

# Projected habitat change of *Manihot esculenta* under different RCP scenarios

Koen Rurenga (s1235869)

---

## Introduction

The Cassava (*Manihot esculenta*) is a woody shrub of great economic importance in the tropics. It is of particular importance in agriculture for developing countries, since it does well on poor soil and in low rainfall conditions. It is the primary food staple of 800 million people worldwide (Stone, 2011).

Luckily, *M. esculenta* is very durable and can stand prolonged periods of drought, sometimes even increasing yield when reapplying water after an induced drought (Connor, Cock and Parra, 1981; El-Sharkawy and Cadavid, 2002). Next to drought, cassava also has a high range of optimum temperature, ranging from 25 to 35°C (El-sharkawy, Pilar and Hershey, 1992). Only below 17°C and above 37°C, the sprouting of shoots becomes delayed (Keating and Everson, 1979).

These characteristics make cassava a crop of interest in the currently changing world, with agricultural conditions projected to deteriorate at various levels. This report examines the different RCP's (Representative Concentration Pathways), set by the IPCC. These scenarios project the change in greenhouse gases and the resulting climatic variables, being rising temperatures and precipitation level. They are measured in radiative forcing, the difference between energy absorbed by the earth and the energy being radiated back into space. For example, RCP2.6 projects a radiative forcing of 2.6W/m<sup>2</sup> in 2100. The scenarios examined are RCP2.6, RCP4.5, RCP6.0 and RCP8.0.

Since *M. esculenta* is such a hardy crop, the working hypothesis is that climate change will not have a significant effect on the suitable habitats of cassava. This will be explored by mapping occurrence data from GBIF into ArcGIS and modelling this data with the WorldClim 1.4 data using MaxEnt.

## Methodology & Results

### *Data Carpentry*

First, the data was collected from GBIF.org. A search was done on *Manihot esculenta*, and data without coordinates was excluded.

The data is cleaned up using Microsoft Excel: all data is removed except for species name, latitude and longitude. Any double observations (species at the exact same longitude and latitude) are removed. Finally, the columns 'Longitude' and 'Latitude' are swapped in order to format the data for further analysis.

The cleaned data was imported into ArcGIS (Figure 1). Using ArcGIS, any occurrence data in locations such as the sea are removed. Important to note is that botanical observations have not been removed from the occurrence map. This could have influence on the modelling by MaxEnt. However, due to time constraints, it was not possible to go back to ArcGIS and remove these observations. Aside from the botanical observations, cassava mainly grows in South America and Africa. Interestingly, despite the hardiness of the species, they grow mainly around the equator. This might be because of more socio-economic reasons, as Thailand and Ghana are among the world's largest cassava producers, being able to export almost enough for the worldwide demand (FAOSTAT).

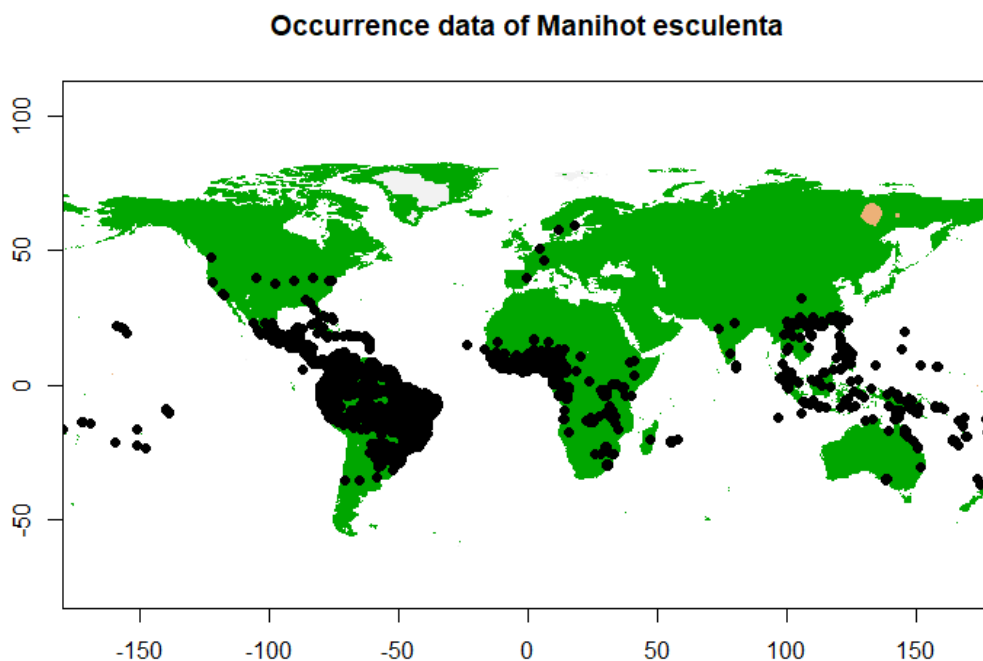


Figure 1: Occurrence data of *Manihot Esculenta* recovered from GBIF.org. A black dot represents an observation.

