*International Food and Agribusiness Management Review*
*Special Issue - Volume 19 Issue A, 2016*

# Targeting Drought-Tolerant Maize Varieties in Southern Africa: A Geospatial Crop Modeling Approach Using Big Data

Kindie Tesfaye[ⓘa], Kai Sonder[b], Jill Cairns[c], Cosmos Magorokosho[d], Amsal Tarekegn[e], Girma T. Kassie[f], Fite Getaneh[g], Tahirou Abdoulaye[h], Tsedeke Abate[i], and Olaf Erenstein[j]

[a, g] *International Maize and Wheat Improvement Center (CIMMYT), ILRI Campus, CMC Road, P.O. Box 5689, Addis Ababa, Ethiopia*

[b, j] *CIMMYT, El Batan, Mexico*

[c, d, e] *CIMMYT, Harare, Zimbabwe*

[f] *International Center for Agricultural Research in the Dry Areas (ICARDA), Addis Ababa, Ethiopia;*

[h] *International Institute of Tropical Agriculture (IITA),*

[i] *CIMMYT, Nairobi, Kenya*

## Abstract

Maize is a major staple food crop in southern Africa and stress tolerant improved varieties have the potential to increase productivity, enhance livelihoods and reduce food insecurity. This study uses *big data* in refining the geospatial targeting of new drought-tolerant (DT) maize varieties in Malawi, Mozambique, Zambia, and Zimbabwe. Results indicate that more than 1.0 million hectares (Mha) of maize in the study countries is exposed to a seasonal drought frequency exceeding 20% while an additional 1.6 Mha experience a drought occurrence of 10–20%. Spatial modeling indicates that new DT varieties could give a yield advantage of 5–40% over the commercial check variety across drought environments while crop management and input costs are kept equal. Results indicate a huge potential for DT maize seed production and marketing in the study countries. The study demonstrates how big data and analytical tools enhance the targeting and uptake of new agricultural technologies for boosting rural livelihoods, agribusiness development and food security in developing countries.

**Keywords:** big data, drought tolerance; geospatial analysis; maize; spatial crop modeling, targeting

[ⓘ]Corresponding author: Tel: +251.116.462324
Email: K. Tesfaye: K.tesfayefantaye@cgiar.org



 

## Introduction

Rain-fed agriculture produces much of the food consumed globally and provides for the livelihoods of rural communities across the developing world. It accounts for more than 95% of farmed land in sub-Saharan Africa (SSA) where the rural populace of predominantly resource-limited families still face poverty, hunger, food insecurity and malnutrition (Wani et al. 2009). Maize is the most important staple food crop in SSA where it is almost entirely grown under rain-fed systems which are dependent on increasingly erratic rainfall. In southern Africa, maize accounts for 77% of the cereal area and 84% of the production, and over 30% of the total calories and protein consumed (FAOSTAT 2015).

However, current maize production in SSA is not sufficient to meet the growing demand in most countries and yields remain among the lowest in the world (Ray et al. 2012) because of an array of biophysical and socioeconomic constraints (Shiferaw et al. 2011). Drought is one of the major constraints under rain-fed systems with an estimated 40% of SSA's maize area facing occasional drought stress causing a yield loss of 10–25%. Around 25% of the maize crop suffers frequent drought resulting in a loss of up to half the harvest (CIMMYT 2013a). In southern Africa, maize yields are typically low due largely to drought and low-N stress (Weber et al. 2012).

Enhancing the productivity of rain-fed agriculture is an important avenue in reducing poverty and food insecurity in rain-fed systems (Rockström and Barron 2007; Wani et al. 2009). For example, adoption of improved maize varieties increases productivity and reduces chronic and transitory food insecurity under rain-fed systems (Kassie et al. 2014). Thus, increasing the use of improved technologies has the potential to enhance the welfare and food security of poor households (Bezu et al. 2014; Kassie et al. 2014). Improved maize technologies have been developed, disseminated and made positive contributions to the livelihood of smallholder farmers in some African countries (e.g., Abate et al. 2015). However, increasing adoption among smallholder farmers in Africa remains a challenge, including for DT maize varieties (Fisher et al. 2015). One of the challenges for wider adoption is the lack of data and tools for targeting new technologies at scale. Targeting is defined here as a process of identifying where a particular technology is the most likely to be successful–i.e. pinpointing the technology geo-spatially to the most likely niches of success. Targeting does not ensure the technology will be adopted there, but it does provide an indication of a potential fit between technology supply and demand in a geo-spatial context; and it is closely associated with recommendation domains (Notenbaert et al. 2013; Tesfaye et al. 2015c). In the context of targeting, data generated from a few research stations and/or on-farm demonstration plots are often not representative enough to address spatial and socioeconomic heterogeneities across scales.

Lately, climate, soil, elevation, and vegetation data sets are widely available at different spatial scales supporting analyses that were much more difficult in the recent past (Hyman et al. 2013). Big data and predictive analytics can make a difference in the agricultural industry (Sabarina and Priya 2015). Crop improvement and adoption research and development efforts have already benefitted from advances in big data, computing technology, and crop modeling for targeting genotypes to diverse environments (Löffler et al. 2005; Hyman et al. 2013). Targeting of crop varieties using a combination of big data and analysis tools has generated interest from public and private seed companies who wish to verify the area of adaptation and the agronomic value of new varieties for planning proper seed marketing and advisory schemes (Annicchiarico 2002). Therefore, the objective of this study is to assess the potential of targeting new DT maize varieties in southern Africa based on adaptation and productivity gains of new DT maize varieties, and present policy implications for seed production planning, marketing, and/or adoption. The study employs geospatial analysis and crop modeling tools that handle high resolution gridded climate, soil and crop data. The study purely focuses on the prospective technology change of using seed of a new DT maize hybrid instead of the prevailing non-DT commercial hybrid seed in areas that already produce maize—keeping other inputs constant. The study, therefore, does not include other productivity enhancing or risk-reducing interventions (be it crop rotation, crop management, and/or input considerations) nor does it assess the general suitability for maize in the study regions or its comparative advantage. The study contributes to a growing field of targeting research to inform agricultural development opportunities–typically linked to specific technologies and agro-ecological characteristics (Homann-Kee Tui et al. 2013; Hyman et al. 2013; Notenbaert et al. 2013; Tesfaye et al. 2015c) and/or socio-economic characteristics (Erenstein et al. 2010; Lang et al. 2013).

## Methodology

*Study Region*

The study was conducted in four major maize-growing countries (Malawi, Mozambique, Zambia and Zimbabwe) in southern Africa. In these countries, maize stands out as the primary crop in terms of area, absolute yield levels, and staple source of food (both calorie and protein) for millions of households (Kassie et al. 2013). Maize production in the region is constrained by several biophysical and socioeconomic factors. Amongst the biophysical factors, drought stands out as the major challenge across the region (Kassie et al. 2012; Weber et al. 2012). The study area is comprised of six Maize Mega–Environments (MME): dry lowland, wet lowland, dry mid-altitude, wet lower mid-altitude, wet upper mid-altitude and highland. MMEs are areas with broadly similar environmental characteristics for maize production delineated using environmental factors (maximum temperature, rainfall, and soil pH) as explanatory factors in capturing genotype by environment interactions (Hodson et al. 2002).

