MAIZE YIELD SENSITIVITY TO CLIMATE VARIABILITY IN SOUTH AFRICA: APPLICATION OF THE ARDL-ECM APPROACH

Rangarirai Roy Shoko1, Petronella Chaminuka2, Abenet Belete1

1University of Limpopo, South Africa  
2Agricultural Research Council, South Africa

Abstract. Climate affects crop production decisions and outcomes in agriculture. From very short-term decisions about which crops to grow, when to plant or harvest a field, to longer-term decisions about farm investments, climate can positively or negatively affect agricultural systems. Although the general effects of climate change on agriculture are broadly understood, there are limited studies that model the relationship between specific crops and climate variables. The study uses the Autoregressive Distributed Lag (ARDL) model to analyze the sensitivity of maize yield to climate variables, fertilizer use and other non-climate variables. The paper uses annual time-series data of 47 observations spanning from 1970 to 2016. The results reveal that rainfall and temperature are important maize yield drivers in South Africa. However, if excessive, they will produce negative effects. The findings of this analysis are relevant for designing long-term interventions to mitigate the effects of climate change on maize production.

Keywords: maize, climate variability, ARDL model, cointegration

INTRODUCTION

Climate affects crop production decisions and outcomes in agriculture. From very short-term decisions about which crops to grow, when to plant or harvest a field, to longer-term decisions about farm investments, climate can positively or negatively affect agricultural systems. Although a natural phenomenon, climate variability has in recent years been exacerbated by anthropogenic interferences, resulting in higher incidences of extreme weather events such as drought, floods and increased temperatures. In most developing countries, including South Africa, the agricultural sector plays a key role in the economy, including by providing food and raw materials to industry, creating employment and being a source of foreign currency earnings. However, it primarily depends on climate variables. Changes in climate variables, particularly rainfall, affect agriculture in multifaceted ways, including changes in yield, reduced total output, changes in trade flows and increased food insecurity (BFAP, 2016). Although climate is accepted to be a primary determinant of agricultural productivity, the relationship between different climate variables and agricultural productivity is complex and uncertain (Adams et al., 1998; Estes et al., 2013).

In 2015, South Africa recorded the lowest national annual rainfall since 1904 (SAWS, 2016), which affected production of most agricultural commodities. Grain South Africa reported that wheat harvests suffered from the lowest rainfall in more than a century, falling to a four-year low in 2015, and an 18% drop from the 2014 yields. Similarly, maize yields were predicted to decline by about 11% in 2016 (BFAP, 2016), resulting in the country having to import maize for the first time since 2008.

1Rangarirai Roy Shoko, Department of Agricultural Economics, University of Limpopo, 0727, Sovenga, South Africa, e-mail: rshoko4@gmail.com, https://orcid.org/0000-0003-3934-8330
Although the general effects of climate change on agriculture are broadly understood, there are limited studies that model the relationship between specific crops and climate variables. Whilst yields are ultimately a result of interaction of several biophysical and socio-economic factors, the former are mostly beyond the control of a farmer or the government (Adams et al., 1998). South Africa is generally considered as a water scarce country with few areas being suitable for production of wheat (Estes et al., 2013). Increased understanding of the interaction between crop yields and climate variability, coupled with sensitivity analysis, in both the short term and the long term, can contribute towards better decision-making at both farm and policy levels. The use of econometric and mathematical models in analyzing the relationship between climate change and agricultural output can help make decisions regarding amendments to climate change policy (Simmelton et al., 2009) and design appropriate insurance packages for farmers.

The main objective of this study is to analyze the relationship between maize yield and climate variables in the South African agricultural sector. The study also incorporates non-climate variables such as fertilizer consumption, producer price of maize and agricultural policy in the analysis. Thus, the study seeks to test the hypothesis that maize yield is not affected by climate and non-climate variables. The paper focuses on the maize crop which is of strategic interest to South Africa in several ways, topmost of which are food security and agricultural trade balance. Secondly, the paper separately considers short-term and long-term relationships between climatic variables and yield, through the use of econometric techniques that have not previously been applied to the problem in South Africa. According to Adams et al. (1998), the commonly used approaches are controlled field experiments or crop simulation models. This paper offers an alternative approach by using data from observed national yields, climate and non-climate agribusiness organizations and the government.

