**Empirical approaches for assessing impacts of climate change on agriculture: The EcoCrop model and a case study with grain sorghum**



**Citation:**

Ramirez-Villegas J, Jarvis A, Läderach P. 2013. Empirical approaches for assessing impacts of climate change on agriculture: The EcoCrop model and a case study with grain sorghum. Agricultural and Forest Meteorology 170:67-78

**Publisher's DOI:**

https://dx.doi.org/10.1016/j.agrformet.2011.09.005

**Access through CIAT Research Online:**

http://hdl.handle.net/10568/42056

**Empirical Approaches for Assessing Impacts of Climate Change on Agriculture: The EcoCrop Model and a Case Study with Grain Sorghum**

**Authors:**

Julian Ramirez-Villegas
[1]International Center for Tropical Agriculture (CIAT), Cali, Colombia, AA6713
[2]Institute for Climatic and Atmospheric Sciences (ICAS), School of Earth and Environment, University of Leeds, Leeds, UK
[3]Challenge Program on Climate Change, Agriculture and Food Security (CCAFS), Copenhagen, Denmark
j.r.villegas@cgiar.org

Andy Jarvis
[1]International Center for Tropical Agriculture (CIAT), Cali, Colombia, AA6713
[3]Challenge Program on Climate Change, Agriculture and Food Security (CCAFS), Copenhagen, Denmark
a.jarvis@cgiar.org

Peter Läderach
[4]International Center for Tropical Agriculture (CIAT), Managua, Nicaragua
p.laderach@cgiar.org

**Target Journal**
Agricultural and Forest Meteorology (AFM)

**Abstract**

Climate has been changing in the last three decades and will continue changing regardless of any mitigation strategy. Agriculture is a climate-dependent activity and hence is highly sensitive to climatic changes and climate variability. Nevertheless, there is a knowledge gap when agricultural researchers intend to assess the production of minor crops for which data or models are not available. Therefore, we integrated the current expert knowledge reported in the FAO-EcoCrop database, with the basic mechanistic model (also named EcoCrop), originally developed by Hijmans et al. (2001). We further developed the model, providing calibration and evaluation procedures. To that aim, we used sorghum (*Sorghum bicolor* Moench) as a case study and both calibrated EcoCrop for the sorghum crop and analyzed the impacts of the SRES-A1B 2030s climate on sorghum climatic suitability. The model performed well, with a high true positive rate (TPR) and a low false negative rate (FNR) under present conditions when assessed against national and subnational agricultural statistics (min TPR=0.967, max FNR=0.026). The model predicted high sorghum climatic suitability in areas where it grows optimally and matched the sorghum geographic distribution fairly well. Negative impacts were predicted by 2030s. Vulnerabilities in countries where sorghum cultivation is already marginal are likely (with a high degree of certainty): the western Sahel region, southern Africa, northern India, and the western coast of India are particularly vulnerable. We highlight the considerable opportunity of using EcoCrop to assess global food security issues, broad climatic constraints and regional crop-suitability shifts in the context of climate change and the possibility of coupling it with other large-area approaches.

*Keywords: climatic suitability, modeling, impacts, adaptation, EcoCrop, sorghum*

## 1. Introduction

Climate has been changing in the last three decades and will continue changing regardless of any mitigation strategy (IPCC, 2001, 2007). By mid-21st century, temperatures are predicted to increase about 3-5ºC (depending on the greenhouse gas emission pathway, though with uncertainties in the climate system response), while precipitation patterns (in amount, seasonality, and intensity) are predicted to shift (Arnell et al., 2004; IPCC, 2007; Meehl et al., 2005). In all the world's economies, agriculture is amongst the most vulnerable of sectors to these changes in climate (Gregory et al., 2005; Jarvis et al., 2010; Thornton et al., 2011), it is the basis for food security and economic sustainability and provides the necessary input for sustaining people's livelihoods, regardless of their economic status (FAO, 2009, 2010c). In developing countries, agriculture is a key driver of national and local economies and the way households live largely depends on what they can grow and how efficiently they can do it.

Several authors report that agricultural production could suffer progressive yield loses in the next hundred years (Challinor et al., 2009, 2010; IPCC, 2007; Lobell et al., 2008; Thornton et al., 2011). While the recent successes in the climate negotiations are promising, it is still unknown how or to what extent the emissions cuts will affect global temperature rises. The effects of a +4ºC warmer world could be disastrous without adequately guided adaptation processes (Thornton et al., 2011). In particular, in the tropics and subtropics, current crop varieties of several crops would be unlikely to produce under extreme conditions (Byjesh et al., 2010; Challinor et al., 2005, 2010), since crop niches in these regions (Fuller, 2007), are highly sensitive to changes and variations in climates (Lane and Jarvis, 2007), and adaptation processes are likely to face numerous constraints (Thornton et al., 2011)

Despite that, there is still no consensus on the magnitude of climate change impacts on crop production, due in part to a lack of understanding of crop growth processes and in other part to a general lack of coordination among crop modelers. To date, more than one hundred crop models exist, and each makes different assumptions and holds different uncertainties (Challinor et al., 2009; Rivington and Koo, 2011). Filling in these information gaps and delivering this key information to guide adaptation processes in the field is not an easy task, particularly for underutilized or neglected crops. Because of the absence of precise methods to evaluate yield response to climate, there are still hundreds of regionally-relevant crops that have been poorly researched. For these crops, suitability indices have been used by several researchers as a proxy to evaluate the response of a variable (or mixture of variables) to a set of environmental factors (Lane and Jarvis, 2007; Nisar Ahamed et al., 2000; Schroth et al., 2009). These indices have been developed as proxies to quantify the relationship between climate and crop performance when no detailed information is available.

In this paper, we integrate the current expert-based ecological ranges data reported in the FAO-EcoCrop database (FAO, 2000) with the basic mechanistic model (also named EcoCrop) originally implemented in DIVA-GIS (Hijmans et al., 2001) to evaluate the likely impacts of climate change on agricultural production. We propose a modification of the original algorithm implemented by Hijmans et al. (2001) and use sorghum (*Sorghum bicolor* Moench) as a case study for developing our model. We choose sorghum on the basis of the crop's importance—it is an important and widely adapted small-grain cereal grown in the tropics and subtropics (Craufurd et al., 1999), ranking 6th globally in total harvested area after wheat, rice, maize, soybean, and barley (FAO, 2010b)—in addition to the availability of calibration and evaluation data. We use current detailed distribution of climates from WorldClim (Hijmans et al., 2005) along with a calibrated set of growing parameters and develop a set of metrics and specific calculations to determine current suitability on a geographic basis over Africa and South-east Asia. We then project the model using a set of 24 statistically downscaled Global Circulation Models (GCMs) for the SRES-A1B emissions scenario (Ramirez and Jarvis, 2010; Tabor and Williams, 2010; Wilby et al., 2009) by 2030s (2020-2049). Finally, we assess the impacts of climate change on sorghum climatic suitability, identify the main caveats and advantages of our approach, compare our results for different regions with the results of other studies, and assess and note the main model- and climate-driven uncertainties.

## 2. Materials and methods

The approach shown in this paper has mainly three different steps: (1) the first step includes the description of the model, its parameterization, and the description of input climatic datasets; (2) the second step involves the implementation of the model in a case study with sorghum in Africa and South Asia; and (3) the third step consists of the post-modeling calculations, and the description of the usage and interpretation of relevant metrics.

### 2.1. Model description

The basic mechanistic model (EcoCrop) we implemented uses environmental ranges as inputs to determine the main niche of a crop and then produces a suitability index as output. The model was originally developed by Hijmans et al. (2001) and named EcoCrop since it was based on the FAO-EcoCrop database (FAO, 2000).

In the model, there are two ecological ranges for a given crop, each one defined by a pair of parameters for each variable (i.e. temperature and rainfall). First, the absolute range, defined by $T_{MIN-C}$ and $T_{MAX-C}$ (minimum and maximum absolute temperatures at which the crop can grow, respectively) for temperature, and by $R_{MIN-C}$ and $R_{MAX-C}$ (minimum and maximum absolute rainfall at which the crop grows, respectively) for precipitation; and second, the optimum range, defined by $T_{OPMIN-C}$ and $T_{OPMAX-C}$ (minimum optimum and maximum optimum temperatures, respectively), and $R_{OPMIN-C}$ and $R_{OPMAX-C}$ (minimum optimum and maximum optimum rainfall, respectively). An additional temperature parameter is used ($T_{KILL}$) to illustrate the effect of a month's minimum temperature (explained below).

