# Effects of climate change on the distribution of the C4-photosynthesizing grass *Sorghum bicolor*

**Jordy G. van der Beek (s1670956)**

**11/12/2019**

## Table of Contents



*** all figures and tables can be found at the end of the report.

## Introduction:

*Sorghum bicolor* (L.) MOENCH is an annual grass species native to the eastern part of Africa. With many cultivated forms, *Sorghum* is the fifth most produced cereal species (FAOSTAT, 2019) and grows, nowadays, on all continents, except Antarctica, at various altitudes, day lengths, and precipitation and temperature regimes (Wagaw, 2019). Naturally, *Sorghum* is, due to its $C_4$-photosynthesis, well adapted to growing in warm climates and under the current low $CO_2$ levels (Christin *et al.*, 2008). However due to extensive cultivation over 7000 varieties of *S. bicolor* have been identified (Kangama and Rumei, 2005) and several 'improved' varieties better adapted to (semi-)arid and tropic environments are available on the market every year (Dicko *et al.*, 2006). Although, *S. bicolor* is grown in the United States, Australia, and some other developed nations as animal food only, in Africa and Asia this cereal crop is used both for human nutrition and animal feed (Dicko *et al.*, 2006).

In developing countries, it is estimated that over 300 million people rely on *Sorghum* as source of energy (Godwin and Gray, 2000). Climate change might threaten food security for these countries in the near future. Globally we are already experiencing that increasing temperatures, the change of precipitation patterns, and more extreme weather events already affecting crop yield (Mbow *et al.*, 2019), and human induced $CO_2$ rises to over a 700 pmm threshold, might lower the benefit $C_4$-photosynthising crops have in the tropics (Bamford, 1997).

This report will focus on the current and future trends in the distribution of *S. bicolor*. The yield of this cereal species is found to decrease with higher temperatures and decreased rainfall (McCarthy and Vlek, 2012; Sultan *et al.*, 2013). In Zambia the amount suitable area of *Sorghum* was found to remain the same, climate change will shift the range northwards (Wang, Tan and Sun, 2015). In this report similar range shifts are expected, because of its large precipitation and temperature tolerance of the different varieties of this crop species. Although certain geographical boundaries might exist.

## Method:

For this research (fig. 1) 138959 species observation records, with the TaxonKey: *Sorghum bicolor* (L.) MOENCH, were downloaded from the Global Biodiversity Information Facility (GBIF.org, 2019). Observations without coordinates were removed from the dataset, as well as observations with duplicate coordinates. This datafile was than imported in ArcGIS (version 10.6.1) and observations that were not located in terrestrial habitats were excluded from the datafile.

For the species distribution models, WorldClim (version 1.4) bioclimatic variables of the present and future were used. These variables had a resolution of 5 arcminutes (appr. 10 km at the equator). The climate change variables used were future bioclimate variables for the Representative Concentration Pathway (RCP) with a climatic forcing of 4.5 $W/m^2$ (RCP4.5), being the intermediate climate prediction with a 1.8 °C predicted temperature rise by 2100 (IPCC, 2013). Which is just in line with the United Nations Paris Agreement goal that the temperature increase above 2 °C above pre-industrial levels.

Bioclimatic variable data was prepared and cropped in RStudio (version 1.2.5019) and R (version 3.6.1) for species distribution modelling. Pairwise correlation between variables was tested via a Pearson's Correlation Test. Based on ecological characteristics of the species, and statistical testing (variation inflation factor, VIF, < 10), bioclimatic variables were selected for species distribution modelling.

Species distribution models were generated using Maxent (version 3.4.1). The results were 3-fold cross-validated. For SDM validation the Area Under the Curve (AUC) of the Receiver Operator Curve (ROC) is used. With a minimum threshold for reliability of AUC > 0.7. Jacknife test was used to specify the importance of each variable by measuring the model performance.

For comparing the present with the 2050 distribution for each cell in the matrix presences and absences calculated in RStudio and R. The *S. bicolor* was defined as present when the predicted habitat suitability was higher than the 'equal test sensitivity and specificity', otherwise the species was defined as absent from a matrix cell.

# Results

## Occurrence data

The largest part of the species observation data is from sub-Saharan Africa and southern Asia (fig. 2. This is in line with the species native range. Most observations, however, are cultivated lineages from both observations from agricultural fields and 'escaped' individuals.

The lack of observations in southwestern part of Africa (e.g. Namibia, Angola, Congo and Cameroon) is most likely not due to sample bias, because the map resembles the map of National Research Council (1996, p. 131). The lack of observations in Latin America cannot be explained by literature, agriculture with Sorghum is reasonably large there (Dicko *et al.*, 2006), and is therefore most likely explained by a major sample bias.