*Dataset for Geospatial Drought-Frequency Analysis*

The frequency of drought occurrence in the maize-growing environments of the study countries during the main cropping season (October–April) was analyzed using a long-term (1960–1998) gridded (0.5 x 0.5 degrees) standardized precipitation index (SPI) calculated using the climate database of the University of East Angelia (UEA) (Mitchell and Jones 2005). The SPI values were downloaded from the online database of the International Research Institute for Climate and Society (IRI 2015). The SPI simply refers to the number of standard deviations that an observed cumulative precipitation deviates from the climatological average (Mckee et al. 1993). The focus of our analysis was on seasonal drought and hence the six–month SPI values used for the study were for the period from November to April, which is the main rainy season in southern Africa.

*Geospatial Drought-Frequency Analysis*

The SPI values can be classified into three wet (SPI $\geq$ 1), three dry (SPI $\leq$ -1) and one normal (1>SPI>-1) classes (Sienz et al. 2012). For simplicity and ease of presentation, the study focused on the frequency of drought occurrence rather than comparing drought severity.

 

Therefore, pixels with values of ≤ -1 were classified as drought years while those with values of > -1 were classified as non-drought years. The frequency analysis was done using the 'equal to frequency' tool in ArcGIS 10.2 software (http://www.esri.com). The tool evaluates the number of times a value in a set of rasters is equal to a reference value raster (drought or non-drought in this case) on a cell-by-cell basis. Therefore, for each cell location in the input reference value raster, the number of occurrences where a raster in the input list has an equal value is counted. This was then converted to percentage frequency that explains the probability of occurrence of a drought or non-drought year for each pixel. A geospatial analysis was used to map and calculate the areas under different drought frequencies (1–10%, 10–20%, 20–30%, and >30%) across the six MMEs.

*Spatial Crop Modeling*

A spatial crop-modeling framework that integrates climate, soil, crop and crop management data was used to assess the performance of new DT maize varieties across environments in southern Africa.

**Model Description**

The Cropping System Model (CSM) used for simulating maize yields was Crop Estimation through Resource and Environment Synthesis, CERES–maize (Jones and Kiniry 1986), which is embedded in the Decision Support System for Agrotechnology Transfer (DSSAT), Version 4.5 (Hoogenboom et al. 2010). CERES–maize is a process-based, management-oriented model that utilizes water, carbon, nitrogen and energy balance principles to simulate the growth and development of maize plants within an agricultural system. The model runs with a daily time step and simulates crop growth, development and yield of specific cultivars based on the effects of weather, soil characteristics and crop management practices (Jones et al. 2003).

**Genetic and Environmental Data for Model Calibration and Evaluation**

Five new DT maize hybrids (CZH0946, CZH0811, CZH0616, CZH0835, and CZH0837) which represent four different maturity groups (extra-early, early, medium and late maturing) and one commercial check hybrid (SC513) that is widely grown in the region were selected for the study. The new hybrids are developed for southern and eastern Africa through a rigorous breeding specifically for yield potential and yield stability in drought-prone environments (Cairns et al. 2013). The CERES-Maize model was calibrated and evaluated using long-term (2005–2011) field data collected from a network of DT maize experiments in southern Africa, particularly from Zimbabwe. Data on crop phenology, yield and crop management (including planting date, plant density, fertilization and irrigation) were obtained from the regional trials database of CIMMYT in Zimbabwe. The data from Chisumbanje (19.800 S, 32.867 E), Chiredzi (21.050 S, 31.667 E) and Harare (17.942 S, 31.090 E) stations were used for model calibration while the data from Kadoma (18.369 S, 30.042 E), Makoholi (19.783 S, 30.750 E), Matopos (20.565 S, 28.453 E) and Ratry Arnold Research Station (17.183 S, 31.103 E) were used for model evaluation. Soil profile data of experimental stations were taken from Nyamapfene (1991). Daily rainfall, maximum and minimum temperature and radiation data of the experimental stations were obtained from the respective research stations or nearby meteorological observatories. Estimated data was provided by National Aeronautics and Space Administration-Prediction of Worldwide Energy Resource

 

(NASA-POWER) (http://power.larc.nasa.gov/) were used whenever radiation data were missing or unavailable.

### Model Calibration and Evaluation

The maize model used for the study requires six genetic coefficients which govern the life cycle and reproductive growth of maize cultivars (Table 1). A stepwise iterative calibration procedure was followed whereby genetic coefficients which determine anthesis and physiological maturity dates (P1, P2, and P5) were adjusted in the first stage of the process, followed by those coefficients which affect yield (G2 and G3) using 38 variety-site-year datasets. Rooting profile and soil fertility factors were adjusted with G2 and G3 whenever necessary. Model evaluation was made using an independent dataset (up to 98 variety-site-years). The agreement between simulated and measured values during calibration and evaluation was assessed using root mean square error (RMSE) and index of agreement (d) (Willmott 1982).

### Data for Spatial Crop Modeling

The calibrated and evaluated model was then used to simulate the yield of newly-released DT and the commercial check maize varieties in the respective countries at a pixel (≈ 10 km x 10 km) level across the maize growing areas in the study countries (Figure 1).

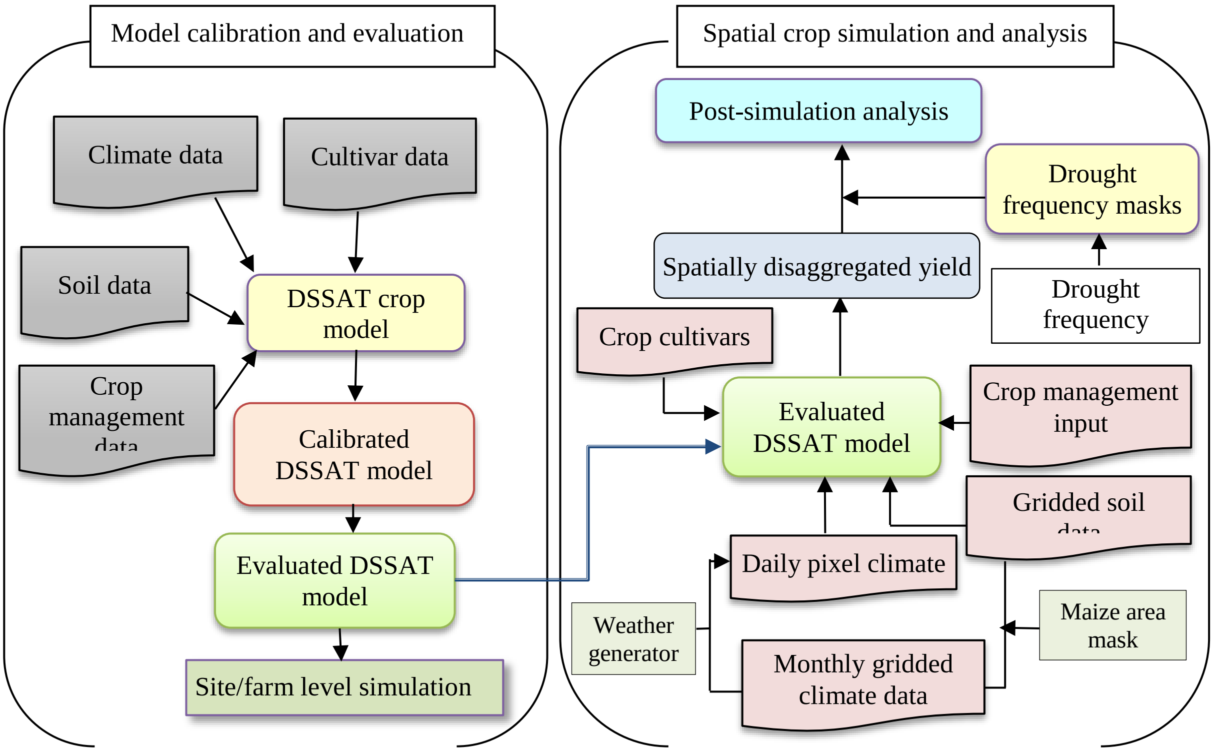


**Figure 1.** The process followed in crop model calibration, evaluation and spatial simulation.

The spatial simulations were made in a High-Performance Computing cluster (HPC) using gridded climate, soil and crop management data obtained from different online sources. The Spatial Allocation Model (SPAM) raster map for maize (You and Wood 2006) was used to select maize-growing areas in the study countries using the Geographic Resources Analysis