When the conditions over the growing season (i.e. temperature, rainfall) at a particular place are beyond the absolute thresholds there are no suitable conditions for the crop (white area, Figure 1A); when they are between absolute and optimum thresholds (dark grey area, Figure 1A) there are a range of suitability conditions (from 1 to 99), and whenever they are within the optimum conditions (light grey area, Figure 1 left) there are highly suitable conditions and the suitability score is 100%. The model performs two different calculations separately, one for precipitation and the other for temperatures and then calculates the interaction by multiplying them (Figure 1B).

**[INSERT FIGURE 1]**

The first parameter that requires definition is the duration of the crop's growing season ($G_{AVG}$, in months). For a given site ($P$), for each month ($i$) of the growing season and for each of the 12 potential growing seasons of the year (assuming each month is potentially the first month of the crop's growing season), the temperature suitability ($T_{SUIT}$) is calculated by comparing the different crop parameters with the climate data at that site (Eqn. 1)

$$T_{SUIT\,i} = \begin{cases} 0 & T_{MIN-Pi} < T_{KILL-M} \\ 0 & T_{MEAN-Pi} < T_{MIN-C} \\ a_{T1} + m_{T1} * T_{MEAN-Pi} & T_{MIN-C} \leq T_{MEAN-Pi} < T_{OPMIN-C} \\ 100 & T_{OPMIN-C} \leq T_{MEAN-Pi} < T_{OPMAX-C} \\ a_{T2} + m_{T2} * T_{MEAN-Pi} & T_{OPMAX-C} \leq T_{MEAN-Pi} < T_{MAX-C} \\ 0 & T_{MEAN-Pi} \geq T_{MAX-C} \end{cases} \qquad \text{[Eqn. 1]}$$

Where $T_{SUITi}$ is the temperature suitability index for the month $i$, $T_{MIN-C}$, $T_{OPMIN-C}$, $T_{OPMAX-C}$ and $T_{MAX-C}$ are defined on a crop basis, $a_{T1}$ and $m_{T1}$ are the intercept and slope (respectively) of the regression curve between [$T_{MIN-C}$, 0] and [$T_{OPMIN-C}$, 100], $a_{T2}$ and $m_{T2}$ are the intercept and slope (respectively) of the regression curve between [$T_{OPMAX-C}$, 100] and [$T_{MAX-C}$, 0]. $T_{MIN-Pi}$ is the minimum temperature of the month $i$ at the site $P$, $T_{MEAN-Pi}$ is the mean temperature of the month $i$, $T_{KILL-M}$ is the crop's killing temperature plus 4ºC. The model assumes that if the minimum temperature of the month in a particular place is below [$T_{KILL}$+4ºC], then the minimum absolute killing temperature will be reached in at least one day of the month, and the crop will freeze and fail. The final temperature suitability ($T_{SUIT}$) is the minimum value of all 12 potential growing seasons.

For precipitation, the calculation is done only once, using the crop's growing season total rainfall (sum of the rainfall in all the growing season's months), and using both the minimum, and maximum absolute and optimum crop's growing parameters (Eqn. 2)

$$R_{SUIT} = \begin{cases} 0 & R_{TOTAL-P} < R_{MIN-C} \\ a_{R1} + m_{R1} * R_{TOTAL-P} & R_{MIN-C} \leq R_{TOTAL-P} < R_{OPMIN-C} \\ 100 & R_{OPMIN-C} \leq R_{TOTAL-P} < R_{OPMAX-C} \\ a_{R2} + m_{R2} * R_{TOTAL-P} & R_{OPMAX-C} \leq R_{TOTAL-P} < R_{MAX-C} \\ 0 & R_{TOTAL-P} \geq R_{MAX-C} \end{cases} \qquad \text{[Eqn. 2]}$$

Where $R_{TOTAL-P}$ is the total rainfall of the crop's growing season at site $P$, $R_{SUIT}$ is the rainfall suitability score, the crop parameters ($R_{MIN-C}$, $R_{OPMIN-C}$, $R_{OPMAX-C}$ and $R_{MAX-C}$) are defined on a crop basis, $a_{R1}$ and $m_{R1}$ are the intercept and the slope of the regression curve between [$R_{MIN-C}$, 0] and [$R_{OPMIN-C}$, 100], and $a_{R2}$ and $m_{R2}$ are the intercept and the slope of the regression curve between [$R_{OPMAX-C}$, 100] and [$R_{MAX-C}$, 0]. Finally, the total suitability score is the product (multiplication) of the temperature and precipitation suitability surfaces calculated separately (Eqn. 3).

$$SUIT = R_{SUIT} * T_{SUIT} \qquad \text{[Eqn. 3]}$$

All the model parameters (i.e. $T_{KILL}$, $T_{MIN-C}$, $T_{OPMIN-C}$, $T_{OPMAX-C}$, $T_{MAX-C}$, $R_{MIN-C}$, $R_{OPMIN-C}$, $R_{OPMAX-C}$, $R_{MAX-C}$) are referred to as "crop ecological parameters" hereafter.

## 2.2. Model calibration

The process we call model calibration is the process of statistically finding the correct ecological parameters for the crop to be modeled, based on point-based crop presence and 20$^{th}$ century spatially explicit climatology data. We selected sorghum in Africa and South Asia as a case study for testing the parameter selection process. We selected these geographical areas because (1) they are of high relevance under the context of climate change and are predicted to receive severe negative impacts (IPCC, 2007), and (2) it has been the focus of several research programs up until now. Similarly, we selected sorghum for several reasons: (1) is an important crop for rural communities in developing countries in Africa and Asia (our study area), (2) there are enough data on it for the proposed calibration, (3) FAOSTAT ranks it 6$^{th}$ in area harvested, so it is very likely that there are ample national statistics for evaluation, and (4) it has been assessed in other studies related to climate change, allowing us to compare our results with these.

### 2.2.1. Present climate data

As the EcoCrop model is intended to be applied over a geographic domain rather than a single point, present climate data for model calibration needs to (1) have enough spatial coverage to permit analysis of the whole region of study, (2) have adequate spatial resolution to provide a decent and realistic representation of current climates and landscape features. Since the end goal is to predict the impacts of progressive climate change, climate data also need to provide a representation of present day climates as an average over a baseline period.

Towards that end, we have selected WorldClim (Hijmans et al., 2005), available at http://www.worldclim.org. These data represent present (1950-2000 averages) monthly climatology (maximum, minimum and mean temperatures, and total monthly precipitation). We downloaded the data at 2.5 arc-minute spatial resolution (approximately 5 km at the equator) for four variables (rainfall, and maximum, minimum and mean temperature) for each of the 12 months of the year.

### 2.2.2. Crop data

We harvested data on the presence of the crop from the GENESYS portal (http://www.genesys-pgr.org). The data consisted of geographic coordinates of 18,955 accessions of sorghum (*Sorghum bicolor*) landraces collected in areas where the crop is grown. The harvested data were carefully verified for the consistency of its geographic coordinates (latitude, longitude) and corrected whenever necessary. We selected only unique locations in 2.5 arc-minute spatial resolution gridcells for all further steps (3,681 locations, “crop dataset” hereafter). We prefer to use crop locations as given by landraces since the alternative approach of using crop distribution gridded data (Monfreda et al., 2008; You et al., 2009) can lead to inaccuracies due to the known biases in those datasets.

We acknowledge that by using a set of landrace accessions we might be capturing a wide range of the crop’s genetic variation, and therefore capturing a wide range of abiotic adaptations. Given the fact that the approach proposed here intends to develop a distributional range for the crop rather than for a particular genotype, we decided to use the whole set of accessions. In some cases, this approach might lead to the detection of different parameterizations yielding different results and differently fitting the data, an issue we cope with in subsequent sections.

### 2.2.3. Determination of ecological parameters

The aim of determining the ecological parameters is to explore the data using some basic statistical concepts and understand the ecological ranges of the crop. We used 80% of the presence points to calculate different ecological parameter sets, and the remaining 20% for selecting the correct parameter set and perform the model runs.

