it is a more general indicator and encompasses political, rules, and regulatory aspects [40] (Table 2).

The nine remaining variables varied significantly from one region to the other in terms of their absolute values (Table 3). For example, variables such as “GDP/capita”, “HDI”, “Fertilizer use (t/ha)” were significantly larger in Southern SSA at the 5th, 50th, and 95th percentiles. “Agricultural GDP (%GDP)” was significantly larger in Central Africa, while variables such as “Food production index” showed approximately similar percentiles in the four SSA regions (Table 3).

Table 3. The nine socio-economic variables in the five categories (economic, human, resource, infrastructure, and governance) selected after the pairwise correlation coefficient analysis (Step 3) (units of each variable are shown in Table 1).

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>5th, 50th, 95th Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic</td>
<td>GDP/capita</td>
<td>Eastern SSA: 177, 330, 700</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Southern SSA: 696, 3025, 4455</td>
</tr>
<tr>
<td></td>
<td>Agriculture GDP</td>
<td>Central SSA: 224, 911, 6060</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Western SSA: 253, 551, 924</td>
</tr>
<tr>
<td>Human</td>
<td>HDI</td>
<td>Eastern SSA: 0.3, 0.39, 0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Southern SSA: 0.46, 0.6, 0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central SSA: 0.3, 0.4, 0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Western SSA: 0.26, 0.39, 0.49</td>
</tr>
<tr>
<td>resource</td>
<td>Agricultural area</td>
<td>Eastern SSA: 0.23, 0.99, 3.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Southern SSA: 0.53, 3.2, 5.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central SSA: 1.1, 2.2, 21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Western SSA: 0.43, 0.88, 12.3</td>
</tr>
<tr>
<td></td>
<td>Fertilizer use</td>
<td>Eastern SSA: 1.7, 5.7, 34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Southern SSA: 3.1, 7.8, 58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central SSA: 0.35, 3.5, 8.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Western SSA: 0.62, 7.8, 12.3</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Water resource access</td>
<td>Eastern SSA: 32, 55, 78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Southern SSA: 54, 79, 95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central SSA: 45, 63, 84.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Western SSA: 44, 63, 83</td>
</tr>
<tr>
<td></td>
<td>Electricity access</td>
<td>Eastern SSA: 0.9, 2.6, 16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Southern SSA: 2.0, 17.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central SSA: 1.1, 13.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Western SSA: 1.8, 2.5</td>
</tr>
<tr>
<td>Governance</td>
<td>Government effectiveness</td>
<td>Eastern SSA: −1.4, −0.7, −0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Southern SSA: −0.7, 0.2, 0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central SSA: −1.8, −1.2, −0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Western SSA: −1.3, −0.8, −0.1</td>
</tr>
</tbody>
</table>

The temporal variability of the nine selected and normalized variables showed different patterns for different variables and countries. For “GDP/capita”, almost all SSA countries, except Zimbabwe, Gambia, and Guinea, showed an increasing trend with a rather steep slope after 2000 (Figure 4). “Agricultural GDP (%GDP)” exhibited different temporal variability across different countries. Both “HDI” and “Food production index” showed an increasing trend all over SSA except Zimbabwe for “HDI” and Democratic Republic of Congo (DRC) for “Food production index”. The data for temporal variability of “fertilizer use (t/ha)” was not available for the period 1990–2000 and we used the average of the years 2001–2012. From the two infrastructure variables, “Water resource access (%)” showed similar trends and values in all countries except Zimbabwe and Sudan, but “Electricity access (%)” was more variable across countries. For “Government effectiveness”, the large variability between countries masked any temporal trends (Figure 4).