 

Support System (GRASS) software (http://grass.osgeo.org/). For each grid cell, soil inputs to the model were obtained from a set of twenty–seven generic soil profiles (HC27) developed by blending and interpreting information from both the Harmonized World Soil Database (HWSD) and the World Inventory of Soil Emission (WISE) database based on texture, rooting depth and organic carbon content (Batjes 2009). Simulations were run for all soils in each grid cell, and the cell-specific output was computed from the area-weighted average, based on the area share of each soil in the grid cell. Long-term climate data (1950-2000) for each simulation grid cell were obtained from the Worldclim gridded dataset (Hijmans et al. 2005) which provided all the required climatic elements needed by the stochastic daily weather generator in DSSAT.

A rule-based automatic planting was used to determine area-specific sowing date. The rule refers to a 70% soil moisture within 30-cm soil depth, monthly maximum temperature of <50 $^{o}$C and minimum temperature of >7 $^{o}$C within a 135-day planting window. The maize varieties were sown at a rate of 5.3 plants m$^{-2}$ and an average of 1000 kg ha$^{-1}$ crop residue was used as initial residue input to the model. All varieties were simulated with two equal split applications of 200 kg ha$^{-1}$ nitrogen. Details on spatial simulation of maize can be found in Tesfaye et al. 2015a.

### *Evaluation of Variety Performance and Seed Requirement Estimation*

The performance of the new DT varieties across the maize growing environments was measured by comparing their yield with the commercial check. Volume of seed required to cover an area of maize with a simulated yield advantage of at least 5% from any of the new DT varieties was determined by multiplying the area by the recent DT maize adoption rate reported for each country using an average seed rate of 25 kg ha$^{-1}$ (CIMMYT 2013b). The seed rate of maize (kg ha$^{-1}$) varies with the required plant population per hectare, seed weight, seed germination percentage and field loss (Macrobert et al. 2014). In Eastern and Southern Africa, 25 kg ha$^{-1}$ is mostly used as a recommended seed rate for maize (Langyintuo et al. 2008) for a target plant population of approximately 44,000–54,000 plants ha$^{-1}$ depending on the seed weight of varieties (Macrobert et al. 2014).

## Results

### *Drought Frequency*

Analysis of drought frequency indicates that all countries in southern Africa are prone to drought during the main cropping season (Figure 2). In the four study countries alone, more than 1.0 million hectares (Mha) of maize growing areas are exposed to seasonal drought events exceeding 20% while an additional 1.6 Mha experience a drought occurrence of 10–20%.

 

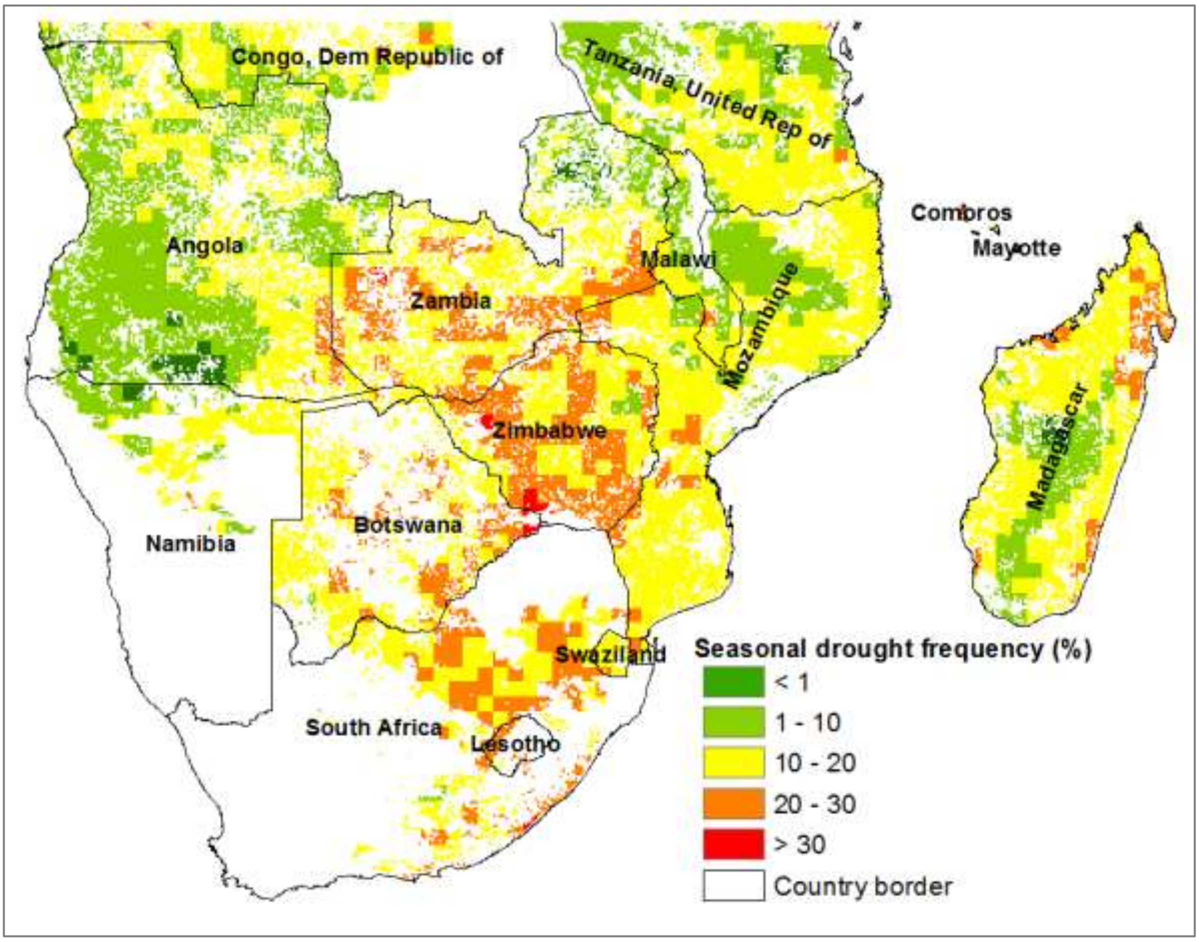


**Figure 2.** Prevalence of drought in the maize growing areas of southern Africa (1960-1998).

Maize area coverage and frequency of drought vary across MMEs in the study countries. The spatial distribution of maize area and drought frequency across countries and MMEs is presented in Figure 3 while the maize area under different drought frequencies across MMEs is summarized in Figure 4. Most of the maize area is found in the wet upper and wet lower mid-altitude MMEs in Malawi and Zambia, whereas it is located in the dry lowland, wet lowland and wet lower mid-altitude MMEs in Mozambique (Figures 3 and 4). Among the four study countries, Zimbabwe is the only country that has considerable maize area in the dry mid-altitude MME but has no maize area at all in the wet lowland MME. Although the maize area under the highland MMEs is extremely small in all countries, Malawi grows more maize in the highland MME than other countries (Figures 3 and 4). In terms of drought prevalence, Zimbabwe and Zambia are prone to more frequent drought events than that of Malawi and Mozambique across all MMEs (Figure 3). In Zimbabwe, most (>10%) of the seasonal droughts occur in the dry lowland, dry mid-altitude, wet lower mid-altitude and wet upper mid-altitude MMEs comprising a total maize area of 1.2 Mha . In Zambia, most of the maize areas (0.50 Mha) that are exposed to drought occurrences of 20% and above are located in the wet lower mid-altitude and wet upper mid-altitude MMEs (Figure 4). Most of the less frequent seasonal droughts (<15%) occur in the wet lower and wet upper mid-altitude MMEs in Malawi, in the wet lowland and wet lower mid-altitude MMEs in Mozambique and in the wet upper mid-altitude MME in Zambia (Figures 3 and 4).