For each of the data points in the crop dataset, we extracted the corresponding values (from the present climate dataset) for maximum and minimum temperature and total rainfall variables and for each of the 12 months of the year. Then, for each of 12 potential growing seasons (assuming all months are equally likely to be the first month of the growing season), we calculate the average maximum and minimum temperatures and total rainfall. For each point, we then calculate the mean (ME), mode (MO), maximum (MX) and minimum (MN) of all growing seasons for each variable and each point. Finally, a total of 12 month-based potential growing seasons (starting in each of the 12 months) and 4 additional “fabricated” seasons (hence totaling 16) derived from initial set of 12 season (ME, MO, MX and MN) are produced for calculating different parameter sets as explained below.

For each of the growing seasons using all the presence points for each of the (3) variables and (16) growing seasons, a histogram is plotted, the mode is calculated, and five thresholds are extracted and assigned as the different ecological parameters to be used for running the EcoCrop model (Figure 2). For temperatures, $T_{KILL}$ is assigned as the 95% class value to the left of the mode, $T_{MIN\text{-}C}$ and $T_{MAX\text{-}C}$ are assigned as the 80% class values to the left and right of the mode, respectively; and $T_{OPMIN\text{-}C}$ and $T_{OPMAX\text{-}C}$ are assigned as 40% of the class values to the left and right of the mode, respectively (Figure 2A). For precipitation, $R_{MIN\text{-}C}$ and $R_{MAX\text{-}C}$ are assigned as the 80% class values to the left and right of the mode respectively, while $R_{OPMIN\text{-}C}$ and $R_{OPMAX\text{-}C}$ are assigned as 40% of the class values to the left and right of the mode, respectively (Figure 2B).

**[INSERT FIGURE 2 HERE]**

All the parameter sets are then used to drive the EcoCrop model. For each of the 16 potential growing seasons, we perform 2 runs of the model, one using the minimum temperature parameter set and the other using the maximum temperature parameter set; both of them use the same precipitation parameter set. Since it was observed in early versions of these analyses that individual parameterizations might not work in all cases, we combined the resulting suitability surfaces obtained from the maximum and minimum temperatures parameter sets (Eqn. 4).

$$SUIT_{TOTAL\ k} = \begin{cases} SUIT_{TMIN\ k} & SUIT_{TMIN\ k} \neq 0;\ SUIT_{TMAX\ k} = 0 \\ SUIT_{TMAX\ k} & SUIT_{TMIN\ k} = 0;\ SUIT_{TMAX\ k} \neq 0 \\ \dfrac{SUIT_{TMIN\ k}^{\ 2} + SUIT_{TMAX\ k}^{\ 2}}{SUIT_{TMIN\ k} + SUIT_{TMAX\ k}} & SUIT_{TMIN\ k} \neq 0;\ SUIT_{TMAX\ k} \neq 0 \end{cases} \quad \text{[Eqn. 4]}$$

The calculation is done on a pixel basis. $SUIT_{TMINk}$ is the suitability of the pixel of the *k-th* growing season, as calculated with the minimum temperature parameter set; $SUIT_{TMAXk}$ is the suitability of the pixel of the *k-th* growing season, as calculated with the maximum temperature parameter set. In this way, a total of 48 suitability surfaces are finally produced. Each one of them is assessed using the 20% remaining of the data.

The distribution of the 20% randomly selected data should resemble the distribution of the crop. The two measures of accuracy used to select the most accurate parameterization are the omission rate (OR, Eqn. 5), and the root mean square error (RMSE, Eqn. 6). A minimization of both values is not sought when assessing the preliminary suitability runs for the reasons given as it is not certain how suitable these environments are and therefore, in the comparison between the randomly selected known presences of the crop and the suitability surfaces we cannot assume a presence point means the crop is 100% suitable.

$$OR = \frac{n_{NZ}}{n} \quad \text{[Eqn. 5]}$$

$$RMSE = \sqrt{\frac{\sum_{p=1}^{n} \left(X_p - 1\right)^2}{n}} \quad \text{[Eqn. 6]}$$

Where $n$ is the total number of points, $X$ is the corresponding suitability value of the point $p$, and $n_{NZ}$ is the number of points that fall in suitable areas ($SUIT > 0$). In general, after observation of preliminary test runs of the model, a model with $OR$>0.1 and $RMSE$>0.5 was observed to heavily restrict the geographic distribution of the crop. Runs with $OR$<=0.1 and $RMSE$<=0.5 are selected. From these, the one with most accurate distributed prediction is chosen by examining the predictions against the known distribution of the crop (Monfreda et al., 2008; You et al., 2009; You et al., 2007). If the best growing season's suitability surface is $SUIT_{TOTAL}$, then this means that despite there is one single niche, climatic constraints act differently depending upon geographies, and hence two possible parameter sets for the crop, one derived from minimum temperatures and the other from maximum temperatures.

### 2.3. Modeling crop suitability

The modeling of the crop's suitability is a process that involves the evaluation of the model and the usage of the selected parameter set(s) to run the model using a certain (set of) climate scenario(s). Here we used a present climate scenario (given by WorldClim) and 24 different downscaled future climate scenarios.

#### 2.3.1. Present day climates run and model evaluation

Present day climate run consisted of applying the algorithm on a pixel basis using the selected parameterization and the climate data in WorldClim. We decided to test model predictions against the known presence of the crop, as reported in national and sub-national agriculture statistics. Four databases were queried, each with different gaps in the existing data (countries and years with data) and with different levels of detail (i.e. country, state, and district):

- FAOSTAT: the Food and Agriculture Organization (FAO) of the United Nations Statistics Database, containing several crops and (almost) all the countries in the world (FAO, 2010b).
- Agro-MAPS: a database developed by different organizations, also supported by FAO. It includes data at the state and district level, but its geographic coverage is not optimal (FAO, 2002).
- CountrySTAT: a database developed by FAO. Contains data at the state and district level, but the availability is not optimal both across time and space (FAO, 2010a).
- International Crops Research Institute for the Semi-Arid Tropics (ICRISAT): a database compiled by ICRISAT's Socio-economic Policy Division. Contains data for the period 1966-2000 for at least 80% of the districts in India (Challinor et al., 2004).

We performed the evaluation procedure at three different spatial levels: country, state, and district. For each of the administrative units for each of the spatial levels, the presence of the crop was assumed if the source reported at least one year with more than 10 ha within the study period (i.e. 1961-2000), and assumed suitable if there was at least one pixel suitable. As evaluation metrics, we calculate the true positive rate (*TPR*, Eqn. 7) as the number of features predicted and marked as suitable by the model (*NTP*) to the total number of available features to assess, and the false negative rate (*FNR*, Eqn. 8) as the number of features predicted by the model to not be suitable for the crop, but marked as cropped in national statistics (*NFN*) to the total number of available features to assess. Since the distribution of a crop is not only driven by climate, but also by political and socio-economic drivers, neither the true negative nor false positive rates could be calculated.

$$TPR = \frac{NTP}{Total} \quad \text{[Eqn. 7]}$$

$$FNR = \frac{NFN}{Total} \qquad \text{[Eqn. 8]}$$

It was observed that the higher the resolution, the less the available data. The exception was India, covered by the ICRISAT dataset, which both was high in resolution and had extensive temporal and within-country geographic coverage (Table 1).

**[INSERT TABLE 1 HERE]**

The available data are rather poor for some datasets, particularly CountrySTAT, which had only 11.8% states in the whole study region. For some datasets (i.e. Agro-MAPS, CountrySTAT) there was no single feature (i.e. state, district) with at least 50% of the years available. We also compared our parameterization with that in the FAO-EcoCrop database, to test the agreement of both and highlight the importance and relevance of the data at FAO.

### 2.3.2. Future climatic data

We downloaded projections of future climate used here from the Coupled Model Intercomparison Project Phase 3 (CMIP3) database. We downloaded monthly time series of maximum, minimum, and mean temperature, and total rainfall from https://esg.llnl.gov:8443/index.jsp for the 20$^{th}$ century and SRES-A1B 21$^{st}$ century simulations, from 24 different coupled global climate models -GCMs (Table 2) used in the IPCC Fourth Assessment Report (IPCC, 2007) for two different periods: (a) baseline (1961-1990), and (b) 2030s (2020-2049). We downscaled the data as described in Ramirez and Jarvis (2010).