The dots on the islands of the Pacific Ocean and Australia might be misidentifications of other *Sorghum* species (Tropicos.org, 2019).

## Environmental data:

The Pearson's pairwise correlation test showed strong correlations between most temperature related (Bio1-Bio11) and precipitation relates variables (Bio12-Bio19) (tab. 1). Based on the ecological characteristics seven bioclimatic variables were selected: Bio1, Bio 6, Bio9, Bio12, Bio14, Bio15 and Bio17. Bio1 (annual mean temperature) and Bio12 (annual precipitation) were selected as being the most widely used variable for temperature and precipitation respectively. Bio6 (min temperature of coldest month) was selected because this might be of interest, because this species is from warm climates. Bio9 (mean temperature of driest quarter) was selected because of its importance for the evapotranspiration of $C_4$-fososynthising plants. Despite the fact that Bio6 and Bio9 strongly correlated with the Bio1 variable (PCC = 0.96 and 0.93 respectively). Bio14 (precipitation of driest month), Bio15 (precipitation seasonality (coefficient of variation)) were selected because of its relevance for tropical regions.

Of these seven variables the Variation Inflation Factor was calculated. After removing Bio6 (VIF = 29.74) all VIF's were less than 10. The five remaining variables are: Bio1, Bio9, Bio12, Bio14, Bio15, were used for the species distribution modelling.

## Model output and SDM projections

The results from modelling in MaxEnt were reliable and did significantly deviate from random species distribution (AUC = 0.732, fig. 3). Most of the contribution to the model (61.7%, 58.4 % with permutation) is the result of bioclimatic variable Bio1, followed by bio12 (25.9%, 34.9% with permutation). Together they contribute for almost 90 % of the models predicting capacity.

Also interesting are the species response curves (fig. 4). We see that in our model the probability of a suitable habitat is zero by extremely low mean annual temperatures (fig. 4a) but is higher at higher temperatures. On the other hand, does high annual precipitation means that habitats are less likely to be suitable habitats for *S. bicolor* than low annual precipitations (fig. 4c). We can see that the bioclimatic variables: Bio9, Bio14 and Bio15, have (almost) no effect on the probability of some location being suitable or not (fig. 4b,d,e). In contrast to than, these bioclimatic variables do show a trend when they were used in a new model where only the corresponding variable was taken into account (fig. 4g,i,j).

For calculating the species presence and absence in RStudio a logistic threshold for 'equal test sensitivity and specificity' of 0.494 was used. The best suitable habitats are predicted for both the present (fig. 5a) and the 2050 (RCP 4.5) (fig. 5b) situation (figure situation in sub-Saharan Africa, subtropical Asia, Australia, some central parts of South America and the Southern United States of America. A slight increase is predicted (fig. 6). We see that the *S. bicolor* is predicted to expand its range more than it will lose possible area. The new habitat is mainly as range extension from the present situation, with the most profound effects found in Europe, China and a band just below the Sahara Desert.

# Discussion and conclusion

## Trends in distribution patterns

In Africa we see a range shift northward towards the Sahara (fig. 6). This particular increase in suitable habitat can be explained by an slight increase of annual rainfall, which can be seen in a map of the RCP4.5 precipitation change map in the Intergovernmental Panel on Climate Change report (IPCC, 2013). This in combination with the predicted effect of precipitation on habitat suitability *of S. bicolor* (fig. 4c), i.e. that the habitat tends to be more suitable when the annual rainfall is not too large, but larger than no rainfall at all.

We see a large increase in area in China. This is an interesting development. In China *Sorghum* is started to become one of the most important crops, being the most used grain for alcoholic beverages, and cultivated mostly in the north (Kangama and Rumei, 2005). Although not so many observation data were present for this area, the northern part of China will gain a large amount of suitable area.

In Europe, many small suitable patches are new when the present and future map are compared (fig. 6). These patches are not connected. However, this species is a crop species and distributed by humans, therefore natural dispersal limitations are no issue. In Europe Sorghum can be cultivated as

Reductions in suitable habitat are edge effects of the distribution and significant locations.

In conclusion, most of the nowadays cultivated area, is suitable in 2050 in a RCP4.5 scenario. Slightly more habitat is predicted to be suitable in the future. The predicting capacity is the best for Africa, where the distribution resembles the current distribution found in literature (National Research Council, 1996, p. 131).

## Reliability of the model

For the validation of the model we used the AUC. However, there is no true absence data entered in the model. The model used pseudo-absences for the validation process. With a crop species with such wide ecological preferences it is hard to say that these pseudoabsences reflect true absences. Therefore, using the AUC for validation of the model is not the best method.