DATA AND ANALYTICAL METHOD

Data and specification of variables
Annual time-series data of 47 observations spanning from 1970 to 2016 is used in this study. Data pertaining to maize yield measured in tons per hectare was extracted from the Abstract of Agricultural Statistics (2017) maintained by the Department of Agriculture Forestry and Fisheries (DAFF, 2016). In addition, average monthly rainfall figures measured in millimeters and average temperature figures measured in degrees Celsius were obtained from the South African weather services. Rainfall and temperature are major determinants of crop yields in rainfed agriculture (Blanc, 2013). Hence, to establish a climate/crop yield association, this study used time series for rainfall and temperature as proxy variables for climate variability. Rainfall is generally found to have a positive relationship with crop yields (Blanc, 2013; Amponsah et al., 2015).

In South Africa, the leading maize producing provinces are the Free State, Mpumalanga and Northwest (DAFF, 2016). In 2015, they contributed 80% to total maize output in South Africa. Therefore, average annual rainfall and temperature values were computed based on monthly rainfall and temperature figures from the three provinces. The values were recorded from 3 weather stations, namely Kroonstad, Potchefstroom and Komatidraai located in the Free State, Northwest and Mpumalanga provinces, respectively. Rather than total annual rainfall and temperature, the figures recorded during the production months are considered important. Hence, the average rainfall and temperature recorded in the six production months for maize (October, November, December, January, February and March) are used in the analysis. Time series values of fertilizer consumption (measured in metric tons) were also included in the analysis. Data on producer price of maize was obtained from the South African Grain Information Services (SAGIS). The variables used in the study were chosen based on economic theory and previous studies related to the subject. Other non-climatic variables such as crop variety selection and crop management were not considered due to lack of long-term annual data series.

Data analysis
The general relationship between the dependent variable (Maize yield) and its associated explanatory variables can be expressed in the form of a simple supply function specified as:

\[ My = f(RF, Tmp, Fcons, Mprice, Policy) \]  

where:

- \( My \) = supply variable measured by maize yield in tons/hectare
- \( RF \) = average annual rainfall variable measured in millimeters
The Autoregressive Distributed Lag (ARDL) model, also known as the bounds testing cointegration technique originally developed by Pesaran et al., 2001 was used to sufficiently assess the effects of climate variability and fertilizer use on maize yield. The model has been successfully used in various studies to estimate the relationship between crop yields and climate variables (Mbaga and Coyle, 2003; Janjua et al., 2014; Amponsah et al., 2015; Sarkodie and Owusu, 2016). The ARDL model provides an efficient platform for testing and estimating long-run relationships based on actual time series data (Hassler and Wolters, 2006) while also being perfectly suited for short time series (Duasa, 2010). Pesaran et al. (2001) suggested that the major advantage of ARDL is its flexibility in analyzing variables of different orders of integration. The cointegration test approach based on Johansen (1991) necessitates that all the variables be integrated of the same order, i.e. I(1). Hence, this approach is not suitable and cannot be applied in this study. The general function of a simple ARDL (1,1) model is specified as:

\[ Y_t = \delta + \theta Y_{t-1} + \Theta_0 X_t + \Theta_1 X_{t-1} + \varepsilon_t \quad (2) \]

\[ T = 1, 2, \ldots, T \]

\[ \varepsilon_t \sim i.i.d \{0, \sigma\}. \]

The model is autoregressive because the lagged values of the dependent variable \( Y_t \) partially explain themselves. A distributed lag component is present in the form of successive lags of the explanatory variable \( X_t \). The model was used in this study to determine the short- and long-run price elasticities for each selected individual gain crop in South Africa. Hence, the model was applied in two steps. The existence of a long-run relationship amongst the variables was determined in the first step whereas the second step consisted in estimating the model’s short-term and long-term coefficients.