To see the spatial variability of selected socio-economic variables across countries in each category, we aggregated variables of each category into one indicator by calculating the average of their normalized values (Step 4). In all five categories, Southern African countries had higher values (higher adaptive capacity). In the economic category (Figure 5a), Western and Eastern SSA showed approximately similar lower values. Other countries like Zimbabwe, Zambia, and Angola were placed in between with aggregated values between −0.35 and 0. The human aspect showed low values in Ethiopia, Angola, Niger, and Mali followed by most Western African countries especially Angola and Namibia (Figure 5b). In the resource category, mostly Central Africa showed the lowest values (between −0.85 and −0.35) (Figure 5c). The infrastructure showed a different picture with mostly low values for most Central and Eastern African countries (Figure 5d). All SSA countries showed very poor capacity in terms of Governance, with the exception of Botswana and South Africa (Figure 5e).
Sustainability 2019, 11, x FOR PEER REVIEW 10 of 18

Figure 4. Temporal variability of the nine selected socio-economic variables during 1990 and 2012; the variables are normalized using Z-score in SSA (DRC: Democratic Republic of Congo and RC: Republic of Congo). Each scatter point is representative of a specific country. The black line highlights countries with different temporal variability.

Figure 5. The spatial variability of aggregated indicators in five categories: (a) economic, (b) human, (c) resource, (d) infrastructure, and (e) governance obtained from the average of selected normalized socio-economic variables in each category.
3.3. Relations between Time-Invariant Socio-Economic Variables and CDVI

We calculated the average of nine selected socio-economic variables (as explanatory variables) and CDVI (as dependent variables) during drought years, i.e., $SPI < 0$ (Step 5). All variables were fixed at this step, and a multiple variable regression model was used. The model was constructed (Model I) to check the suitability of the mean independent and dependent variables for model explanation and then to determine which of the nine selected socio-economic variables can explain the vulnerability of maize to drought. The relatively low $R^2$ value of 0.30 (Table 4) indicates that the time-invariant model is probably not sufficiently explanatory. Of the nine variables, only three including “Agricultural GDP (%GDP)”, “Food production index”, and “Electricity access (%)” were identified as statistically significant variables for reducing the vulnerability to drought. The highest $\beta$ value for “Agricultural GDP (%GDP)” with a value of $-0.72$ indicated that this variable was the most influential for reducing vulnerability followed by “Electricity access (%)” ($\beta = 0.34$) and “Food production index” ($\beta = 0.32$).

Table 4. The time-invariant socio-economic factors influencing maize drought vulnerability obtained from Model I. The averages of socio-economic variables and CDVI were used in the analysis. Only variables that were significant at 1%, 5%, or 10% levels were included in the model. The empty rows (-) pertain to those that were not significant (SE: standard error) (units of each variable are shown in Table 1).

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\beta$</th>
<th>SE</th>
<th>t-stat</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>$-1.25$</td>
<td>0.79</td>
<td>$-1.58$</td>
<td>0.9</td>
</tr>
<tr>
<td>GDP/capita</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Agriculture GDP</td>
<td>$-0.72$</td>
<td>0.25</td>
<td>$-2.87$</td>
<td>0.007</td>
</tr>
<tr>
<td>HDI</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Food production index</td>
<td>$0.32$</td>
<td>0.16</td>
<td>1.99</td>
<td>0.054</td>
</tr>
<tr>
<td>Agricultural land</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Fertilizer use</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Water resource access</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Electricity access</td>
<td>$0.34$</td>
<td>0.18</td>
<td>1.90</td>
<td>0.065</td>
</tr>
<tr>
<td>Government effectiveness</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

3.4. Assessing the Relationship between Time-Variant Socio-Economic Variables and CDVI

The temporal variability of socio-economic variables (as explanatory variables) and CDVI (as dependent variables) was examined during 1980–2012. As variables had both time and space (individual country) dimensions, panel regression models were used. We tested and compared two different models. In the first model (Model IIa), variables during drought and non-drought years (1980–2012) were included, while in the second model (Model IIb), only drought years were selected. The main reason for comparing these two models was to check if influential factors differ depending on climate stresses. Besides, drought has a long term effect, and the influence of some factors might be revealed in non-drought years. We believe that this comparison provides more detailed insights into the influencing factors of drought vulnerability.