The occurrence data is saved as a .csv file and opened in RStudio. This data is formatted for use in MaxEnt by transforming it into Raster data together with bioclimatic data collected from WorldClim.org. The data from WorldClim 1.4 was used for present knowledge and future scenarios. Scenarios downloaded and prepared were RCP2.6, RCP 4.5, RCP6.0 and RCP8.0. This is done by using the manual provided, found on:

<https://github.com/naturalis/mebioda/tree/master/doc/week2/w2d5>.

After preparing the environmental data for MaxEnt, specific bioclimatic variables were selected for analysis. Originally, the bioclimatic variables selected were visible in the table below.

Table 1: Original chosen Bioclimatic Variables for analysis

Bioclimatic variable number	Description
<b>Bio2</b>	Mean Diurnal Range (mean of monthly (max temp- min temp))
<b>Bio4</b>	Temperature seasonality (SD * 100)
<b>Bio6</b>	Min Temperature of Coldest Month
<b>Bio13</b>	Precipitation of Wettest Month
<b>Bio14</b>	Precipitation of Driest Month
<b>Bio17</b>	Precipitation of Driest Quarter

Understandably, when testing, Bio14 and Bio17 resulted in an extremely high correlation (Figure 2). After all, they concern very similar data- precipitation during a specific time frame in the year. Therefore, Bio17 was removed from this list.

	A	B	C	D	E	F	G	H
1		Bio02	Bio04	Bio06	Bio13	Bio14	Bio17	
2	Bio02	1	0,0256039	0,0076382	-0,363576	-0,472597	-0,48672041	
3	Bio04	0,0256039	1	-0,936094	-0,536694	-0,219844	-0,24826677	
4	Bio06	0,0076382	-0,936094	1	0,4721643	0,1737880	0,200108577	
5	Bio13	-0,363576	-0,536694	0,4721643	1	0,3832343	0,418596122	
6	Bio14	-0,472597	-0,219844	0,1737880	0,3832343	1	0,993706492	
7	Bio17	-0,486720	-0,248266	0,2001085	0,4185961	0,9937064	1	
8								

Figure 2: Correlation table of the chosen Bioclimatic variables

When testing for multicollinearity, Bio4 and Bio6 had a rather high VIF of over 8. While it is not too surprising when looking at the correlation table (Figure 2), it was not removed yet. The VIF results confirmed the need for removing either Bio4 or Bio6. The removed variable was Bio6 (Temperature seasonality)

Table 2: Final chosen Bioclimatic Variables for analysis

Bioclimatic variable number	Description
<b>Bio2</b>	Mean Diurnal Range (mean of monthly (max temp- min temp))
<b>Bio4</b>	Temperature seasonality (SD * 100)
<b>Bio13</b>	Precipitation of Wettest Month
<b>Bio14</b>	Precipitation of Driest Month

After testing for multicollinearity again, the following results were produced:

Variables	VIF
1 Bio02	1.439739
2 Bio04	1.492140
3 Bio13	1.738005
4 Bio14	1.396854

These low VIF values indicate that these factors would be interesting to work with using MaxEnt.

### Model Settings

MaxEnt was made to produce Response Curves in order to see the modelled effect of each environmental variable on *M. esculenta*.

Furthermore, MaxEnt is asked to make pictures of the predictions so the data is better interpretable.

A Jackknife was done to measure the importance of every variable.

A random seed was used to make sure every replication MaxEnt produces is a newly randomised model. MaxEnt was made to run 5 replications and summarize these in a larger file.

Extrapolating was ticked in order to simulate the species over the entire world.

### Model Output

Running MaxEnt resulted in the following data.