The ARDL model used in this study to measure the long-run relationship among the variables is specified as:

\[ \ln M_{yt} = \alpha_0 + \sum_{i=1}^{\kappa} \alpha_i \ln M_{yt-i} + \sum_{i=1}^{\kappa} \beta_i \ln RF_{t-i} + \sum_{i=1}^{\kappa} \gamma_i \ln Tmp_{t-i} + \sum_{i=1}^{\kappa} \delta_i \ln Fcons_{t-i} \quad (3) \]

If the variables are cointegrated, then there exists an error correction representation. The short-run coefficients were estimated by the following ARDL error correction model (ECM):

\[ \ln M_{yt} = \alpha_0 + \sum_{i=1}^{\kappa} \alpha_i \Delta \ln M_{yt-i} + \sum_{i=1}^{\kappa} \beta_i \Delta \ln RF_{t-i} + \sum_{i=1}^{\kappa} \gamma_i \Delta \ln Tmp_{t-i} + \sum_{i=1}^{\kappa} \delta_i \Delta \ln Fcons_{t-i} + \sum_{i=1}^{\kappa} \omega_i \Delta Policy_i + \alpha_7 \Delta ECT + u_t \quad (4) \]

\[ \forall i = 1, 2, \ldots, k \]

where \( \alpha_7 \) represents the speed of adjustment (ECM term) which measures the deviation of \( M_{yt} \) from the long-run equilibrium level. All the variables except \( D \) were expressed in natural logarithms. The log transformation was employed to obtain a more homogeneous variance of a series (Luetkepohl and Xu, 2009). Eviews 10 econometric software was used to carry out the analysis, with the optimum lag lengths chosen based on the Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC).

**Unit roots test**

One of the requirements of the ARDL method is that none of the variables be integrated of order 2. The model collapses in the presence of any variable of second order (Granger and Newbold, 1974). Hence, the Augmented Dickey Fuller (ADF) and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) unit root tests were applied to check for the order of integration.

**Diagnostic checks**

The consequences of model misspecification in regression analysis can be severe in terms of the adverse effects on the sampling properties of both estimators and tests (Greene, 2017). Thus, to validate the goodness of
fit of the ARDL model, relevant diagnostic tests were applied, such as the Jarque-Bera test for normality, Breusch-Godfrey LM test for serial correlation and the White test for heteroskedasticity.

Stability tests
The Cumulative Sum (CUSUM) and Cumulative sum of squares (CUSUM squares) tests will be used to test for model stability. These tests have been used by several authors such as Janjua et al. (2014) to examine whether the parameters of a model are stable across various data sub-samples.

RESULTS AND DISCUSSION

Descriptive statistics
Understanding the properties of the variables involved in the analysis is an essential prerequisite for modeling time series data. Thus, various descriptive statistics (including mean, standard deviation, kurtosis, skewness, minimum and maximum) for all variables involved in the maize model are summarized in Table 1. On average, 2.5 tons per hectare of maize are produced at national level in South Africa. Temperatures and rainfall peak in summer, the time during which maize is produced (between October and March). For the sampled period, average annual rainfall and temperature are 89.81 millimeters and 13.39°C, respectively. The climate variables show relatively high standard deviation values, indicating that the data points are spread out over a large range of values. This is also an indication of high climate variability. The sample kurtosis and skewness values reflect non-normality in some of the variables. This was corrected by logarithmic transformation and first differencing.

Unit root test results
The results of the ADF and KPSS unit root tests, as presented in Table 2, show that all variables are non-stationary at level, except for rainfall and maize yield. As expected, all the non-stationary series became stationary after first difference.

The null hypothesis of unit root cannot be rejected at level since not all series are stationary at level. However, the hypothesis of unit root in all series was rejected at a 5% level of significance for all series after first difference. These results demonstrate that the series are integrated of order one, 1(1) and order zero, 1(0). Since none of the variables are integrated of second order difference, the ARDL model is estimated and a valid bounds test is applied.

ARDL model results
The short-run regression results of the ARDL model for the period 1970–2016 are presented in Table 3. Based on the minimum Akaike Information Criterion (AIC), the ARDL (1,1,1,0,0) model was considered to be the best fitted.

The value of the adjusted R² is 0.72, the F-statistic is 13.94 and is significant at 5% and 10%, which is acceptable to show overall fitness of the model.