**[INSERT TABLE 2 HERE]**

Although we acknowledge this type of downscaling is referred to as “unintelligent” (Thornton et al., 2011; Wilby et al., 2009), it is often the only option when assessing impacts at higher spatial scales than GCM resolutions and in areas with considerable variability in orography (Ramirez and Jarvis, 2010; Tabor and Williams, 2010). Finally, we obtained a total of 24 future scenarios at the same spatial resolution of WorldClim data (i.e. 2.5 arc-minutes). Each of these scenarios represent monthly means of maximum, minimum, and mean temperatures, and total rainfall, for the SRES-A1B emission scenario by 2030s. We selected this time-slice and scenario because the 2030s are at a close time horizon by which most of the necessary adaptation strategies for climate change-vulnerable crops should be in place. Additionally, by 2030s there is not much difference between the different SRES storylines (Arnell et al., 2004; IPCC, 2007)

### 2.3.3. Relations to yield, assessing impacts and uncertainties

Using crop distributions as reported by You et al. (You et al., 2009) we compared the numerical output of EcoCrop with yield by extracting 1,000 random points over the study area from areas that did not have optimal (100%) or no (0%) suitability in EcoCrop. The latter was done to avoid biases, as there are other factors that drive crop yields and it is likely that 100% suitable areas would have low values due to other factors. We then did a basic exploration of the data using quantile plots and dispersion diagrams.

For the selected parameter set(s), we drove the EcoCrop model using the 24 future climate scenarios in the same way we did with WorldClim (Sect. 2.3.1). We calculated some uncertainty metrics to accompany the climate change impact metrics. For each of the 24 future suitability results, we calculated the change in suitability as the difference between the future scenario and the baseline. We then calculated the average (of all GCMs) on a pixel basis of these changes as measure of the general trend and the geographic distribution of among-GCM variability. In addition, for each GCM-specific result, we calculated the overall percent increase and decrease in area suitable assuming both migration and no migration of agriculturally suitable lands.

To illustrate uncertainties, we constructed four maps: (a) a map of the standard deviation of all GCMs; (b) a map showing the average of the first 25% of the GCMs per pixel; (c) a map showing the average of the last 25% of the GCMs per pixel; and (d) a map showing the percent of models that predict changes in the same direction of the average prediction (IPCC, 2007; Schlenker and Lobell, 2010).

## 3. Results

### 3.1. Model calibration and parameterization

First, we chose the duration of the growing season. According to different studies (Craufurd et al., 1999; FAO, 2000; Geleta and Labuschagne, 2005; Mishra et al., 2008), sorghum can be harvested between 90 and 300 days after sowing, depending on the variety, with the most frequent range being 150-200 days (Craufurd et al., 1999; FAO, 2000; Geleta and Labuschagne, 2005). As the growing season length in EcoCrop is defined in months, we decided to test for different growing seasons (between 3 and 10 months). The best performance was achieved with a growing season of 6 months (data not shown), although differences in results of present suitability using this value and 7, 8 and 9 months were negligible. Shorter growing seasons always showed poor performance, although we acknowledge this in reality depends on the temperatures and radiation available to the crop and that often sorghum is harvested between 4 and 5 months after sowing (Geleta and Labuschagne, 2005; Mishra et al., 2008). Two conclusions were drawn from this result (1) our model is not highly sensitive to the length of the growing season (i.e. a flaw in EcoCrop), and (2) the considerable variability in the landrace dataset is likely to have a mixture of different growing seasons, and it is likely we are capturing the most frequent of it (i.e. 6 months).

We found that only 10.4% of the parameterizations were highly accurate (i.e. OR<0.1 and RMSE between 0.25 and 0.5). The combined parameterizations (derived from Eqn. 4) were the most accurate, suggesting that despite there is only one possible niche for the crop there could be two different environmental constraints (i.e. minimum and maximum temperature as principal limiting factors), each producing a different climate-suitability geographical gradient. These responses can be considered as within-crop among-landrace genetic variability.

**[INSERT FIGURE 3 HERE]**

The selected parameter set (Table 3) indicated that the crop's distributional range is meant to be subjected to two climate constraints. The first one indicates the crop is located in low-temperature stressed areas (i.e. sub-tropical environments and highlands, figure not shown) and it would thus freeze if minimum temperature during the growing season goes below 0.5ºC [+4ºC], is not suited below 4.1ºC, thrives optimally between 13.6ºC and 24.6ºC and is heat stressed in temperatures above 26ºC. On the other hand, the parameter set derived from seasonal maximum temperature data indicates that the crop landraces in these areas to high-temperature stresses (i.e. mainly across the Sahelian belt, figure not shown). In this case the crop would die if the minimum temperature of at least one month goes below 14.5[+4ºC], is not suited for a mean temperature below 17.8ºC, grows optimally in the range 26.7–37.4ºC, and will not grow if temperatures are above 39.1ºC. This result stressed the difficulty in fitting one single parameter set to (1) a large number of environments and (2) a genetically-variable landrace dataset, and also stressed the importance of considering the different constraints in space (and time).

**[INSERT TABLE 3 HERE]**

Regarding precipitation, the crop is harmfully stressed if the total rainfall during the growing season is less than 160 mm (drought) or above 2,780 mm (excess water, or waterlogging). Sorghum develops best between 500 and 1,800 mm of rainfall during the growing season.

### 3.2. Present day suitability and model evaluation

As expected, the greatest constraint to sorghum distribution is the very hot and dry weather above the Sahel region in Africa (Figure 4). Suitability is mostly below 50% in areas under high temperature and/or rainfall stress in southern Mauritania, central Mali, Niger, Chad, Sudan and Eritrea, southeastern Ethiopia, central Somalia, northeastern Kenya, Namibia and Botswana. In contrast, in India, the crop was found to be highly suitable across nearly the whole country.

**[FIGURE 4 HERE]**

The areas where the crop is most intensively grown are located in the borders between Niger and Nigeria, and in the states of Maharashtra and Karnataka (Monfreda et al., 2008; Portmann et al., 2010; You et al., 2009), which are also areas of high suitability in our prediction.

The TPR and FNR in general showed high and low values, respectively, regardless of the dataset from which they were calculated (Table 4). TPR ranged from 0.967 (FAOSTAT dataset) to 1.0 (CountrySTAT and ICRISAT datasets), while FNR ranged from 0 (CountrySTAT and ICRISAT dataset) and 0.026 (AgroMAPS state-level dataset).

**[INSERT TABLE 4 HERE]**

### 3.3. Relations to yield, future predictions of suitability and impacts

Relationships between suitability with yields were not clear from the actual values of both suitability and yields, and we could not find a way to numerically relate both outputs in absolute terms. A linear regression is not statistically significant, although it has a positive slope. Also, we clearly observed through a quantile plot that high values of suitability corresponded to high values of yield more likely than they corresponded to low values, although the relationship is not linear.

Changes ranged between –93 and 61% (Figure 5) and lower GCM-specific averages (Table 5). Tropical humid areas are likely to present the most significant losses, whilst subtropical regions (i.e. the north-east Indo-Gangetic Plains, Nepal, and central Botswana) present some gains (Figure 7). There are also gains in some areas in East Africa (i.e. eastern Ethiopia) and in the semi-arid regions of Mali, Niger, Chad and Sudan. East Africa and the Indian subcontinent appear as the most affected regions (Figure 5) and considerable between-GCM variability (Table 5).

**[INSERT TABLE 5 HERE]**

There were particularly negative impacts in central Ethiopia, Uganda, south-eastern Kenya and Tanzania, where between 50–80% of the suitable areas could decrease in climatic suitability even when assuming agriculturally suitable lands can move to new environments (Figure 5).

**[INSERT FIGURE 5 HERE]**

The most significant decrease in the amount of suitable area and in the average suitability occurred in the range of 80-90%, particularly in areas where the crop is already marginal (*SUIT*<50%). On the other hand, only a limited expansion of suitable croplands was predicted, and this was observed mainly in currently very low suitability areas (where cropping is unsustainable) or in areas where suitability is optimal (Figure 5).

### 3.4. Climate-driven uncertainties

The great majority of croplands within the study region present rates of agreement between models ranging between 60 and 80% (Figure 6C), mostly covering Sub-Saharan Africa, and several parts of India. Low confidence (AG<50%) is observed in the Congo and Central African Republic, as well as in Namibia, Botswana and Zimbabwe and the Sahel. The analysis shows a considerably high confidence in negatively impacted areas (Figure 6); however, there is less certainty when the predicted impacts are positive.