 

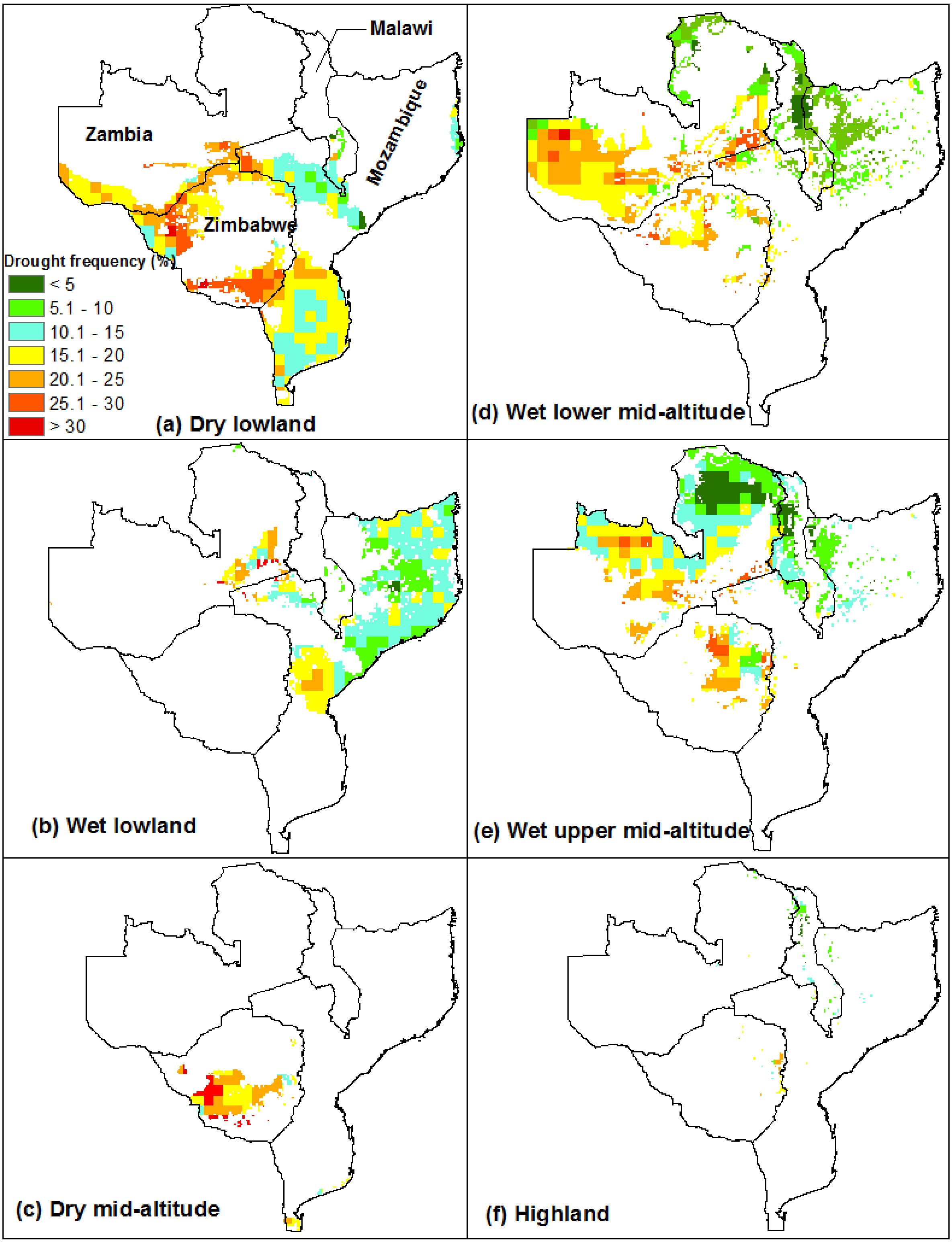


**Figure 3.** Prevalence of seasonal (November–April) drought (1960–1998) across six maize mega-environment in four southern Africa countries.

 

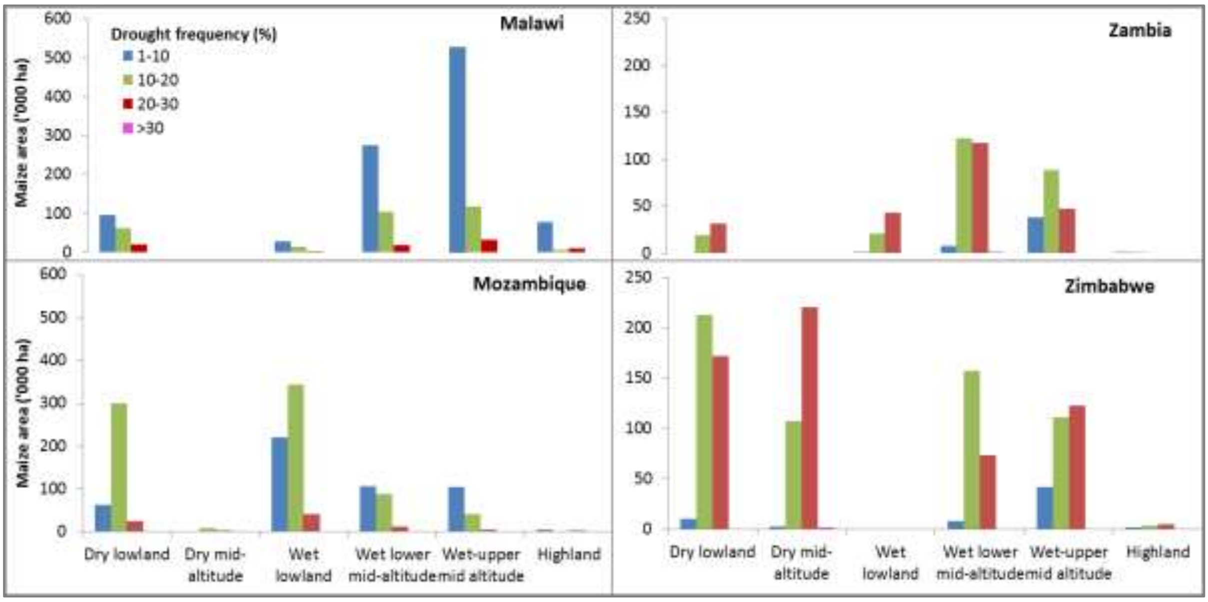


**Figure 4.** Seasonal drought frequencies across maize mega-environments in four southern Africa countries.

*Model Calibration and Evaluation*

A comparison of measured and simulated days to anthesis and maturity of the studied maize varieties showed good agreement between the measured and simulated values for both the calibration and evaluation datasets. The average RMSE of days to anthesis and maturity respectively was 4.2 and 7.7 days for the calibration dataset and 3.9 and 2.3 days for the evaluation dataset. The d-index values were 0.94 and 0.74 for days to anthesis and 0.67 and 0.95 for days to physiological maturity in the calibration and evaluation datasets, respectively (see Figure 1 for a plot of measured and simulated values). For grain yield, the average RMSE was 1.6 and 1.0 t $ha^{-1}$ for the calibration and evaluation datasets, respectively. The average simulated yield of the studied varieties across all site-years was closely related to measured grain yield with a d-index of >0.89 both in the calibration and evaluation datasets (see Figure 2 for a plot of measured and simulated grain yield). In general, the indices used for comparing the measured and simulated values of days to anthesis and physiological maturity and grain yield indicate that the CSM–CERES–maize model has captured the response of the DT maize varieties to different growing environments.