Table 1. Summary statistics for variables used in the ARDL model (1970–2016)

| Statistical option | My  | RF  | Tmp | Fcons | Mprice |
|--------------------|-----|-----|-----|-------|--------|
| Mean               | 2.5204 | 89.8186 | 13.3941 | 378 201.5300 | 856.2289 |
| Standard deviation | 0.9678 | 18.0312 | 1.1363 | 71 217.9200 | 212.9310 |
| Kurtosis           | −0.4468 | −0.4796 | −0.1751 | 1.0164 | −1.1145 |
| Skewness           | 0.4263 | −0.0149 | 0.2658 | −0.9937 | 0.0166 |
| Range              | 4.0212 | 73.6690 | 5.0000 | 336 584.0000 | 837.3522 |
| Minimum            | 0.8180 | 55.0281 | 11.2000 | 180 685.0000 | 419.3103 |
| Maximum            | 4.8392 | 128.6971 | 16.2000 | 517 269.0000 | 1 256.6625 |

Source: own elaboration.
Diagnostic tests results

The validity of the ARDL model was confirmed through relevant diagnostic tests. The results of the tests, as presented in Table 4, strongly suggest that the model is adequate in terms of its specifications.

Stability test results

The Cumulative sum (CUSUM) and Cumulative sum of squares (CUSUM of squares) tests were applied to validate the stability of the ARDL approach. The results of the tests are presented in graphical form (see Fig. 2 and 3). The output shows that the CUSUM lines in both

---

**Table 2. Results of unit root tests at levels**

| Variables | ADF test statistics | KPSS test statistics | 5% significance level | 10% significance level |
|-----------|---------------------|----------------------|-----------------------|------------------------|
|           | Level               | 1st difference       | Level                 | 1st difference         |                         |
| LnMy      | –1.0878             | –7.4958              | 0.06157               | 0.119663               | I(0)                   | I(0)                   |
| LnFcons   | –3.7491             | –9.2532              | 0.513067              | 0.64837                | I(1)                   | I(0)                   |
| LnRF      | –5.9547             | –9.3789              | 0.06157               | 0.19025                | I(0)                   | I(0)                   |
| LnTmp     | –2.4562             | –8.5940              | 0.126720              | 0.159251               | I(1)                   | I(1)                   |
| LnMprice  | –4.6873             | –7.6966              | 0.47072               | 0.50000                | I(1)                   | I(1)                   |

Notes:
All variables include intercept and linear trend.
All variables are in logarithmic form.

---

**Table 3. ARDL (1,1,1,0,0) regression results**

| Variable    | Coefficient | Std. error | t-Statistic | Prob.* |
|-------------|-------------|------------|-------------|--------|
| Ln(My-1))   | 0.29300     | 0.11981    | 2.44563     | 0.0195**|
| Ln(Tmp)     | –0.41580    | 0.56510    | –0.73580    | 0.4666 |
| Ln(Tmp(-1)) | 0.81275     | 0.31930    | 2.54542     | 0.0153**|
| Ln(RF)      | 1.04708     | 0.14372    | 7.28553     | 0.0000**|
| Ln(RF(-1))  | –0.27907    | 0.19931    | –1.40019    | 0.1700 |
| Ln(Fcons)   | 0.42837     | 0.18552    | 2.30899     | 0.0268**|
| Ln(MPrice)  | 0.29195     | 0.16572    | 1.76173     | 0.0866* |
| Policy      | 0.21007     | 0.11886    | 1.76734     | 0.0856* |
| Constant    | –10.1847    | 2.67522    | –3.80707    | 0.0005**|
| Trend       | 0.00642     | 0.00354    | 1.81463     | 0.0779* |
| R-squared   | 0.777146    | Durbin-Watson stat | 2.0298  |
| Adjusted R-squared | 0.721433 | | |
| F-statistic | 13.94898   | | |
| Prob(F-statistic) | 0.00000 | | |

*Significant at 10% level. **Significant at 5% level.
Note: the dependent variable is maize yield (My).
Figures are positioned between the critical bound of 5% significance level over time, indicating that the model is largely stable.