**[INSERT FIGURE 6 HERE]**

More than 50% of the countries showed particularly low amounts of area with high certainty (AG>80%), and high proportions of area with very low certainty (AG=50%), particularly in Eastern and Southern Africa. Despite that, differences in conservative (upper 25%, Figure 6C) and non-conservative GCMs (lower 25%, Figure 6B) are considerable in some regions, particularly in those where very negative (SUIT change < 30%) impacts are observed. In these areas, the different models depict completely different pictures on impacts.

## 4. Discussion

### 4.1. Modeling approach and model-evaluation results

The benefits of a more simplistic approach are considerable, despite some caveats and uncertainties (see Sect. 4.3) that require further research and work. An approach as the one proposed here reduces the parameterizations to a minimum while at the same time making sense of the biology of the crop species (Hijmans et al., 2001). Here, the ecological parameters are related to crop growth as they represent the thresholds at which the crop can grow and produce harvestable product.

Although a calibration procedure has been provided, crop experts and/or literature must be queried to gather the ecological parameters required to perform EcoCrop. The FAO-EcoCrop database (FAO, 2000) contains ~1,800 different crops' ecological parameterizations. While these ecological parameterizations have not been validated, they are based on either literature or expert views on the crop and can provide a relatively accurate estimate of the crop's adaptive capacity and ecological niche. We compared our predictions done with default parameters for three types of sorghum genotypes (as reported in FAO-EcoCrop, Table 3) and found that for high altitude, medium altitude and low altitude sorghum the agreement was very high ($R^2$=0.865, $R^2$=0.878, and $R^2$= 0.854, all at $p<0.0001$), though the default parameters tended to exclude areas in southern Africa, very likely due to the difficulty in capturing seasonal climates, an advantage of the calibration using crop locations.

Although it is difficult to quantitatively compare results from other studies mainly because these use (a) a different emissions scenario, (b) a different set of GCMs, (c) a different period, or (d) a combination of (a), (b) and (c). When comparing EcoCrop results with the studies of Chipanshi et al. (2003), Lobell et al. (2008), Schlenker and Lobell (2010), and Srivastava et al. (2010), we found that results on a country and region basis agreed 88.4% of the times. Negative impacts were predicted 92.5% of the times whereas positive impacts were predicted 33.3% of the times (but only 3 cases with positive impacts were found in the reviewed studies). In addition, we compared the actual estimates of the different studies (Figure 7) and found despite all estimates are heavily subjected to uncertainties, there were considerable similarities in our estimates of changes in suitable area and suitability *per se* and the changes in yields reported in the other studies, both expressed as percentages.

**[INSERT FIGURE 7 HERE]**

Central Africa (CAF), southern Africa (SAF), East Africa (EAF), are the areas where we found the greatest agreement, whereas the Sahel (SAH), Southern Asia (SAS) and western Africa (WAF) show higher variability within and between studies yet showing up to 75% agreement in the direction of changes.

### 4.2. Climate constraints, future impacts and adaptation

Sorghum is adapted to a wide range of environmental conditions, but the main factor operating against the expansion of sorghum croplands in the tropics is seasonal precipitation (Folliard et al., 2004; Kouressy et al., 2008; Neild et al., 1983). Sorghum is particularly sensitive to shortages in water in late development stages, and hence, sowing time, although flexible, is critical for avoiding crop failure (Smith and Frederiksen, 2000). Additionally, it is very likely that increases in temperatures (as found in this study) will not pose a strong pressure in areas where sorghum grows optimally (though yields could reduce if the temperature rises beyond the +2ºC limit), but that is not the case in marginal areas (Lobell et al., 2008).

In vulnerable areas in Sub-Saharan Africa and the Indo Gangetic Plains, adaptation needs to happen before negative impacts become too severe or too costly. There is an opportunity for simple strategies to minimize yield losses. For instance, delayed sowing can help crops avoid water stress during initial growth phases (Srivastava et al., 2010). Nevertheless, biological adaptation also needs to happen. The sorghum genetic pool contains a wide range of traits that might be useful under changing climate conditions (Geleta and Labuschagne, 2005; Kameswara Rao et al., 2003; Mekbib, 2008). In terms of sorghum adaptation in Sub-Saharan Africa and India, both growing cycle duration and drought tolerance are two of the most important abiotic traits meriting research focus (Kouressy et al., 2008; Krishna Kumar et al., 2004; Srivastava et al., 2010)

Other strategies such as crop substitution and targeting have also been suggested in different studies (Chipanshi et al., 2003; Jarvis et al., 2010; Lane and Jarvis, 2007). Expansion to new agriculturally suitable areas is another adaptive pathway under climate change, since some environments with particularly low temperatures will likely become suitable in the future; in our observations, these areas were in the highlands of the semi-arid tropics.

### 4.3.Uncertainties, caveats and further improvements

Figures and results obtained from these types of approaches are subject to both inaccuracies and uncertainties, and this suggests that they could be improved. Below we summarize the most relevant sources of uncertainty in our approach and point out some ways in which these could be addressed.

#### 4.3.1. Climate data

Two different sources of climate data were used in this study: WorldClim and GCM data. Although not quantified in the present study, in WorldClim, uncertainties can arise from the location of the weather stations (latitudinal, altitudinal biases, see Hijmans et al. 2005), from the interpolation algorithm (Hutchinson and de Hoog, 1985), from the quality of historical records, and/or from the geographic distances between stations.

GCM data accounted for a significant amount of uncertainty (Figure 5 and Sect. 3.4), mainly because the predicted changes in climates (i.e. temperatures, rainfall) exhibit considerable variability among GCMs (Pierce et al., 2009; Quiggin, 2008). In areas where GCM predictions do not reach an admissible certainty threshold, options are basically to further climate research to improve calibration or to develop and/or calibrate regional models (RCMs) that can yield better results.

Finally, the process of spatial downscaling performed is also a source of uncertainty (Wilby et al., 2009). Further research needs to be done to improve GCM and RCM predictions for areas where convection processes are complex and cannot be easily captured with parameterization schemes (Wagner and Graf, 2010). Meanwhile, assessments of the quality of downscaled GCM data, in relation to a possible "degradation" and "misinterpretation" of the GCM data, need to be addressed.

#### 4.3.2. Model calibration and evaluation data

Cleansing of the occurrence data used for calibration of the present approach is critical in order to properly identify the actual areas where the crop is suitable (Hill et al., 2009; Yesson et al., 2007). Therefore, cross-checking, verification and retrieval of accurate coordinates are necessary when performing this type of approach.

Using expert data or literature to identify the ecological parameters needed to perform the EcoCrop model as done in the FAO-EcoCrop database can also induce errors. Hence, it is important to query as many different sources as possible when deriving the ecological ranges, as well as to interact with experts in the crop to visually inspect and further refine the suitability result.

Given that evaluation data are mainly a mixture of different political-level agricultural statistics, the evaluation proposed here is dependent on both the availability and the precision of such data. Further development and improvement of global online platforms such as FAOSTAT and CountrySTAT is therefore fundamental to proper evaluation of model's performance.

#### 4.3.3. Model formulation

The implementation of the EcoCrop model proposed here is subjected to some limitations:

(1) The biological sense of the model's parameters: For temperature, optimal and marginal thresholds are also included in mechanistic crop models and are used to derive growing degree days (Yang et al., 2004). In process-based models, water flow is first analyzed in the soil and then in the plant as absorbed until it is transpired in the leaves. Although responses in plants vary, lack or excess of water cause lower yields, and there is a level of available soil water above and below which plants fail to flower, flower too early, do not fill grains, or die (Whitmore

and Whalley, 2009). In rainfed systems, these values depend upon rainfall. The simplistic approach in EcoCrop tries to simulate the non-linear effects of these stresses, but it fails to capture the whole set of interactions occurring within the plant at the physiological level. Therefore, suitability indices and their likely changes need to be interpreted carefully as the ability of a certain environment to allow the growth of a certain species in a broad sense.