We started with variables in the economic category i.e., “GDP/capita” and “Agricultural GDP (%GDP)” and then added variables of other categories one by one in a forward stepwise manner to check their significance. The comparison was made using the Likelihood Ration test. The results show significant improvement as the $R^2$ values were 0.63 in Model IIa and 0.60 in Model IIb (Table 5) over that of Model I ($R^2 = 0.3$). Model IIa slightly outperformed Model IIb. The SE values of the socio-economic variables were slightly lower compared to Model I in Table 4. This indicates that Models IIa,b are better able to identify factors that are important to maize vulnerability to drought. Including temporal variability of indicators is, therefore, critical to better characterize the effects of socio-economic factors on vulnerability mitigation.
A comparison of influential variables of Model I with Models IIa,b shows that more factors were significant in Models IIa,b. Apart from the three significant variables of Model I (“Agriculture GDP (%GDP)”, “Food production index”, and “Electricity access (%))”, “HDI”, “Agricultural land (ha/capita)”, and “Fertilizer use (t/ha)” were significant in Models IIa,b at 1%, 5%, or 10% levels. “Government effectiveness” was only significant in Model IIa, and “Agricultural land (ha/capital)” was only significant in Model IIb. Besides, the influential factors of Model I are significant at the 1% level (higher confidence), indicating their importance for vulnerability mitigation. All other variables were significant at the 5% level. In both Models, IIa,b, “Food production index” and “Electricity access (%)” showed higher β values (respectively, 1.94 and 2.787 in Model IIa; 2.01 and 3.101 in Model IIb).

Comparison of Model IIa with Model IIb showed that “Government effectiveness” was only statistically significant in Model IIa when all years were included. The β value for “Fertilizer use (t/ha)” was slightly higher in Model IIb (with a value of 0.141) compared with Model IIa (with a value of 0.132), indicating that as drought is occurring, the influence of fertilizer application is becoming more critical as an adaptation strategy. Besides, the β values of the human category, i.e., “HDI” and “Food production index” were higher, suggesting that their influence might be more critical during drought (Table 5). The two variables of “GDP/capita”, and “Water resource access (%))” were not significant in the three models. The country-level intercept (α) values obtained from two fixed-effect models (Model IIa and Model IIb) show different values for different countries and also in two models (Table S1).
4. Discussion

4.1. Changes in the Crop Drought Vulnerability

In this study, we assessed the relationship between maize drought vulnerability and socio-economic variables influencing vulnerability at the national level for SSA countries during 1990–2012. We identified significant socio-economic factors that predispose an area’s maize harvest to be resilient or sensitive to drought. The residuals of the simulated and observed yields were identified as prominent indicators of resilience concerning drought vulnerability. The response of a crop model adequately reflects the dynamic of climate on the rainfed maize yield. This is advantageous over studies that use third or fourth-order regression modeling to simulate yield [9,37], as such a de-trending procedure does not reflect the undeniable influence of climate variability on yield.

The overall spatiotemporal patterns of drought and vulnerability resemble the insights that different countries are coping differently with the occurring drought. An interesting observation is that maize vulnerability to drought has become less severe in recent years of our study period. For example, while the drought exposure is higher in Eastern and Southern African countries due to lower amounts of rainfall after 2000, the vulnerability of maize yield has declined. This corroborates the results of Naumann et al., [21], who concluded that these regions were less vulnerable. It suggests that there might be generic socio-economic factors that help mitigating vulnerability in different regions.

We should emphasize that the period of our analysis excluded extreme drought events that occurred in the late 1980s over the entire SSA. Including this period might significantly improve the assessment of changes in CDVI and factors influencing the trend. However, since most socio-economic variables are not available before the 1990s, this period was not included, which is a limitation of this study.