Although, not so many occurrences were scored in Latin America and China, the model does reasonably large predictions for these areas. The reliability of the model for these area's is therefore low.

Another In the model certain bioclimatic variables were used in the model although strongly correlated. Some of them did not contribute much to the performance of the model. More careful selection of bioclimatic variable, based on both statistical and ecological arguments, would probably strengthen the predicting performance of the model.

The model predicts the habitat suitability via climatic variables, however in the developed countries humans alter fields suitability for certain crop species in various ways. On top of natural precipitation, farmers may use irrigation to provide enough water for the plants. Soil temperature can be higher on agricultural fields than under natural conditions by lowering ground water tables. Protective blankets are used to protect seedlings for the cold in early spring. These agricultural measurements provide artificial suitable conditions for *S. bicolor*, although the model links this species occurrence to natural bioclimatic variables. This can lead to an overestimation of its fundamental niche.

The future predictions of the model take only the occupied area into account. However, as mentioned in the introduction more than 7,000 varieties of *S. bicolor* are available on the market. The change in usage of certain varieties or the development of new, better adapted, varieties will change the habitat preferences of this species and therefore places where the crop will grow in the future.

In conclusion, the methodology used in this report is widely used for species distribution modelling of wild organisms. For the above-mentioned reasons, using these methodologies for (widely used) crop species will not result in the most reliable and useful output.

## References:

Bamford, A. (1997) 'Plants and climate change', *Biological Sciences Review*, 10(2), pp. 18–23.

Christin, P.-A. *et al.* (2008) 'Oligocene CO2 decline promoted C4 photosynthesis in grasses.', *Current biology : CB*. Elsevier, 18(1), pp. 37–43. doi: 10.1016/j.cub.2007.11.058.

Dicko, M. H. *et al.* (2006) 'Sorghum grain as human food in Africa: Relevance of content of starch and amylase activities', *African Journal of Biotechnology*, 5(5), pp. 384–395. doi: 10.5897/AJB.

FAOSTAT, (Food and Agriculture Organization of the United Nation) (2019) *FAOSTAT Database.* Available at: http://faostat.fao.org/site/291/default.aspx (Accessed: 8 December 2019).

GBIF.org (2019) 'GBIF Occurrence'. doi: 10.15468/dl.gsqtyi.

Godwin, I. and Gray, S. (2000) 'Overcoming productivity and quality constraints in sorghum: the role for genetic engineering', in O'Brien, L. and Henry, R. J. (eds) *Transgenic cereals*. St. PAul, Minnesota, USA: AACC publishing, pp. 153–177.

IPCC (2013) 'Summary for Policymakers', in Stocker, T. F. et al. (eds) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Chance.* Cambridge, United Kingdom and New York, NY, United States: Cambridge University Press.

Kangama, C. O. and Rumei, X. (2005) 'Introduction of sorghum (*Sorghum bicolor* (L.) Moench) into China', *African Journal of Biotechnology*, 4(7), pp. 575–579.

Mbow, C. *et al.* (2019) 'Food Security', in Shukla, P. R. et al. (eds) *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation,* *sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*. Intergovernmental Panel on Climate Change (IPCC), pp. 437–550.

McCarthy, D. S. and Vlek, P. L. G. (2012) ‘Impact of climate change on *Sorghum* production under different nutrient and crop residue management in semi-arid region of Ghana: A modeling perspective’, *African Crop Science Journal*, 20(Issue Supplement s2), pp. 243–259.

National Research Council (1996) *Lost Crops of Africa. Volume I: Grains.* Washington, D.C.: National Academy Press.

Sultan, B. *et al.* (2013) ‘Assessing climate change impacts on sorghum and millet yields in the Sudanian and Sahelian savannas of West Africa’, *Environmental Research Letters*, 8(1). doi: 10.1088/1748-9326/8/1/014040.

Tropicos.org (2019) *Tropicos | Sorghum Moench*, *Missouri Botanical Garden*. Available at: http://www.tropicos.org/Name/40002210 (Accessed: 9 December 2019).

Wagaw, K. (2019) ‘Review on Mechanisms of Drought Tolerance in Sorghum (*Sorghum bicolor* ( L .) Moench) Basis and Breeding Methods’, *Academic Research Journal of Agricultural Science and Research*, 7(March), pp. 87–99. doi: 10.14662/ARJASR2019.007.