(2) For perennial crops, it is harder to calibrate the modeling approach, since the rainfall and temperatures during the growing season are equal to the annual rainfall and temperature, which results in neglecting climate seasonality. A good option to overcome this issue would be the development of a function to involve the concept of degree days (Neild et al., 1983),

(3) The model does not account for soil conditions and becomes less accurate when estimating suitability in very well-drained soils in high-rainfall areas where waterlogging could be but is not a constraint. Here we decided to not use soil data since (a) there is not enough spatial resolution in the available soil datasets and (b) it would be complicated to derive soil conditions when predicting future crop suitability;

(4) The model does not account for drought, waterlogging, excessive heat or cold during key physiological periods (i.e. fruit filling, flowering), leading to a climatic suitability over-estimation;

(5) The application of the model relies upon monthly data, whilst stressful conditions may occur in shorter periods (i.e. one week or two). In addition, the model does not provide an indication of the relationship between suitability and yield.

(6) The fact that the model has a fixed duration of the growing season facilitates the selection of ecological parameters, but poses a constraint as physiologically crops do not have always the same growing season. Clustering of data into agro-ecological zones can solve this problem, accompanied by a derivation of growing season duration on these agro-ecologies.

We consider that given the flexibility of the approach, it can be continuously improved, and some additional processes can be incorporated. We acknowledge that other environmental, social, cultural, and political conditions likely also affect the resulting yield of a field plot. More research is therefore required towards the clear identification of the relationship between our climatic suitability rating and the resulting attainable yield obtained in fields.

### 4.4. Future focus and research priorities

Further mining of datasets to find a clearer relationship between yield and the suitability index is necessary for EcoCrop results to be comparable with results from other models and studies, whose responses are in terms of yield (Aggarwal et al., 2006; Challinor et al., 2004; Jones and Thornton, 2003; Steduto et al., 2009; Thornton et al., 2009; Thornton et al., 2011).

Policy makers may invest in the most effective measures with the least risk (win-win strategies). The caveats in the modeling and the agreement between different GCMs' are key to deciding where, when and how much to invest. Despite the limitations, which we have tried to mention at the maximum extent possible, the approach implemented here provides an initial broad picture of what the effects could be of changing conditions on the regional suitability of the sorghum crop. Moreover, the EcoCrop model can be used for the same purpose for basically any existing crop, as long as the ecological range is determined.

## 5. Conclusions

Here we have proposed a simple model to assess the impacts of progressive climate change. The model can be tuned either by using the known presences of a crop or using expert knowledge, or by directly drawing data from the FAO-EcoCrop database. The model was found to perform well when predicting suitable areas under present conditions, although some questions as to how accurate its predictions of future impact and how predictions relate to yield remain unresolved. In the present study, we found that these are similar to other studies, though it depends upon the region of study.

Using the model, we predicted the impacts of climate change on sorghum-growing areas and found that in general the crop is performing well in the areas where it grows optimally. Vulnerabilities in countries where sorghum cultivation is already marginal are likely (with a high degree of certainty). The western Sahel region, southern Africa, northern India, and the western coast of India are particularly vulnerable. The same pattern is observed in southern Africa, where suitable areas could be reduced by some 20% by the 2030s. Uncertainty was found to play an important role, with a large area under the high uncertainty range (Figure 6). Our results could benefit considerably from better GCM parameterizations and results.

We highlight the considerable potential of this approach to assess global and regional food security issues, broad climatic constraints and regional crop-suitability shifts in the context of climate change, as well as the possible linkage of the approach with other broad-scale approaches such as large-area process-based crop models or statistical and/or empirical approaches.

## Acknowledgments

Authors thank the anonymous reviewers for their useful and insightful comments aimed at improving this manuscript, and Myles Fisher (despite he withdraw from reviewing the manuscript) for the helpful discussions held via e-mail and at the International Center for Tropical Agriculture (CIAT). Authors also thank Charlotte Lau, from the International Center for Tropical Agriculture (CIAT) for her editorial work, Robert Hijmans, the developer of the initial algorithm (implemented through DIVA-GIS) used as the basis for this research, and David Lobell for providing data from his studies. This work was supported by the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS).

## References

Aggarwal, P.K., Kalra, N., Chander, S. and Pathak, H., 2006. InfoCrop: A dynamic simulation model for the assessment of crop yields, losses due to pests, and environmental impact of agro-ecosystems in tropical environments. I. Model description. Agricultural Systems, 89(1): 1-25.

Arnell, N.W. et al., 2004. Climate and socio-economic scenarios for global-scale climate change impacts assessments: characterising the SRES storylines. Global Environmental Change, 14(1): 3-20.

Byjesh, K., Kumar, S. and Aggarwal, P., 2010. Simulating impacts, potential adaptation and vulnerability of maize to climate change in India. Mitigation and Adaptation Strategies for Global Change, 15(5): 413-431.

Challinor, A.J., Ewert, F., Arnold, S., Simelton, E. and Fraser, E., 2009. Crops and climate change: progress, trends, and challenges in simulating impacts and informing adaptation. Journal of Experimental Botany, 60(10): 2775-2789.

Challinor, A.J., Simelton, E.S., Fraser, E.D.G., Hemming, D. and Collins, M., 2010. Increased crop failure due to climate change: assessing adaptation options using models and socio-economic data for wheat in China. Environmental Research Letters, 5(3): 034012.

Challinor, A.J., Wheeler, T.R., Craufurd, P.Q. and Slingo, J.M., 2005. Simulation of the impact of high temperature stress on annual crop yields. Agricultural and Forest Meteorology, 135(1-4): 180-189.

Challinor, A.J., Wheeler, T.R., Craufurd, P.Q., Slingo, J.M. and Grimes, D.I.F., 2004. Design and optimisation of a large-area process-based model for annual crops. Agricultural and Forest Meteorology, 124(1-2): 99-120.

Chipanshi, A.C., Chanda, R. and Totolo, O., 2003. Vulnerability Assessment of the Maize and Sorghum Crops to Climate Change in Botswana. Climatic Change, 61(3): 339-360.

Craufurd, P.Q. et al., 1999. Adaptation of sorghum: characterisation of genotypic flowering responses to temperature and photoperiod. TAG Theoretical and Applied Genetics, 99(5): 900-911.

FAO, 2000. The Ecocrop database. In: FAO (Editor), Rome, Italy.

FAO, 2002. AgroMaps. In: FAO (Editor), Rome, Italy.

FAO, 2009. The State of Food and Agriculture. FAO, Rome, Italy.

FAO, 2010a. CountrySTAT. In: FAO (Editor), Rome, Italy.

FAO, 2010b. FAOSTAT. In: FAO (Editor), Rome, Italy.

FAO, 2010c. The State of Food Insecurity in the World. FAO, Rome, Italy.

Folliard, A., Traoré, P.C.S., Vaksmann, M. and Kouressy, M., 2004. Modeling of sorghum response to photoperiod: a threshold-hyperbolic approach. Field Crops Research, 89(1): 59-70.

Fuller, D.Q., 2007. Contrasting Patterns in Crop Domestication and Domestication Rates: Recent Archaeobotanical Insights from the Old World. Annals of Botany, 100(5): 903-924.

Geleta, N. and Labuschagne, M.T., 2005. Qualitative Traits Variation in Sorghum (<i>Sorghum Bicolor</i> (L.) Moench) Germplasm from, Eastern Highlands of Ethiopia. Biodiversity and Conservation, 14(13): 3055-3064.

Gregory, P.J., Ingram, J.S.I. and Brklacich, M., 2005. Climate change and food security. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1463): 2139-2148.

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

Hijmans, R.J., Guarino, L., Cruz, M. and Rojas, E., 2001. Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. Plant Genetic Resources Newsletter, 127: 15-19.

Hill, A. et al., 2009. Location, location, location: utilizing pipelines and services to more effectively georeference the world's biodiversity data. BMC Bioinformatics, 10(0): 1-9.

Hutchinson, M.F. and de Hoog, F.R., 1985. Smoothing noisy data with spline functions. Numerische Mathematik, 47: 99-106.

IPCC, 2001. IPCC Third Assessment Report: Climate Change 2001 (TAR). IPCC, Geneva, Switzerland.

IPCC, 2007. IPCC Fourth Assessment Report: Climate Change 2007 (AR4). IPCC, Geneva, Switzerland.

Jarvis, A., Ramirez, J., Anderson, B., Leibing, C. and Aggarwal, P., 2010. Scenarios of Climate Change Within the Context of Agriculture. Climate Change and Crop Production. CAB International.

Jones, P.G. and Thornton, P.K., 2003. The potential impacts of climate change on maize production in Africa and Latin America in 2055. Global Environmental Change, 13(1): 51-59.