*Simulated Performance of DT Maize Varieties Across Environments*

The simulated relative yield performance of each of the new four DT varieties over that of the standard commercial check is shown in Figure 5. The simulated maize yield across different drought environments indicates that new DT varieties could give a yield advantage of 5% – 40% over the check variety (Figure 6). Although the performance of the new DT varieties varied across environments, they could give an average yield advantage of 16% and 12% under highly (>30% frequency) and less (<10% frequency) drought-prone environments, respectively. Specifically, new DT varieties give 11.4%, 12.9%, 13.6% and 14.7% higher yield than the check across environments with different drought frequencies in Malawi, Mozambique, Zambia and Zimbabwe, respectively. Average yield advantage among new DT varieties ranges from 5%-11% (CZH0946, CZH0811, and CZH0835), 15%-20% (CZH0616) and 28–40% (CZH0837). However, the new DT varieties do not beat the check universally (Figure 5). The coefficient of variation (CV) of yield showed that the new DT varieties could reduce annual yield variability by 3–7% as compared to the commercial check.

 

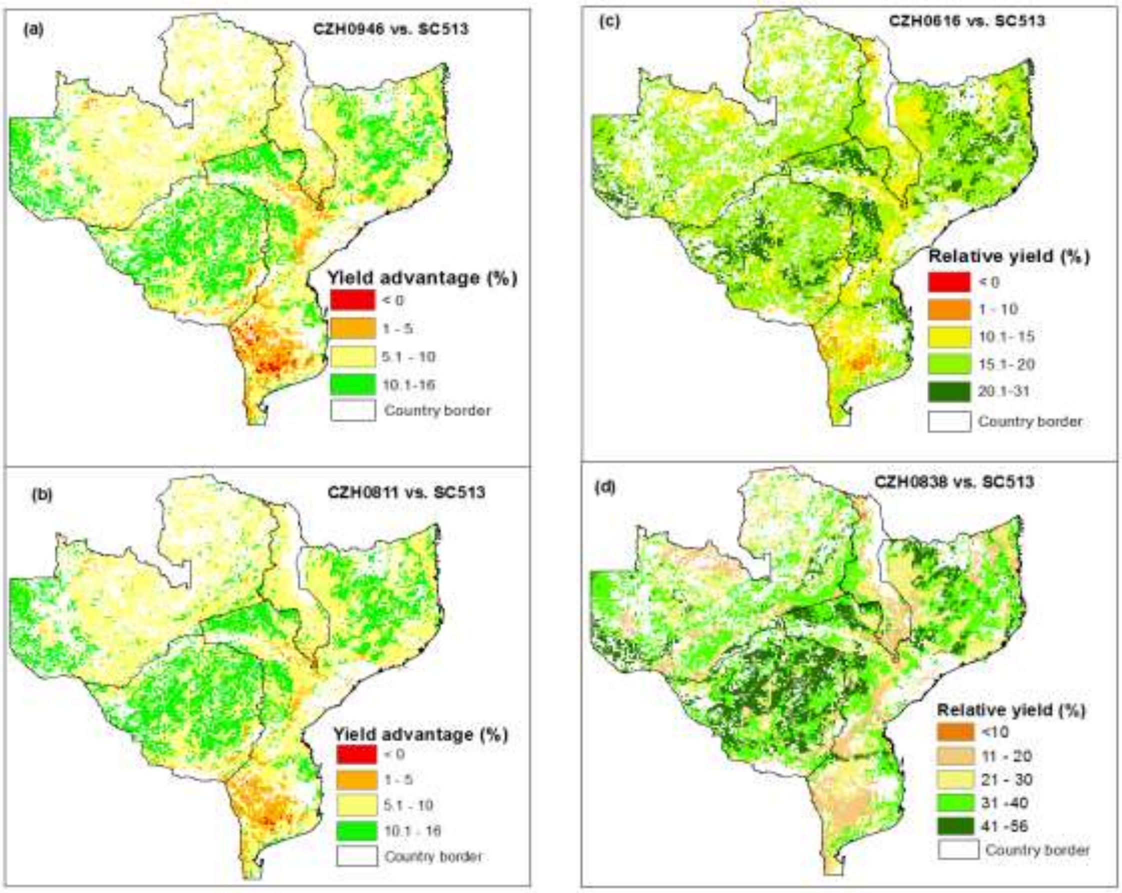


**Figure 5.** Spatial distribution of simulated relative yields of four new drought-tolerant varieties (a. extra early, b. early, c. medium and d. late maturity) compared to a commercial check (SC513) in four southern Africa countries.

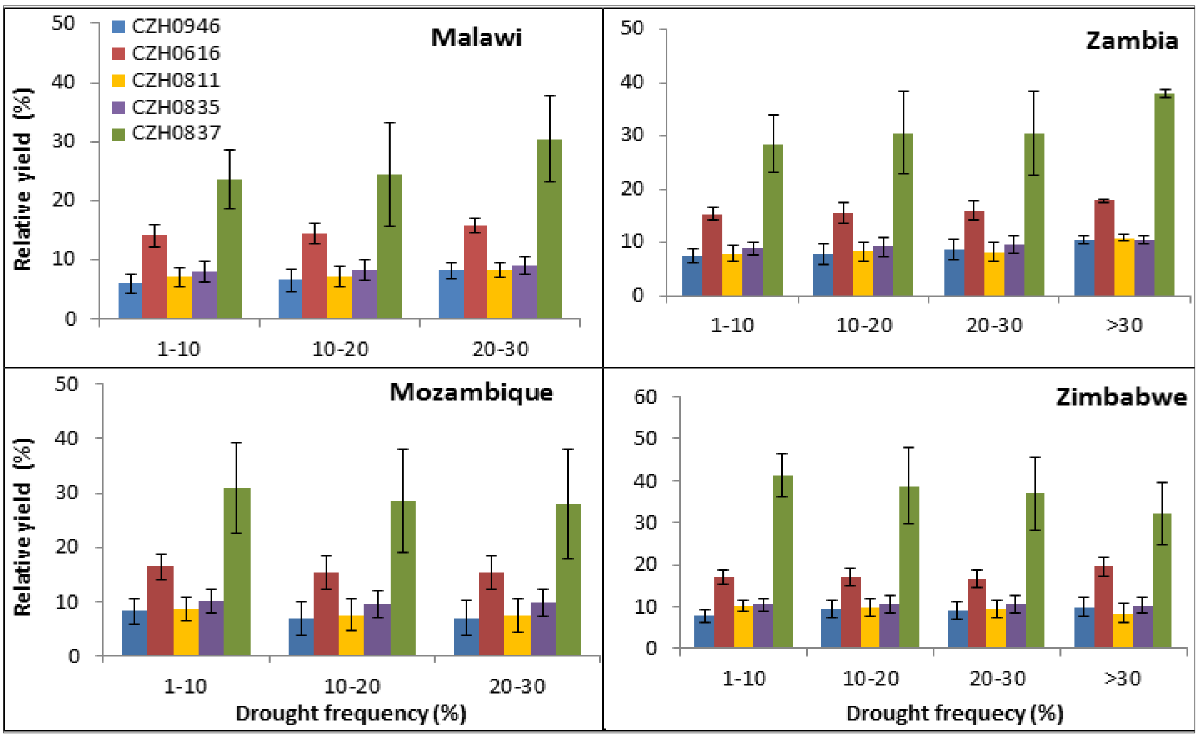


**Figure 6.** Simulated relative yield advantage and variance of five new drought-tolerant varieties over a commercial check (SC513) across different drought frequency environments in southern Africa. Vertical bars indicate standard deviations.