Wang, Y., Tan, Z. and Sun, G. (2015) ‘The Impact of Climate Change on the Potential Suitable Distribution of Major Crops in Zambia and the Countermeasures’, in. Springer, Cham, pp. 460–472. doi: 10.1007/978-3-319-19620-6_52.

## Pictures and tables

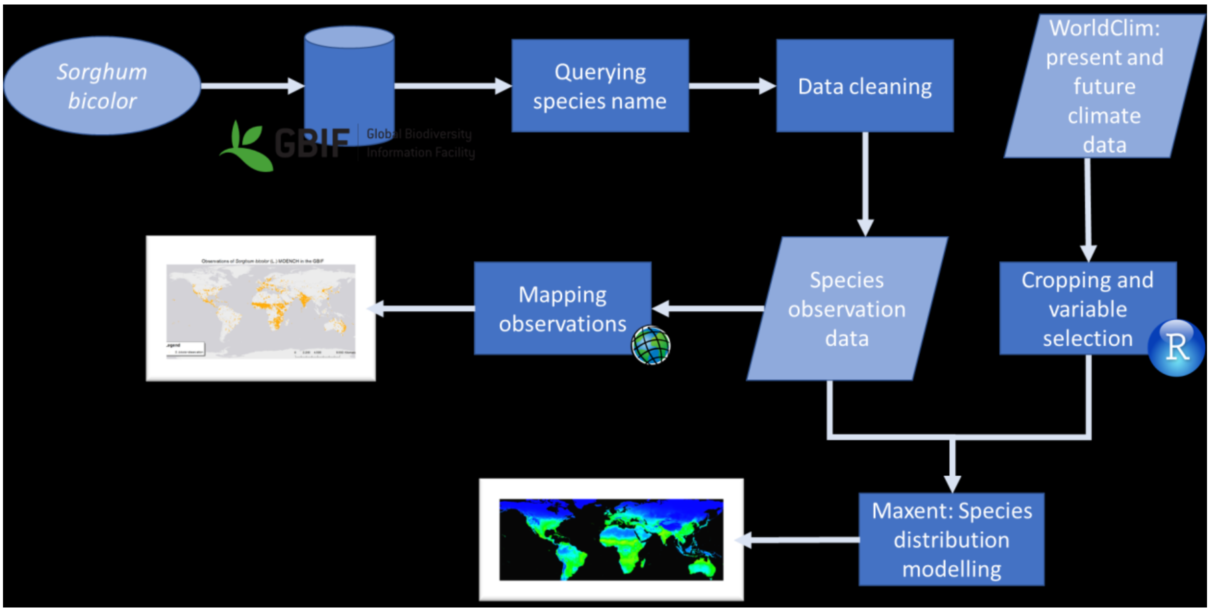


***Figure 1**. Schematic overview of the whole workflow.*

Observations of *Sorghum bicolor* (L.) MOENCH in the GBIF

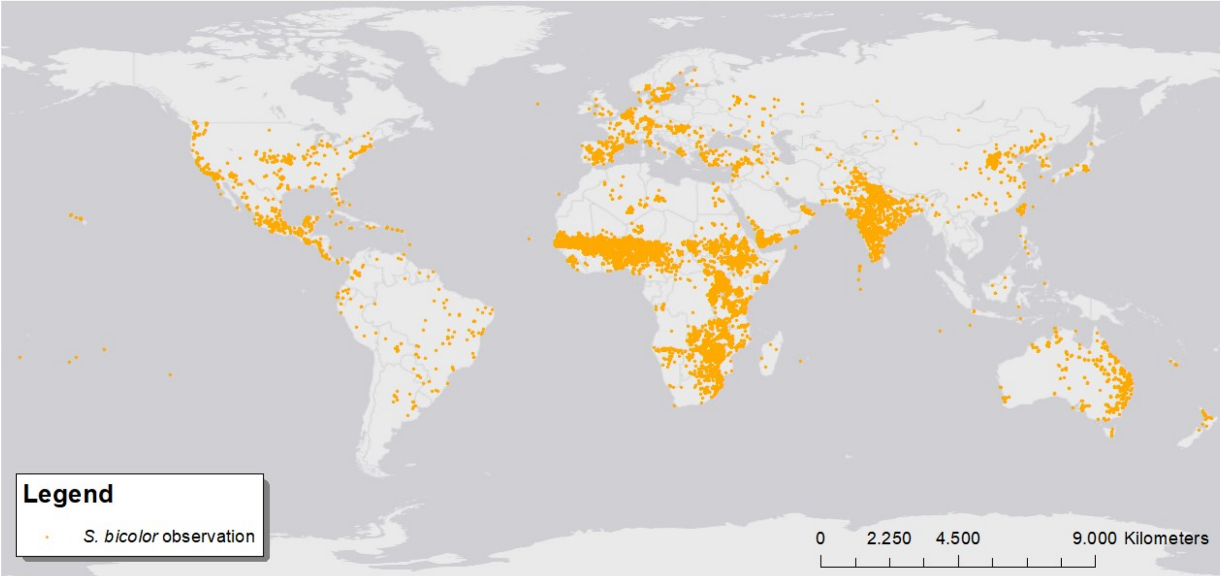


***Figure 2****. Spatial distribution of species observation data of Sorghum bicolor (L.) Moench in GBIF. Map produced in ESRI ArcMap software © 2019 OpenStreetMap contributors and the GIS user community*

***Table 1****. Pearson's Pairwise Correlation Matrix for all bioclimatic WorldClim variables. Green means a strong positive correlation (PCC > 0.7), red resembles a small positive correlation (PCC < -0.7). Bio1 = Annual Mean Temperature. Bio2 = Mean Diurnal Range (Mean of monthly (max temp - min temp)). Bio3 = Isothermality (Bio2/Bio7) (* 100). Bio4 = Temperature Seasonality (standard deviation *100). Bio5 = Max Temperature of Warmest Month. Bio6 = Min Temperature of Coldest Month. Bio7 = Temperature Annual Range (Bio5-Bio6). Bio8 = Mean Temperature of Wettest Quarter. Bio9 = Mean Temperature of Driest Quarter. Bio10 = Mean Temperature of Warmest Quarter. Bio11 = Mean Temperature of Coldest Quarter. Bio12 = Annual Precipitation. Bio13 = Precipitation of Wettest Month. Bio14 = Precipitation of Driest Month. Bio15 = Precipitation Seasonality (Coefficient of Variation). Bio16 = Precipitation of Wettest Quarter. Bio17 = Precipitation of Driest Quarter. Bio18 = Precipitation of Warmest Quarter. Bio19 = Precipitation of Coldest Quarter.*

| | **Bio01** | **Bio02** | **Bio03** | **Bio04** | **Bio05** | **Bio06** | **Bio07** | **Bio08** | **Bio09** | **Bio10** | **Bio11** | **Bio12** | **Bio13** | **Bio14** | **Bio15** | **Bio16** | **Bio17** | **Bio18** | **Bio19** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Bio01** | 1.00 | 0.28 | 0.81 | -0.86 | 0.85 | 0.97 | -0.79 | 0.74 | 0.93 | 0.91 | 0.98 | 0.34 | 0.40 | 0.05 | 0.32 | 0.40 | 0.08 | 0.13 | 0.24 |