After having confirmed the validity of the model, the bounds test was conducted to determine if there is a long-run relationship among the variables.

**Bounds test results**

One of the main purposes of estimating an ARDL model is to use it as a basis for the bounds test. The approach relies on the F-test statistic to determine the significance of lagged levels of the variables. From Table 5, the $F$-statistic value of 6.2025 is greater than the upper bound critical values at 1%, 2.5%, 5% and 10% levels, respectively. Accordingly, the null hypothesis of no long-run relationship is rejected. The results indicate that there is a long-run relationship among the estimated variables.

**Long-run elasticities of the ARDL model**

The presence of a long-run relationship among variables validates the estimation of an ARDL long-run model to obtain the long-run parameters. The results of

| Test statistic | Value | Level of significance | Lower bounds I(0) | Upper bounds I(1) |
|----------------|-------|-----------------------|-------------------|-------------------|
| $F$-statistic  | 6.2025| 10%                   | 2.75              | 3.79              |
| $k$            | 5     | 5%                    | 3.12              | 4.25              |
|                |       | 2.5%                  | 3.49              | 4.67              |
|                |       | 1%                    | 3.93              | 5.23              |

Null hypothesis: no long-run relationship.
the long-run model are presented in Table 6. The long-run coefficient of rainfall is positive and significant at 5%, suggesting that a 10% increase in rainfall (RF) will lead to an increase in maize yield by about 10.8% in the long run. The results are consistent with the findings by (Mwaura and Okoboi, 2014) who observed a positive long-run association between rainfall and crop yields in Uganda. Several empirical studies have also acknowledged the importance of the rainfall variable in explaining crop yield in rainfed agriculture (Blanc, 2012; Shakoor et al., 2015). In South Africa, maize production is largely rainfed and hence understanding the relationship between maize yield and rainfall is crucial in supporting informed policy decision-making.

The long-run elasticity of maize yield with respect to fertilizer consumption is positive and significant at 5%, indicating that an increase in fertilizer use by 10% will lead to an increase in maize yield by about 6.06%. Fertilizers are an important output driver in the South African agricultural sector and play a pivotal role in explaining crop yields. Agreeably, Janjua et al. (2014) argued that in the long run, fertilizers enhance land fertility causing an increase in agricultural production. Hence, the results validate the importance of fertilizer use in explaining maize yield changes in the long run.

The coefficient of the maize price variable is positive and significant at 10%, indicating that a 10 percent increase in the price of maize will be followed by an increase in maize production of about 4.1 percent in the long run. Maize prices play a pivotal role in stimulating crop production in South Africa. Economic theory suggests that farmers generally respond to high crop prices by adopting better farming methods which in turn improve crop yields. The long-run parameters obtained in this study are also comparable to Alemu et al. (2003) who recorded long-run price elasticities of 0.51 for maize in Ethiopia.

The coefficient of the policy variable is positive and significant at 5%. This finding suggests a positive long-run effect of the agricultural marketing policy on maize production in South Africa. The introduction of agricultural policies in the 1990s freed agricultural markets from stringent regulations, and this in turn created a free market system which allowed farmers to trade goods at competitive prices and have access to better farming technologies. This environment improved crop yields and enhanced food production in the country. The temperature variable is insignificant in the long-run, possibly suggesting that other variables (such as rainfall, fertilizer consumption) are better maize yield drivers than temperature in South Africa.

### Results of the error correction model

The presence of a long-run relationship among the variables validates the estimation of the error correction model to compute the short-run parameters. The results of the short-run dynamic error correction model (ECM), as presented in Table 7, indicate that in the short run, all variables are significant at 5% and 10%.