 

*Potential DT Maize Area and DT Seed Demand*

The potential DT maize area and DT seed demand were derived based on the simulated yield advantage (>5%) of the new DT varieties over the commercial check (Table 1). The results show DT maize to have substantial promise in terms of market opportunity for seed companies in the study countries. The level of adoption of new maize varieties varies among countries and so does the potential annual seed requirement: from 5,276 metric tons in Mozambique to 22,302 metric tons in Zimbabwe (Table 1).

**Table 1.** Potential DT maize area and DT seed demand

| Country | Potential DT maize area (ha pa)* | Current DT maize adoption rate (%)** | Potential DT seed demand (metric tons pa) | Current DT seed supply (metric tons pa)*** |
|---|---|---|---|---|
| Malawi | 1,387,790 | 47.3 | 16,411 | 4,416 |
| Mozambique | 1,366,799 | 15.4 | 5,276 | 855 |
| Zambia | 537,092 | 72.6 | 9,748 | 3,422 |
| Zimbabwe | 1,251,157 | 71.3 | 22,302 | 7,618 |

* Based on crop simulation, including all current maize area with a simulated yield advantage of >5% from new DT varieties over commercial check.
** **Source.** CIMMYT (2013b)
*** **Source.** Abate (2013).

## Discussion

The highly variable yield of rain-fed crops is the most important downside risk that farmers face in SSA essentially due to the uncertainty surrounding the frequency, intensity, and temporal and spatial distribution of drought (Kassie et al. 2012; Shiferaw et al. 2014). Understanding the nature of drought in a given area is the first step towards managing the risks associated with it (Kassie et al. 2012). Therefore, using long-term gridded data, this study identified the frequency and spatial distribution of seasonal drought during the main cropping season in the major maize growing countries in southern Africa. The results indicated that all the study countries are prone to drought despite variations in drought frequencies. Maize-growing areas in Zambia and Zimbabwe experience more frequent drought events than those in Malawi and Mozambique. The dry lowland and dry mid-altitude MMEs are generally prone to higher drought frequency than the rest of the MMEs, but the size of maize area affected by frequent drought within each MME varies among the study countries. Although all MMEs in Zimbabwe are prone to frequent droughts, the largest drought prone (≥20% frequency) maize area is found in the dry lowland and dry mid-altitude MMEs. In Zambia, however, the largest drought prone maize area is found in the wet lower mid-altitude MME. Therefore, the spatially explicit drought frequency maps generated in this study could be used to design appropriate drought risk management strategies in the respective countries such as targeting DT maize varieties.

Crop models have emerged as potential tools in agricultural research and development and in the exploration of management and policy decisions (Boote et al. 1996), and they have been used to assess spatial and temporal yield variability over different environmental conditions (Batchelor et al. 2002). However, the credibility of outputs of crop models depends on their calibration and evaluation within target environments (Timsina and Humphreys 2006; Xiong et al. 2008). In this study, the CERES–Maize model was calibrated and evaluated for selected DT maize varieties using measured data from a network of maize experiment stations in Zimbabwe. The evaluation results indicate that the model performed well in simulating the phenology and yield of maize after it is calibrated, and results agreed with previous studies that utilized field trial data from different environments to estimate maize genetic coefficients (Gungula et al. 2003; Yang et al. 2009).

This study provided a framework for evaluating the performance of new DT varieties across environments in southern Africa using geospatial analysis and spatial crop modeling tools that allow for an integrated analysis of big datasets (climate, soil, crop and management). Geospatial analysis tools play a valuable role in genotype targeting and can unravel genotype-by-environment interactions by providing high-resolution spatial and temporal data. Spatial analysis is key to identifying environmental frequencies and mapping out target environments that ultimately lead to a more effective deployment of germplasm (Hyman et al. 2013). As shown in this study and previous ones (Hyman et al. 2013; Tesfaye et al. 2015b), spatially explicit crop modeling takes into account changes in year-to-year environmental conditions across environments and could facilitate delivery of the right genotypes to farmers. Since crop varieties or genotypes could perform differently in different environments, a combination of crop simulation models and geographic information systems (GIS) are useful to understand the spatial and temporal aspects of genotype-by-environment interactions (Löffler et al. 2005). In this study, for example, the new DT varieties outperformed the commercial check variety across several environments, but they did not perform better than the check in all environments. Similarly, all new DT varieties did not perform the same way in the same environment, indicating the need for proper targeting of each variety.

Like other modeling studies (e.g., Challinor et al. 2009; Ruane et al. 2013), our study involved some important assumptions. Firstly, except for the varietal change—all other things were assume constant. Given the change of one hybrid seed for another at basically the same seed cost is a common practice in the study region; this appears to be a reasonable assumption. The seed change would not also initially trigger a different crop management practices given the stochastic nature of drought. Over time, however, one would expect farmers to realize the reduced risk inherent in DT maize and possibly adapt maize management practices that potentially increase DT maize benefits further. Secondly, our study focused only on sole maize cultivation and does not simulate other cropping systems such as crop rotation, intercropping or double cropping. Thirdly, the study assumed that plant nutrients other than nitrogen are applied or available in enough quantity so that they do not limit maize growth and development. Our interest in this study is on drought which is more difficult to manage than other crop management practices under rain-fed systems, and hence, our assumptions avoid confounding effects of other factors with drought. This indicates scope for future studies in addressing the assumptions made in this study.

The maps generated in this study show how the new DT varieties perform relative to the commercial check in different environments where maize is currently grown. The results reported in this simulation study are in agreement with previous studies that compared the performance of new DT varieties with commercial checks using field experiments. For example, in less drought prone environments (environments with a yield of ≥3 t/ha), the best DT hybrids yielded 15–25% more than SC513 under on-farm trials in Southern Africa (Setimela et al. 2013). Under severe drought stress environments, DT hybrids gave up to 40% yield advantage compared to commercially available hybrids in the farmers' fields (Setimela et al. 2012; Setimela et al. 2013). Moreover, the field experiments indicated that the best new DT hybrids out-yielded the farmers' own varieties by an average of 35% and 25% under high and low drought conditions in southern Africa, respectively (Setimela et al. 2013). In general, the yield gap between the commercial and the new DT varieties is higher under stressful conditions than non-stressed ones (Bänziger et al. 2006; Edmeades 2013; Setimela et al. 2013) indicating that more progress has been made in developing varieties for drought conditions compared to optimum environmental conditions.