| **Bio02** | | 1.00 | 0.20 | -0.06 | 0.53 | 0.12 | 0.15 | 0.25 | 0.24 | 0.39 | 0.20 | -0.43 | -0.29 | -0.46 | 0.48 | -0.31 | -0.46 | -0.38 | -0.36 |
| **Bio03** | | | 1.00 | -0.90 | 0.50 | 0.86 | -0.85 | 0.57 | 0.76 | 0.56 | 0.87 | 0.55 | 0.55 | 0.24 | 0.23 | 0.55 | 0.27 | 0.31 | 0.43 |
| **Bio04** | | | | 1.00 | -0.48 | -0.95 | 0.97 | -0.50 | -0.85 | -0.57 | -0.94 | -0.53 | -0.54 | -0.23 | -0.18 | -0.54 | -0.25 | -0.29 | -0.39 |
| **Bio05** | | | | | 1.00 | 0.71 | -0.36 | 0.71 | 0.77 | 0.98 | 0.74 | 0.00 | 0.11 | -0.18 | 0.39 | 0.09 | -0.17 | -0.15 | 0.00 |
| **Bio06** | | | | | | 1.00 | -0.91 | 0.64 | 0.93 | 0.79 | 0.99 | 0.46 | 0.49 | 0.17 | 0.22 | 0.48 | 0.20 | 0.21 | 0.35 |
| **Bio07** | | | | | | | 1.00 | -0.44 | -0.80 | -0.48 | -0.89 | -0.61 | -0.58 | -0.33 | -0.06 | -0.59 | -0.36 | -0.36 | -0.46 |
| **Bio08** | | | | | | | | 1.00 | 0.50 | 0.77 | 0.67 | 0.24 | 0.34 | -0.04 | 0.40 | 0.32 | -0.02 | 0.22 | 0.05 |
| **Bio09** | | | | | | | | | 1.00 | 0.81 | 0.94 | 0.31 | 0.34 | 0.08 | 0.22 | 0.34 | 0.11 | 0.04 | 0.28 |
| **Bio10** | | | | | | | | | | 1.00 | 0.81 | 0.12 | 0.21 | -0.09 | 0.36 | 0.19 | -0.07 | -0.03 | 0.08 |
| **Bio11** | | | | | | | | | | | 1.00 | 0.42 | 0.47 | 0.12 | 0.27 | 0.46 | 0.15 | 0.19 | 0.30 |
| **Bio12** | | | | | | | | | | | | 1.00 | 0.89 | 0.70 | -0.21 | 0.92 | 0.74 | 0.78 | 0.74 |
| **Bio13** | | | | | | | | | | | | | 1.00 | 0.39 | 0.11 | 0.99 | 0.42 | 0.72 | 0.57 |
| **Bio14** | | | | | | | | | | | | | | 1.00 | -0.53 | 0.42 | 0.99 | 0.55 | 0.66 |
| **Bio15** | | | | | | | | | | | | | | | 1.00 | 0.06 | -0.53 | -0.14 | -0.29 |
| **Bio16** | | | | | | | | | | | | | | | | 1.00 | 0.46 | 0.74 | 0.60 |
| **Bio17** | | | | | | | | | | | | | | | | | 1.00 | 0.57 | 0.69 |
| **Bio18** | | | | | | | | | | | | | | | | | | 1.00 | 0.34 |
| **Bio19** | | | | | | | | | | | | | | | | | | | 1.00 |