The estimated long-run elasticity of supply of maize to rainfall is 1.04. This suggest that a 10 percent increase in rainfall increases maize yield by 10.4% in the short run. The sign of the estimated coefficient is positive and in line with economic theory. Interestingly, the coefficient of the temperature variable is negative and significant at 10%. This finding suggests that an increase in temperature levels by 10% decreases maize yield by 4.1%. The finding agrees with Meng et al. (2016) who

### Table 6. ARDL model long-run results

| Variable | Coefficient | std. error | t-statistic | Prob. |
|----------|-------------|------------|-------------|-------|
| Ln(Tmp)  | 0.5615      | 1.1275     | 0.4980      | 0.6215|
| Ln(RF)   | 1.0863      | 0.3070     | 3.5387      | 0.0011*|
| Ln(FC)   | 0.6059      | 0.2776     | 2.1823      | 0.0357*|
| Ln(MPrice)| 0.4129      | 0.2483     | 1.6628      | 0.1050**|
| Policy   | 0.2971      | 0.1395     | 2.1298      | 0.0401*|

*Significant at 10% level. **Significant at 5% level.

Note: the dependent variable is maize yield (My).
showed that the effects of temperature changes can differ between the crops, and that high temperatures can be harmful in some periods, but not necessarily in all stages from planting to harvest. The coefficient of the error correction term $\text{ECT}(-1)$ has a value of 0.7070 and is negative and statistically significant at 5%. It demonstrates that after a 10-percent shock to the system, the long-run equilibrium relationship of maize yield is quickly reestablished at the rate of about 70% per annum. The result demonstrates a fairly high adjustment process. Given the results obtained in this study, the null hypothesis stated earlier in the paper is therefore rejected, and the conclusion is that climate and non-climate variables affect maize yield in South Africa.

CONCLUSION

The study investigated whether the newly evolving risk of climate variability is influencing maize yields in South Africa. Annual time-series data of 47 observations spanning from 1970 to 2016 was used. The Autoregressive Distributed Lag (ARDL) model was applied to examine the impact of climate and non-climate variables on maize yield. The results reveal that rainfall and temperature are important maize yield drivers in South Africa. However, if excessive, they will produce negative effects. The study has also shown that non-climatic variables such as fertilizer and policy also significantly influence maize yields. Long-run results revealed that fertilizer and rainfall will play a huge role in improving and maintaining maize yields in the future. The findings of this analysis are relevant for designing long-term interventions to mitigate the effects of climate variables on the country’s staple crop. While government policies cannot affect natural conditions like rainfall and temperature, they can compensate for the negative impact of climate change by increasing investments in irrigation, promoting efficient use of water and encouraging the adoption of drought-resistant varieties. Also, improving the farmers’ access to seasonal weather information can be another tool of effective adaptation for climate variability.

SOURCE OF FINANCING

University of Limpopo, South Africa.

REFERENCES

Adams, R. M., Hurd, B. H., Lenhart, S., Leary, N. (1998). Effects of global climate change on agriculture: an interpretative review. Clim. Res., 11, 19–30.

Alemu, Z. G., Oosthuizen, K. Schalkwyk, H. D. V. (2003). Grain-Supply Response in Ethiopia: An Error-Correction Approach, Agrekon, 42(4), 389–404.

Amponsah, L., Hoggar, K. G., Asuamah, S. Y. (2015). Climate change and agriculture: modelling the impact of carbon dioxide emission on cereal yield in Ghana. Agric. Food Sci. Res., 2(2), 32–38.

Breitenbach, M. C., Fényes, T. I. (2000). Maize and wheat production trends in South Africa in a deregulated environment. Agrekon, 39(3), 292–312.

Table 7. Error correction model results

| Variable   | Coefficient | Std. error | t-Statistic | Prob.    |
|------------|-------------|------------|-------------|----------|
| Constant   | -10.18477   | 1.56149    | -6.5225     | 0.0000*  |
| Trend      | 0.00642     | 0.00238    | 2.6910      | 0.0107*  |
| DLn(Tmp)   | -0.41580    | 0.24992    | -1.6637     | 0.1049** |
| DLn(RF)    | 1.04708     | 0.11407    | 9.1790      | 0.0000*  |
| ECT(-1)*   | -0.7070     | 0.10860    | -6.5103     | 0.0000*  |
| Adjusted R-squared | 0.71954 | Durbin-Watson stat | 2.0298 |
| $F$-statistic | 29.8621    |            |             |          |
| Prob($F$-statistic) | 0.00000    |            |             |          |

*Significant at 10% level. **Significant at 5% level.

Note: The dependant variable is maize yield (My).