This paper presents the first comprehensive vulnerability assessment at a large scale by linking a physically-based spatially explicit crop model with indicators of adaptive capacity. The study is unique in the current literature, as most analyses on agriculture vulnerability have been case-specific and done at national or community levels [42–45]. Other large-scale analyzes lack the integration of a biophysical crop models with socio-economic factors [46,47]. Despite the challenges, complexity, and associated uncertainty, we made the best use of the existing data by taking into account both temporal and spatial dimensions of socio-economic components to have a broader understanding at the continental level.

4.2. Major Factors Influencing Drought Vulnerability

The 17 potential socio-economic variables in the five categories helped to disclose different aspects of adaptive capacities and the related potential for reducing vulnerability. While this classification is not ideal and some indicators may fall in more than one category, it gives more details on adaptive capacity of different factors. We also mapped the spatial and temporal dynamics of selected variables. Out of 17 variables, 9 remained after the ‘multicollinearity’ analysis from which 3–7 variables were significant in the regression models depending on drought and non-drought years included in the study. The following conclusions can be drawn for the variables in each of the five categories:

**Economic category:** “GDP/capita” was not a statistically significant factor. This is probably because the definition of HDI already encompasses the economic status of a country. “Agricultural GDP (%GDP)” as a more specific variable for agriculture, was associated with high vulnerability of maize in all models (Model I, Model IIa, and Model IIb) at 1% or 5% levels. The negative sign of β coefficient indicates that the large share of “Agricultural GDP (%GDP)” in total GDP can have a significant effect on reducing the ability to cope with drought. A strong economy secures the system by facilitating implementation of coping strategies against environmental risk and drought exposure. It also provides possibility for higher investment in weather forecasting, which may help farmers to better prepare for drought. A weak economy, often represented by the large share of agricultural sector in GDP, has the opposite effect [9,17,48].

**Human category:** “Food production index” was identified as a statistically significant factor in both time-variant and time-invariant models with a relatively high positive β coefficient. The index is...
a measure of food security, which is consistent with the situation in SSA where maize is one of the staple crops, and its vulnerability can significantly influence human health status in Africa. It is also representative of economic conditions of a country. The effect of “HDI” in reducing vulnerability becomes apparent when time-variant factors are included. This suggests that more investment in increasing life expectancy and education (as components of “HDI”) will reduce vulnerability. Bahadur Kc et al., [12] also showed that increasing “HDI” can result in an additional 6.8 million tons of maize production at the global level and the increase is more remarkable for developing countries such as Africa.

Resource category: Three variables were representative of the system’s natural resources from which only “Agricultural land (ha/capita)” and “Fertilizer use (t/ha)” had significant roles. Both variables were more influential during drought periods (higher β coefficients in Model IIa and Model IIb). This suggests that fertilizer is increasingly used to mitigate the impact of drought on maize. The lower β coefficient of “fertilizer use (t/ha)” compared to other variables might be related to missing values for the period 1990–2000 when most significant droughts occurred. We used the average of 2001–2012 in the model to substitute the missing data. Obtaining more accurate data for fertilizer application in SSA will help to better understand the potential benefits. The statistical significance of “Agricultural land (ha/capita)” indicates that the population growth together with the limited land resources will be a significant threat to maize-based food security in the future.

Infrastructure category: “Water resource access (%)” showed no statistical significance in all models. This might be due to the temporal variation of this variable (in Figure 4), which shows only a linear increasing trend with no significant variability across countries and inter-annually. By contrast, “Electricity access (%)” with more regional variability (Figure 4) and with high β coefficients in all models is a better representative.

Governance category: “Governance effectiveness” was also identified as a critical factor for reducing vulnerability in Model IIa. As mentioned by Keshavarz et al., [49], adaptation at a governmental level can help create a set of practical long-term plans and policies which may enhance the capacity to develop, revise, and execute drought policies. The “Governance effectiveness” was less significant in Model IIb because the missing values for the period of 1990–1995 (when extreme droughts occurred) were filled with constant mean values. Therefore, no significant variability was noticed in the variable “Governance effectiveness” during extreme drought years. Hence, a deeper implication of this variable requires data at higher spatial and temporal resolutions.