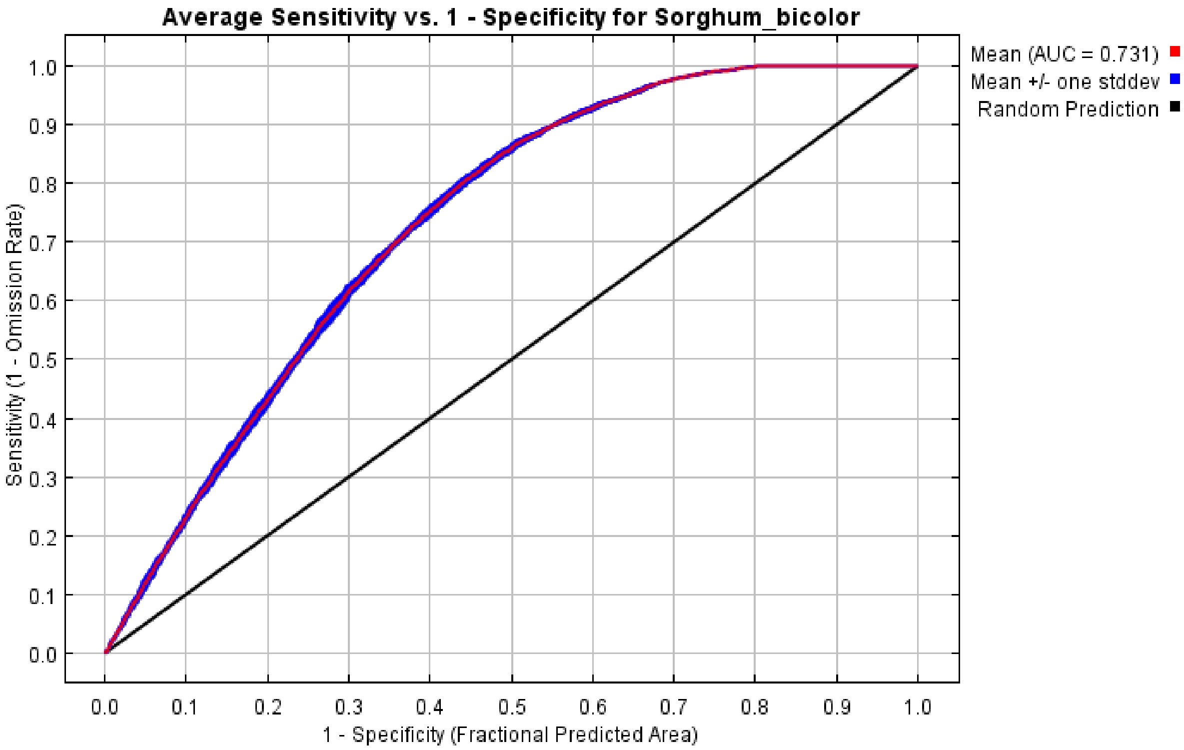


***Figure 3***. *The receiver operating characteristic (ROC) curve for the model.*

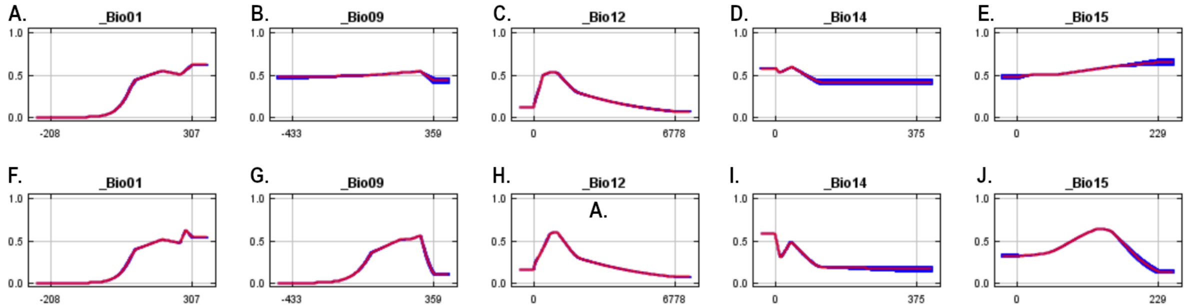


***Figure 4**. Species response curves. The x-axis contains the range of the bioclimatic variables; however, these are in the WorldClim version 1.4 sometimes hard to interpret. The y-axis shows the probability that this value of the bioclimatic variable is a suitable condition. The curves A-E show how each environmental variable affects the Maxent prediction in our model. The curves F-G show their effect in a Maxent model created using only the corresponding variable.*