Kameswara Rao, N., Reddy, L.J. and Bramel, P.J., 2003. Potential of wild species for genetic enhancement of some semi-arid food crops. Genetic Resources and Crop Evolution, 50(7): 707-721.

Kouressy, M., Dingkuhn, M., Vaksmann, M. and Heinemann, A.B., 2008. Adaptation to diverse semi-arid environments of sorghum genotypes having different plant type and sensitivity to photoperiod. Agricultural and Forest Meteorology, 148(3): 357-371.

Krishna Kumar, K., Rupa Kumar, K., Ashrit, R.G., Deshpande, N.R. and Hansen, J.W., 2004. Climate impacts on Indian agriculture. International Journal of Climatology, 24(11): 1375-1393.

Lane, A. and Jarvis, A., 2007. Changes in Climate will modify the Geography of Crop Suitability: Agricultural Biodiversity can help with Adaptation. Journal of the Semi-Arid Tropics, 4(1): 1-12.

Lobell, D.B. et al., 2008. Prioritizing Climate Change Adaptation Needs for Food Security in 2030. Science, 319(5863): 607-610.

Meehl, G.A. et al., 2005. How Much More Global Warming and Sea Level Rise? Science, 307(5716): 1769-1772.

Mekbib, F., 2008. Genetic erosion of sorghum (<i>Sorghum bicolor</i> (L.) Moench) in the centre of diversity, Ethiopia. Genetic Resources and Crop Evolution, 55(3): 351-364.

Mishra, A. et al., 2008. Sorghum yield prediction from seasonal rainfall forecasts in Burkina Faso. Agricultural and Forest Meteorology, 148(11): 1798-1814.

Monfreda, C., Ramankutty, N. and Foley, J.A., 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global Biogeochem. Cycles, 22(1): GB1022.

Neild, R.E., Logan, J. and Cardenas, A., 1983. Growing season and phenological response of sorghum as determined from simple climatic data. Agricultural Meteorology, 30(1): 35-48.

Nisar Ahamed, T.R., Gopal Rao, K. and Murthy, J.S.R., 2000. GIS-based fuzzy membership model for crop-land suitability analysis. Agricultural Systems, 63(2): 75-95.

Pierce, D.W., Barnett, T.P., Santer, B.D. and Gleckler, P.J., 2009. Selecting global climate models for regional climate change studies. Proceedings of the National Academy of Sciences, 106(21): 8441-8446.

Portmann, F.T., Siebert, S. and Döll, P., 2010. MIRCA2000 --Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. Global Biogeochem. Cycles, 24(1): GB1011.

Quiggin, J., 2008. Uncertainty and Climate Change Policy. Economic Analysis & Policy, 38(2): 203-210.

Ramirez, J. and Jarvis, A., 2010. Downscaling Global Circulation Model Outputs: The Delta Method. Decision and Policy Analysis Working Paper No. 1. Decision and Policy Analysis. International Center for Tropical Agriculture (CIAT), Cali, Colombia.

Rivington, M. and Koo, J., 2011. Report on the Meta-Analysis of Crop Modelling for Climate Change and Food Security Survey, CGIAR Research Program on Climate Change, Agriculture and Food Security.

Schlenker, W. and Lobell, D.B., 2010. Robust negative impacts of climate change on African agriculture. Environmental Research Letters, 5(1): 014010.

Schroth, G. et al., 2009. Towards a climate change adaptation strategy for coffee communities and ecosystems in the Sierra Madre de Chiapas, Mexico. Mitigation and Adaptation Strategies for Global Change, 14(7): 605-625.

Smith, C.W. and Frederiksen, R.A., 2000. Sorghum: Origin, History, Technology and Production. John Wiley & Sons, USA.

Srivastava, A., Naresh Kumar, S. and Aggarwal, P.K., 2010. Assessment on vulnerability of sorghum to climate change in India. Agriculture, Ecosystems & Environment, 138(3-4): 160-169.

Steduto, P., Hsiao, T.C., Raes, D. and Fereres, E., 2009. AquaCrop—The FAO Crop Model to Simulate Yield Response to Water: I. Concepts and Underlying Principles. Agron. J., 101(3): 426-437.

Tabor, K. and Williams, J.W., 2010. Globally downscaled climate projections for assessing the conservation impacts of climate change. Ecological Applications, 20(2): 554-565.

Thornton, P.K., Jones, P.G., Alagarswamy, G. and Andresen, J., 2009. Spatial variation of crop yield response to climate change in East Africa. Global Environmental Change, 19(1): 54-65.

Thornton, P.K., Jones, P.G., Ericksen, P.J. and Challinor, A.J., 2011. Agriculture and food systems in sub-Saharan Africa in a 4°C+ world. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369(1934): 117-136.

Wagner, T.M. and Graf, H.-F., 2010. An ensemble cumulus convection parameterisation with explicit cloud treatment. Journal of the Atmospheric Sciences, 0(0): null.

Whitmore, A.P. and Whalley, W.R., 2009. Physical effects of soil drying on roots and crop growth. Journal of Experimental Botany, 60(10): 2845-2857.

Wilby, R.L. et al., 2009. A review of climate risk information for adaptation and development planning. International Journal of Climatology, 29(9): 1193-1215.

Yang, H.S. et al., 2004. Hybrid-maize--a maize simulation model that combines two crop modeling approaches. Field Crops Research, 87(2-3): 131-154.

Yesson, C. et al., 2007. How Global Is the Global Biodiversity Information Facility? PLoS ONE, 2(11): e1124.

You, L., Wood, S. and Wood-Sichra, U., 2009. Generating plausible crop distribution maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. Agricultural Systems, 99(2-3): 126-140.

You, L., Wood, S., Wood-Sichra, U. and Chamberlin, J., 2007. Generating Plausible Crop Distribution Maps for Sub-Saharan Africa Using a Spatial Allocation Model. Information Development, 23(2-3): 151-159.

## Figure captions

**Figure 1** Two- (A) and three-dimensional (B) diagram of the mechanistic model used in the analysis.

**Figure 2** Example of parameter selection for a certain distribution over a particular growing season for (A) temperature and (B) precipitation

**Figure 3** Assessment of preliminary predictions for parameter selection. OR: Omission rate and RMSE: Root mean square error. Areas in the chart indicate the optimal ranges for both accuracy parameters: highly under-estimative (HU), highly over-estimative (HO), moderately accurate (MA), and highly accurate (HA)

**Figure 4** Present suitability and known distribution of the crop. (A) Sorghum suitability calculated with EcoCrop and parameter set (small bottom-right map), (B) sorghum distribution as reported in You et al. (2009), (C) sorghum distribution as reported in Monfreda et al. (2008), (D) sorghum distribution as in Portmann et al. (2010)

**Figure 5** Predicted changes in suitability across the region as an average of 24 GCMs.

**Figure 6** Uncertainties in suitability prediction across the region. (A) Standard deviation of 24 GCMs and standard deviation among predictions, (B) Average of the first 25% GCMs, (C) average of the last 25% GCMs, (D) Agreement among GCMs (fraction of GCMs agreeing direction).

**Figure 7** Agreement of the estimates of impacts in the present study with those reported in previous studies. CSA: change in suitable area (in percent), CS: change in suitability (in percent), LO 2008: Lobell et al. (2008), SL 2010: Schlenker and Lobell (2010), CH 2003: Chipanshi et al. (2003), SR 2010: Srivastava et al. (2010), all in percent. The boxplot represent the distribution of all available outputs (country means, and GCM-specific results, if available) for each study as found in the original papers or as provided by the authors (i.e. SL2010, LO2008). Black horizontal lines are the median, boxes show the first and third quartile and whiskers extend 5 and 95% of the distributions. Zone typology is the same as in Lobell et al. (2008) (see http://www.sciencemag.org/cgi/content/full/319/5863/607/DC1)

**Table 1** Proportion of data available relative to the total potential data in the different databases

| Database | Level* | Countries with data (%) | Features in database (%) | Features with data (%) | Features with >50% of data (%) | Maximum percent of data (%) |
|---|---|---|---|---|---|---|
| FAOSTAT | C | 75.7 | 100.0 | 75.7 | 72.9 | 100.0 |
| CountrySTAT | S | 10.0 | 14.2 | 11.8 | 4.5 | 65.0 |
| CountrySTAT | D | 4.3 | 3.9 | 3.2 | 0.0 | 37.5 |
| Agro-MAPS | S | 40.0 | 84.5 | 32.3 | 0.0 | 47.5 |
| Agro-MAPS | D | 12.9 | 66.2 | 6.1 | 0.0 | 12.5 |
| ICRISAT | D | 1.4 | 100.0 | 66.7 | 62.1 | 92.5 |