4.3. Comparison of the Models Explaining the Relationship between CDVI and Socio-Economic Variables

We tested the suitability of time-variant and time-invariant variables and concluded as expected that including the temporal dimension of variables was necessary for the determination of socio-economic factors influencing drought. Drought is a time-dependent phenomenon and is characterized by climate conditions for a specified period. Panel regression was the method of choice to evaluate the influence of socio-economic variables. As socio-economic variables differ significantly from one year to another, aggregating the severity of multiple drought years by calculating their average disguises the influence that a specific socio-economic variable might have in a specified year. Implementing panel data results in a more accurate inference of model parameters as these models usually have more degrees of freedom and more sample variability.

5. Conclusions and Limitation

Overall, our results underline the suitability of regression models for identifying how socio-economic factors influence drought-affected maize production between 1990–2012 in SSA. Despite the usefulness, some limitations to the data used in this study call for caution in the interpretation and further empirical efforts to improve data quality. Our spatial scale lacks details at the sub-national level as socio-economic data and maize yield were reported at the national level. As also mentioned by Conway et al., [50] and Simelton et al., [51], regional or gridded data could identify which
regions contribute most to national food insecurity. The World Bank databases reporting country level data were the only available sources with time series of socio-economic variables. Another limitation of this study is related to the lack of high-quality data on socio-economic variables. Official statistics in many SSA countries are not reliable and subject to criticism [52–54]. Therefore, there is still urgent need to invest in improving data quality.

There are also some other factors such as pest or disease that might have resulted in harvest loss and therefore increased crop drought vulnerability. However, we did not have access to these types of crop failure data. Such levels of information will depend on farm-level surveys. Another limitation is that the selected explanatory variables are not crop-specific or even agriculture-specific. Other factors, such as heat or cold spells, might have influenced crop vulnerability. However such specifications demand more work at the farm scale, which takes into account other drivers of vulnerability.

In conclusion, the current study was a preliminary but novel effort in identifying influential socio-economic factors on drought vulnerability across SSA. The results and the approaches developed can be used as a baseline study for further research to analyze crop drought vulnerability and its mitigation. As the quality and resolution of the data improve, a better understanding of the interaction of variables and their effects on drought vulnerability will be achieved by upgrading the calibrated crop model and also updating the analyses through inclusion of more recent years.

Supplementary Materials: The following are available online at http://www.mdpi.com/2071-1050/11/21/6135/s1, Figure S1: The visual representation of scatter plots generating correlation coefficient between normalized variables of the economic category; Figure S2: The visual representation of scatter plots generating correlation coefficient between normalized variables of the human category; Figure S3: The visual representation of scatter plots generating correlation coefficient between normalized proxies of the governance category; Table S1: The country level intercept (α) values obtained from two fixed effect models (Model IIa and Model IIb).

Author Contributions: Conceptualization, H.Y. and B.K.; methodology, B.K.; software, K.C.A.; validation, B.W., K.C.A. and B.K.; formal analysis, B.K.; investigation, B.K.; resources, H.Y.; data curation, B.K.; writing—original draft preparation, B.K.; writing—review and editing, B.K., B.W., H.Y. and K.C.A.; visualization, B.K.; supervision, B.W., H.Y. and K.C.A.; project administration, Swiss Federal Institute for Aquatic Science and Technology (Eawag); funding acquisition, H.Y.

Funding: Swiss National Science Foundation.

Acknowledgments: We are thankful to the Swiss National Science Foundation (SNF No.: CR21I3_146430. Dec.2013-Nov.2016) for financial support.

Conflicts of Interest: All co-authors of this manuscript declare that they have no conflict of interest.

References

10. Fraser, E.D.G.; Simelton, E.; Termansen, M.; Gosling, S.N.; South, A. “Vulnerability hotspots”: Integrating socio-economic and hydrological models to identify where cereal production may decline in the future due to climate change induced drought. *Agric. For. Meteorol.* **2013**, *170*, 195–205. [CrossRef]


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).