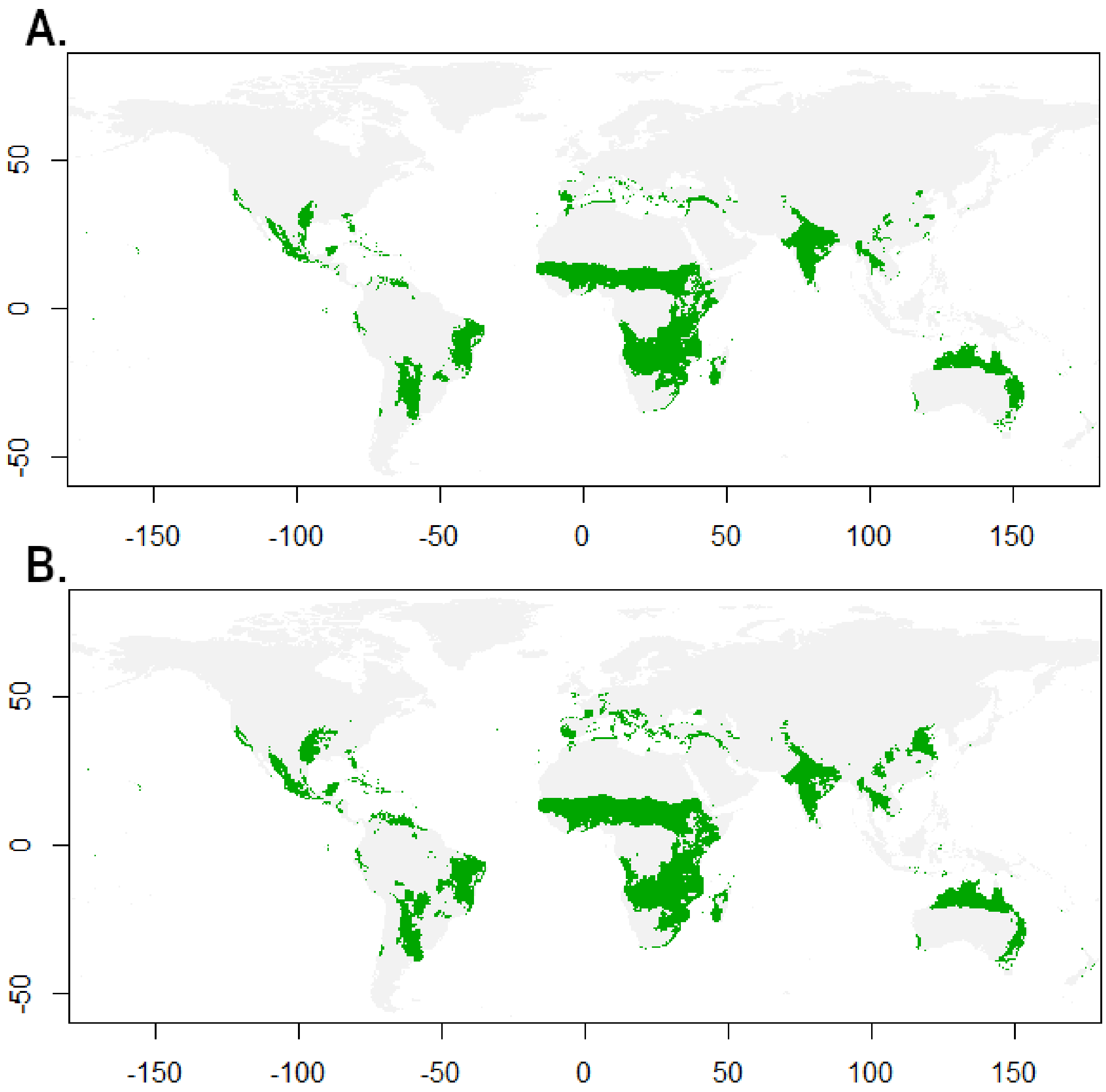


***Figure 5****. Global predicted species distribution for Sorghum bicolor (L.) MOENCH based on habitat suitability. The green area is the area with a predicted habitat probability above the ‘equal test sensitivity and specificity’-treshold. A. Is the predicted present situation. B. Is the predicted distribution in 2050 based for the RCP 4.5 scenario.