*C=Country, S=State, D=District

**Table 2** Global Circulation Models used in the analyses

| Model | Country | Atmosphere* | Ocean* |
|---|---|---|---|
| BCCR-BCM2.0 | Norway | T63, L31 | 1.5x0.5, L35 |
| CCCMA-CGCM3.1 (T47) | Canada | T47 (3.75x3.75), L31 | 1.85x1.85, L29 |
| CCCMA-CGCM3.1 (T63) | Canada | T63 (2.8x2.8), L31 | 1.4x0.94, L29 |
| CNRM-CM3 | France | T63 (2.8x2.8), L45 | 1.875x(0.5-2), L31 |
| CSIRO-Mk3.0 | Australia | T63, L18 | 1.875x0.84, L31 |
| CSIRO-Mk3.5 | Australia | T63, L18 | 1.875x0.84, L31 |
| GFDL-CM2.0 | USA | 2.5x2.0, L24 | 1.0x(1/3-1), L50 |
| GFDL-CM2.1 | USA | 2.5x2.0, L24 | 1.0x(1/3-1), L50 |
| GISS-AOM | USA | 4x3, L12 | 4x3, L16 |
| GISS-MODEL-EH | USA | 5x4, L20 | 5x4, L13 |
| GISS-MODEL-ER | USA | 5x4, L20 | 5x4, L13 |
| IAP-FGOALS1.0-G | China | 2.8x2.8, L26 | 1x1, L16 |
| INGV-ECHAM4 | Italy | T42, L19 | 2x(0.5-2), L31 |
| INM-CM3.0 | Russia | 5x4, L21 | 2.5x2, L33 |
| IPSL-CM4 | France | 2.5x3.75, L19 | 2x(1-2), L30 |
| MIROC3.2-HIRES | Japan | T106, L56 | 0.28x0.19, L47 |
| MIROC3.2-MEDRES | Japan | T42, L20 | 1.4x(0.5-1.4), L43 |
| MIUB-ECHO-G | Germany/Korea | T30, L19 | T42, L20 |
| MPI-ECHAM5 | Germany | T63, L32 | 1x1, L41 |
| MRI-CGCM2.3.2A | Japan | T42, L30 | 2.5x(0.5-2.0) |
| NCAR-CCSM3.0 | USA | T85L26, 1.4x1.4 | 1x(0.27-1), L40 |
| NCAR-PCM1 | USA | T42 (2.8x2.8), L18 | 1x(0.27-1), L40 |
| UKMO-HADCM3 | UK | 3.75x2.5, L19 | 1.25x1.25, L20 |
| UKMO-HADGEM1 | UK | 1.875x1.25, L38 | 1.25x1.25, L20 |

*Horizontal (T) resolution indicates number of cells in which the globe was divided for each component of the coupled climate model (i.e. atmosphere, ocean). Vertical (L) resolution indicates the number of layers in which the atmosphere was divided. When a model is developed with different latitudinal and longitudinal resolutions, the respective cellsizes (LonxLat) in degrees are provided instead of a unique value

**Table 3** Selected parameter set for suitability calculation and reported parameters in the FAO-EcoCrop database

| Source | Variable | Originating data | Kill | Min | Opmin | Opmax | Max |
|---|---|---|---|---|---|---|---|
| Calibration | Temperature | Min temp. | 0.5 | 4.1 | 13.6 | 24.6 | 26.0 |
| Calibration | Temperature | Max. temp. | 14.5 | 17.8 | 26.7 | 37.4 | 39.1 |
| Calibration | Precipitation | Precipitation | NA | 160 | 500 | 1,800 | 2,780 |
| FAO (2000) | Temperature | HAS* | 0 | 12.0 | 22.0 | 32.0 | 35.0 |
| FAO (2000) | Precipitation | HAS* | NA | 300 | 500 | 1,000 | 3,000 |
| FAO (2000) | Temperature | MAS* | 0 | 8.0 | 27.0 | 32.0 | 40.0 |
| FAO (2000) | Precipitation | MAS* | NA | 300 | 500 | 1,000 | 3,000 |
| FAO (2000) | Temperature | LAS* | 0 | 10.0 | 24.0 | 35.0 | 40.0 |
| FAO (2000) | Precipitation | LAS* | NA | 300 | 500 | 1,000 | 3,000 |

*HAS: High altitude sorghum, MAS: medium altitude sorghum, LAS: low altitude sorghum

**Table 4** Selected parameter set evaluation metrics for all evaluation datasets

| Database | Level* | TPR | FNR |
|---|---|---|---|
| Agro-MAPS | S | 0.974 | 0.026 |
| Agro-MAPS | D | 0.984 | 0.016 |
| CountrySTAT | S | 1.000 | 0.000 |
| CountrySTAT | D | 1.000 | 0.000 |
| FAOSTAT | C | 0.967 | 0.033 |
| ICRISAT | D | 1.000 | 0.000 |

**Table 5** Regional changes in suitability for each individual GCM

| Climate model | OSC* (%) | SCPIA* (%) | PIA* (km$^2$ x 10$^6$) | SCNIA* (%) | NIA* (km$^2$ x 10$^6$) |
|---|---|---|---|---|---|
| BCCR-BCM2.0 | 0.62 | 11.84 | 9.60 | -15.01 | 6.49 |
| CCCMA-CGCM3.1-T47 | 1.41 | 11.35 | 10.13 | -13.99 | 5.81 |
| CCCMA-CGCM3.1-T63 | 1.18 | 11.99 | 10.12 | -15.19 | 6.03 |
| CNRM-CM3.0 | 2.86 | 13.15 | 11.01 | -15.50 | 4.91 |
| CSIRO-MK3.0 | -1.76 | 11.64 | 9.17 | -19.95 | 7.24 |
| CSIRO-MK3.5 | -2.29 | 13.15 | 9.27 | -22.21 | 7.72 |
| GFDL-CM2.0 | -3.28 | 12.22 | 8.66 | -21.24 | 8.37 |
| GFDL-CM2.1 | -5.19 | 13.30 | 8.34 | -24.60 | 9.21 |
| GISS-AOM | -1.10 | 10.10 | 9.29 | -17.48 | 6.67 |
| GISS-MODEL-EH | 0.25 | 11.82 | 10.54 | -19.65 | 5.88 |
| GISS-MODEL-ER | -3.08 | 11.74 | 8.39 | -20.26 | 8.36 |
| IAP-FGOALS1.0-G | -7.32 | 6.29 | 5.12 | -18.34 | 10.93 |
| INGV-ECHAM4 | -5.44 | 8.72 | 6.95 | -19.93 | 9.30 |
| INM-CM3.0 | -1.64 | 13.04 | 10.21 | -23.73 | 6.96 |
| IPSL-CM4 | -0.35 | 10.90 | 10.19 | -20.03 | 5.81 |
| MIROC3.2-HIRES | -1.17 | 15.61 | 10.84 | -28.37 | 6.64 |
| MIROC3.2-MEDRES | 3.71 | 17.24 | 10.96 | -17.93 | 5.45 |
| MIUB-ECHO-G | -2.97 | 10.22 | 8.23 | -18.81 | 8.08 |
| MPI-ECHAM5 | -1.14 | 7.78 | 8.59 | -12.28 | 7.41 |
| MRI-GCGM2.3.2A | 0.60 | 6.37 | 9.66 | -8.64 | 5.39 |
| NCAR-CCSM3.0 | 1.72 | 15.31 | 10.94 | -22.03 | 5.64 |
| NCAR-PCM1 | -29.92 | 15.68 | 7.74 | -63.17 | 12.81 |
| UKMO-HadCM3 | -4.04 | 12.26 | 8.63 | -22.78 | 8.62 |
| UKMO-HadGEM1 | -5.55 | 9.55 | 7.35 | -19.66 | 10.07 |

*OSC: overall suitability change, SCPIA: suitability change in positively impacted areas, PIA: amount of positively impacted area, SCNIA: suitability change in negatively impacted areas, NIA: amount of negatively impacted area.