The results of this study also do shed light on the location and volume of potential demand for DT seed and, therefore, could help boost the dissemination of varieties to the farmers that need them. Targeting of new genotypes is not only important to farmers, but it is also critical for public and private seed companies for planning proper marketing and advisory schemes for their varieties (Annicchiarico 2002). The results from this study indicate that the potential annual DT seed volumes in areas where the new DT varieties outperform provide a substantial market opportunity in the four study countries. This helps identify market opportunities for seed companies in southern Africa where varietal replacement is still very slow. However, the potential annual seed volume varies among the countries due to differences in adoption rate; for example, Mozambique has a very large maize area where the new DT varieties could perform well but with relatively low seed requirement. This reiterates that technology adoption is not only dependent on the biophysical suitability of the technology itself but also on socio-economic, political, cultural and institutional factors that may be of equal or greater importance (Notenbaert et al. 2013). Therefore, this type of analysis not only helps seed companies to determine potential annual seed demand in high adoption areas but also to identify areas where adoption is low so that they will be able to plan for addressing the low adoption problems. The relevance of geospatial crop modeling in agribusiness can be further strengthened by integrating socioeconomic factors into the modeling framework (e.g. Tesfaye et al. 2015b).

## Conclusion

The availability of big data—soil, climate, elevation and crop distribution—keeps improving over time and there is a growing interest in analytical tools that enable users to handle such data for agricultural applications. This study used geospatial and crop-modeling tools to processes and analyze big datasets for the characterization of drought prevalence and evaluation of the performance of new DT varieties across environments in southern Africa. This type of analysis helps target new DT varieties where they perform well and benefit most and identifies market opportunities. Big data and analytical tools thus can improve the effectiveness of targeting and enhance the uptake of new agricultural technologies that are required in boosting rural livelihoods, agribusiness development and food security in developing countries.

## Acknowledgement

We acknowledge a financial support from the Drought Tolerant Maize for Africa (DTMA) project funded by the Bill and Melinda Gates Foundation, USA. We also acknowledge the comments received from anonymous reviewers and the editor which helped improve the quality of the paper. The views expressed here are those of the authors and do not necessarily reflect the views of the authors' institutions or donor. The usual disclaimer applies.

 

## References


Abate, T. 2013. DTMA III Highlights for 2012/13: An Overview. Presented at the Drought Tolerant Maize for Africa (DMA) Annual Meeting. September, Nairobi, Kenya.

Abate, T. B. Shiferaw, A. Menkir, D. Wegary, Y. Kebede, K. Tesfaye, M. Kassie, G. Bogale, B. Tadesse, and T. Keno. 2015. Factors that transformed maize productivity in Ethiopia. *Food Security* 7(5): 965–981.

Annicchiarico, P. 2002. *Genotype x Environment Interactions: Challenges and Opportunities for Plant Breeding and Cultivar Recommendations*. 174. Food and Agriculture Organization, Rome, Italy. http://www.fao.org/docrep/005/y4391e/y4391e00.HTM.

Bänziger, M., P.S. Setimela, D. Hodson, and B. Vivek. 2006. Breeding for improved abiotic stress tolerance in maize adapted to southern Africa. *Agricultural Water Management* 80(1): 212–224.

Batchelor, W. D., B. Basso, and J. O. Paz. 2002. Examples of strategies to analyze spatial and temporal yield variability using crop models. *European Journal of Agronomy* 18:141–158.

Batjes, N. H. 2009. Harmonized soil profile data for applications at global and continental scales: updates to the WISE database. *Soil Use and Management* 25(2):124–127.

Bezu, S., G. T. Kassie, B. Shiferaw, and J. Ricker-Gilbert. 2014. Impact of improved maize adoption on welfare of farm households in Malawi: A panel data analysis. *World Development* 59(C): 120–131.

Boote, K. J., J. W. Jones, and N. B. Pickering. 1996. Potential uses and imitations of crop models. *Agronomy Journal* 88: 704–716. doi:10.2134/agronj1996.00021962008800050005x.

Cairns, J.E., J. Crossa, P.H. Zaidi, P. Grudloyma, C. Sanchez, J.L. Araus, S. Thaitad, D. Makumbi, C. Magorokosho, M. Bänziger, A. Menkir, S. Hearne, and G.N. Atlin. 2013. Identification of drought, heat, and combined drought and heat tolerant donors in maize. *Crop Science* 53(4): 1335–1346.

Challinor, A.J., F. Ewert, S. Arnold, E. Simelton, and E. Fraser. 2009. Crops and climate change: progress, trends, and challenges in simulating impacts and informing adaptation. *Journal of Experimental Botany* 60(10): 2775–89.

CIMMYT. 2013a. The Drought Tolerant Maize for Africa project. DTMA Brief, September. http://dtma.cimmyt.org/index.php/about/background [accessed October 14, 2015].

CIMMYT. 2013b. Country level improved maize adoption monitoring surveys. International Maize and Wheat Improvement Center, Mexico, D.F.

Edmeades, G.O. 2013. Progress in Achieving and Delivering Drought Tolerance in Maize - An Update. *International Service for the Acquisition of Agri-biotech Applications* 1–44. http://www.isaaa.org [accessed May14, 2016].


 

Erenstein, O., J. Hellin, and P. Chandna. 2010. Poverty mapping based on livelihood assets: A meso-level application in the Indo-Gangetic Plains, India. *Applied Geography* 30(1): 112–125.

FAOSTAT. 2015. FAO statistical databases. Food and Agriculture Organization of the United Nations. Statistics Division. Rome, Italy. http://faostat3.fao.org/home/E.

Fisher, M., T. Abate, R. W. Lunduka, W. Asnake, Y. Alemayehu, and R. B. Madulu. 2015. Drought tolerant maize for farmer adaptation to drought in sub-Saharan Africa: Determinants of adoption in eastern and southern Africa. *Climatic Change* 133(2): 283–299.

Gungula, D. T., J. G. Kling, and A. O. Togun. 2003. CERES-maize predictions of maize phenology under nitrogen-stressed conditions in Nigeria. *Agronomy Journal* 95(4): 892–899. doi:10.2134/agronj2003.8920.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25(15): 1965–1978. doi:10.1002/joc.1276.

Hodson, D. P., E. Martínez-Romero, J. W. White, J. D. Corbett, and M. Bänziger. 2002. Maize Research Atlas, version 3.0. Available from http://www.cimmyt.org/en/about-us/job-opportunities/1441-k12255-germplasm-data-coordinator-kenya [accessed October 15, 2015].

Homann-Kee Tui, S., M. Blümmel, D. Valbuena, A. Chirima, P. Masikati, A.F. van Rooyen, and G.T. Kassie. 2013. Assessing the potential of dual-purpose maize in southern Africa: A multi-level approach. *Field Crops Research* 153: 37–51.

Hoogenboom, G. et al. 2010. Decision Support System for Agrotechnology Transfer version 4.5 [CD-ROM]. University of Hawaii, Honolulu, HI.

Hyman, G., D. Hodson, and P. Jones. 2013. Spatial analysis to support geographic targeting of genotypes to environments. *Frontiers in Physiology* 4: 1-13.

IRI. 2015. Dataset: IRI Analyses SPI: Standardized Precipitation Index analyses of multiple global precipitation datasets. http://iridl.ldeo.columbia.edu/SOURCES/.IRI/.Analyses/.SPI/SOURCES/.IRI/.Analyses/.SPI/ [accessed April 19, 2015].