*

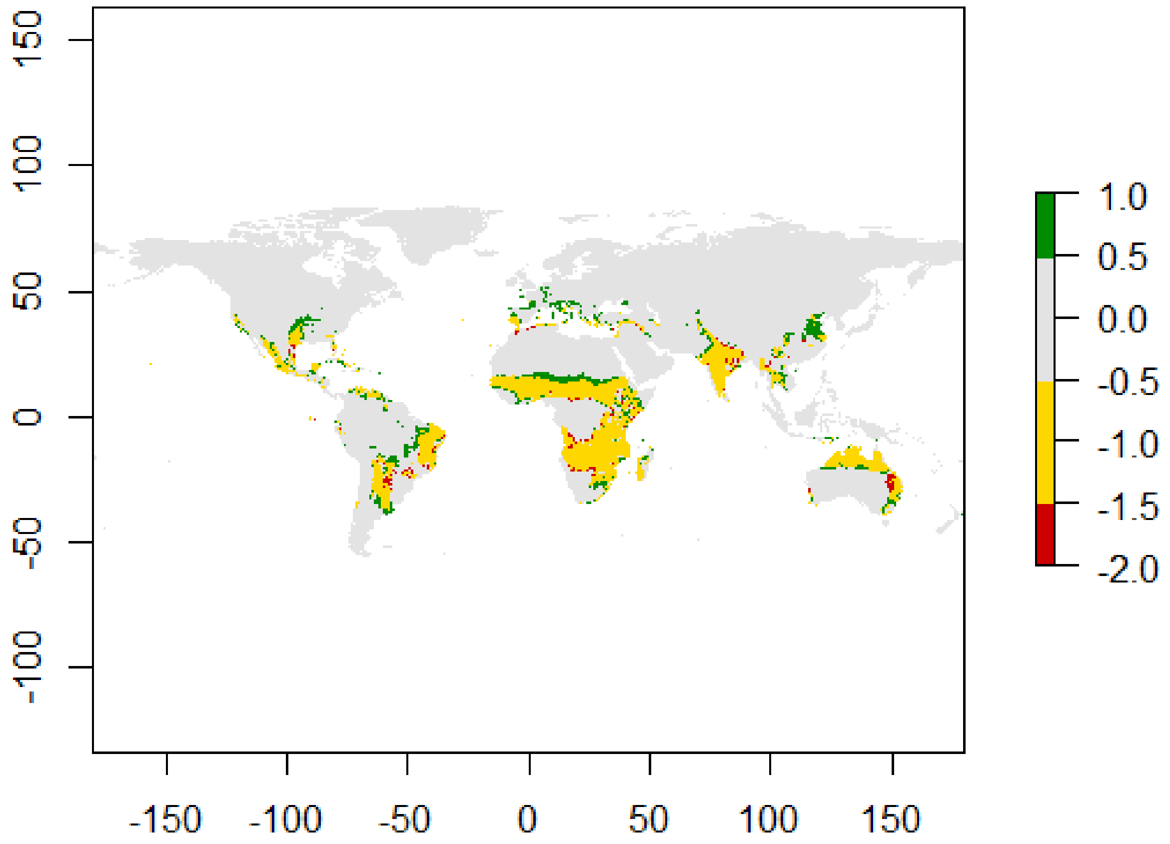


***Figure 6***. *Change in habitat suitability map for S. bicolor between the present and the year 2050 (under RPC4.5 conditions). The predicted loss of territory is shown in red, and green indicates a predicted gain of occurrence area. Yellow are habitat conditions that remain unchanged.*