Jones, C. A., and J. R. Kiniry. 1986. CERES-Maize: A simulation model of maize growth and development. In *CERES Maize a Simulation Model of Maize Growth and Development,* edited by C. A. Jones and J. R. Kiniry. Texas A&M University Press, Texas.

Jones, J. W., G. Hoogenboom, C. H. Porter, K. J. Boote, W. D. Batchelor, L. A. Hunt, P. W. Wilkens, U. Singh, A. J. Gijsman, and J. T. Ritchie. 2003. The DSSAT cropping system model. *European Journal of Agronomy* 18: 235–265.

Kassie, G. T., O. Erenstein, W. Mwangi, R. Larovere, P. Setimela, and A. Langyintuo. 2012. Characterization of Maize Production in Southern Africa : Synthesis of CIMMYT / DTMA Household Level Farming System Surveys. 4. Mexico, D.F.

 

Kassie, G. T., R. La Rovere, W. Mwangi, O. Erenstein, A. Langyintuo, and K. Sonder. 2013. Drought Risk and Maize Production in Southern Africa. *Journal of Asian Scientific Research* 3: 956–973.

Kassie, M., M. Jaleta, and A. Mattei. 2014. Evaluating the impact of improved maize varieties on food security in Rural Tanzania: Evidence from a continuous treatment approach. *Food Security* 6: 217–230.

Lang, C., C.B. Barrett, and F. Naschold. 2013. Targeting Maps: An Asset-Based Approach to Geographic Targeting. *World Development* 41(1):232–244.

Langyintuo, A.S., A.O. Diallo, J.F. MacRobert, J. Dixon, and M. Bazinger. 2008. An analysis of the bottlenecks affecting the production and deployment of maize seed in Eastern and Southern Africa. International Maize and Wheat Improvement Center (CIMMYT), Mexico, D.F.: CIMMYT.

Löffler, C. M., J. Wei, T. Fast, J. Gogerty, S. Langton, M. Bergman, B. Merrill, and M. Cooper. 2005. Classification of maize environments using crop simulation and geographic information systems. *Crop Science* 45(5):1708. doi:10.2135/cropsci2004.0370.

Macrobert, J.F., P. Setimela, J. Gethi, and M.W. Regasa. 2014. Maize Hybrid Seed Production Manual. International Maize and Wheat Improvement Center (CIMMYT), Mexico, D.F.: CIMMYT.

Mckee, T. B., N. J. Doesken, and J. Kleist. 1993. The relationship of drought frequency and duration to time scales: Proceedings of the 8th Conference on Applied Climatology. 17(22) California.

Mitchell, T. D., and P. D. Jones. 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology* 25:693–712.

Notenbaert, A., M. Herrero, H. De Groote, L. You, E. Gonzalez-Estrada, and M. Blummel. 2013. Identifying recommendation domains for targeting dual-purpose maize-based interventions in crop-livestock systems in East Africa. *Land Use Policy 3*0: 834–846.

Nyamapfene, K. W. 1991. *The Soils of Zimbabwe*. Nehanda Publisher, Harare.

Ray, D. K., N. Ramankutty, N. D. Mueller, P. C. West, and J. A. Foley. 2012. Recent patterns of crop yield growth and stagnation. *Nature Communications* 3: 1293. doi: 10.1038/ncomms2296.

Rockström, J., and J. Barron. 2007. Water productivity in rainfed systems: overview of challenges and analysis of opportunities in water scarcity prone savannahs. *Irrigation Science* 25(3):299–311.

Ruane, A.C., L.D. Cecil, R.M. Horton, R. Gordón, R. McCollum, D. Brown, B. Killough, R. Goldberg, A.P. Greeley, and C. Rosenzweig. 2013. Climate change impact uncertainties for maize in Panama: Farm information, climate projections, and yield sensitivities. *Agricultural and Forest Meteorology* 170(March): 132–145.

 

Sabarina, K., and N. Priya. 2015. Lowering data dimensionality in big data for the benefit of precision agriculture. *Procedia Computer Science* 48: 548–554.

Shiferaw, B., B. M. Prasanna, J. Hellin, and M. Bänziger. 2011. Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. *Food Security* 3: 307–327.

Shiferaw, B., K. Tesfaye, M. Kassie, T. Abate, B. M. Prasanna, and A. Menkir. 2014. Managing vulnerability to drought and enhancing livelihood resilience in sub-Saharan Africa: Technological, institutional and policy options. *Weather and Climate Extremes* 3: 67–79.

Setimela, P., J. MacRobert, G. N. Atlin, C. Magorokosho, A. Tarekegne, D. Makumbi and G. Taye. 2012. Performance of elite maize varieties tested on-farm trials in eastern and southern Africa. Presented at the International Annual Meeting, American Society of Agronomy/ Crop Science Society of America/ Soil Science Society of America Meetings. October. Cincinnati, Ohio.

Setimela, P.S., G. Tesfahun, O. Erenstein, and O.F. Ndoro. 2013. Performance of Elite Drought Tolerant Maize Varieties Tested On-Farm in Eastern and Southern Africa. Presented at the International Annual Meeting, American Society of Agronomy/ Crop Science Society of America/ Soil Science Society of America Meetings. November. Tampa, Florida.

Sienz, F., O. Bothe, and K. Fraedrich. 2012. Monitoring and quantifying future climate projections of dryness and wetness extremes: SPI bias. *Hydrology and Earth System Sciences* 16: 2143–2157.

Tesfaye, K., S. Gbegbelegbe, J. E. Cairns, B. Shiferaw, B. M. Prasanna, K. Sonder, K. Boote, D. Makumbi, and R. Robertson. 2015a. Maize systems under climate change in sub-Saharan Africa. *International Journal of Climate Change Strategies and Management* 7(3): 247–271.

Tesfaye, K., M. Jaleta, P. Jena, and M. Mutenje. 2015b. Identifying potential recommendation domains for conservation agriculture in Ethiopia, Kenya, and Malawi. *Environmental Management* 55(2): 330-46. doi: 10.1007/s00267-014-0386-8.

Timsina, J., and E. Humphreys. 2006. Performance of CERES-Rice and CERES-Wheat models in rice–wheat systems: A review. *Agricultural Systems* 90:5–31.

Wani, S. P., T. K. Sreedevi, J. Rockstroma, and Y. S. Ramakrishna. 2009. Rainfed Agriculture – Past Trends and Future Prospects. In *Rainfed Agriculture: Unlocking the Potential*. 1–35. http://oar.icrisat.org/2428/ [accessed October 15, 2015].

Weber, V. S., A. E. Melchinger, C. Magorokosho, D. Makumbi, M. Bänziger, and G. N. Atlin. 2012. Efficiency of managed-stress screening of elite maize hybrids under drought and low nitrogen for yield under rainfed conditions in Southern Africa. *Crop Science* 52:1011. doi: 10.2135/cropsci2011.09.0486.

 

Xiong, W., I. Holman, D. Conway, E. Lin, and Y. Li. 2008. A crop model cross calibration for use in regional climate impacts studies. *Ecological Modelling* 213(3-4):365–380.

Yang, Z., G. G. Wilkerson, G. S. Buol, D. T. Bowman, and R. W. Heiniger. 2009. Estimating genetic coefficients for the CSM-CERES-Maize model in North Carolina environments. *Agronomy Journal* 101(5):1276-1285.

You, L., and S. R. Wood. 2006. An entropy approach to spatial disaggregation of agricultural production. *Agricultural Systems* 90:329–347. doi:10.1016/j.agsy.2006.01.008.

 