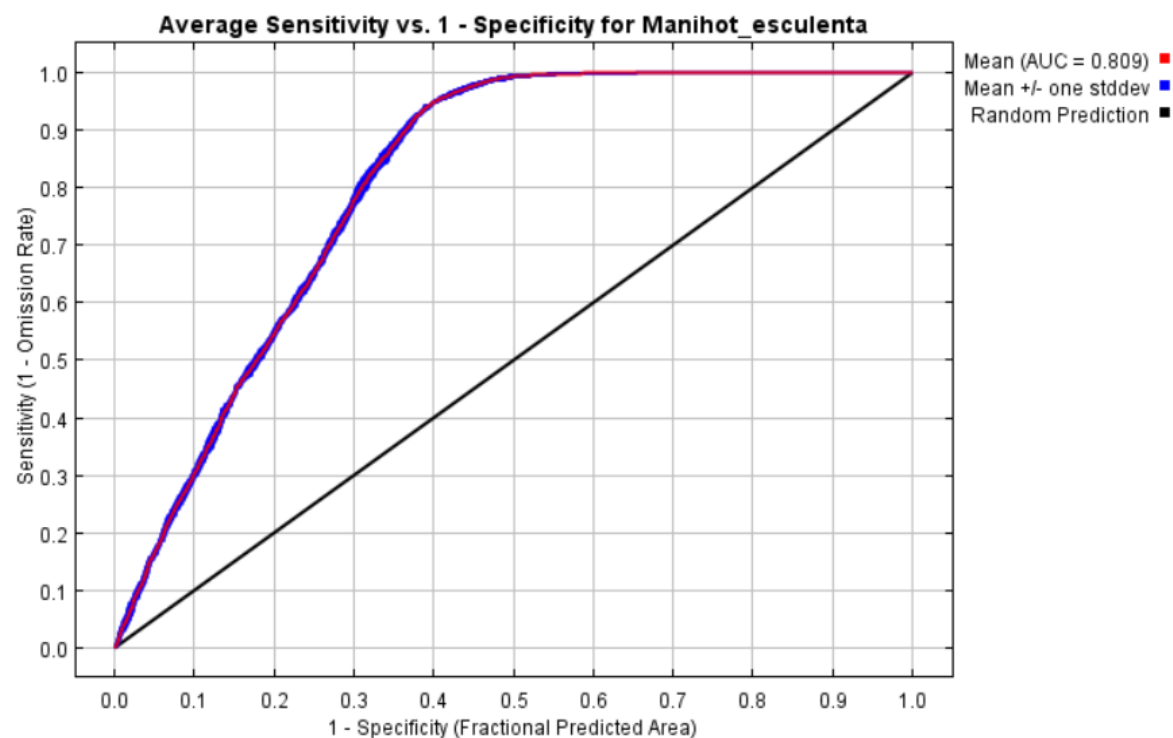


Figure 3: ROC curve of the modeled data for *M. esculenta*

The mean AUC for the 5 replications is 0.809 with a standard deviation of 0.002. This indicates that the data is modelled well enough and there is minimized risk of false positives.

The Response curves for each bioclimatic variable show the influence of every variable on the predicted probability of occurrence predictions by MaxEnt. Looking at these curves for both the complete model and the models for individual bioclimatic variable analyses, factor Bio4 (Temperature seasonality) is predicted by MaxEnt to have the most influence in whether an area is suitable for *Manihot esculenta* to grow.

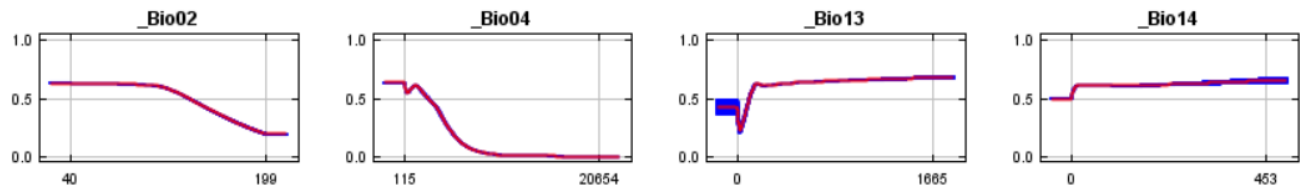


Figure 4: Response curves for all variables under full model analysis

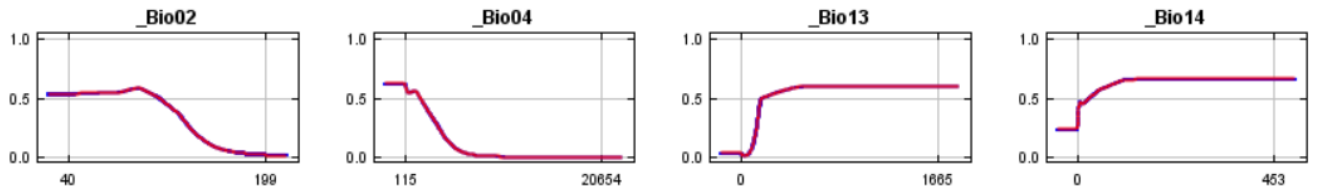


Figure 5: Response curves for all variables, modelled independently from one another.

This graphical data is confirmed by the Jackknife analysis. The environmental variable with the highest gain when used in isolation is Bio4, which therefore appears to have the most useful information by itself.

The data results from MaxEnt were once again imported into R, in order to make changemaps. These resulted in the following maps. A threshold was set in order to create these changemaps. This threshold was calculated using the mean Maximum training sensitivity plus specificity of every replicate. This resulted in a threshold of 0.3254.

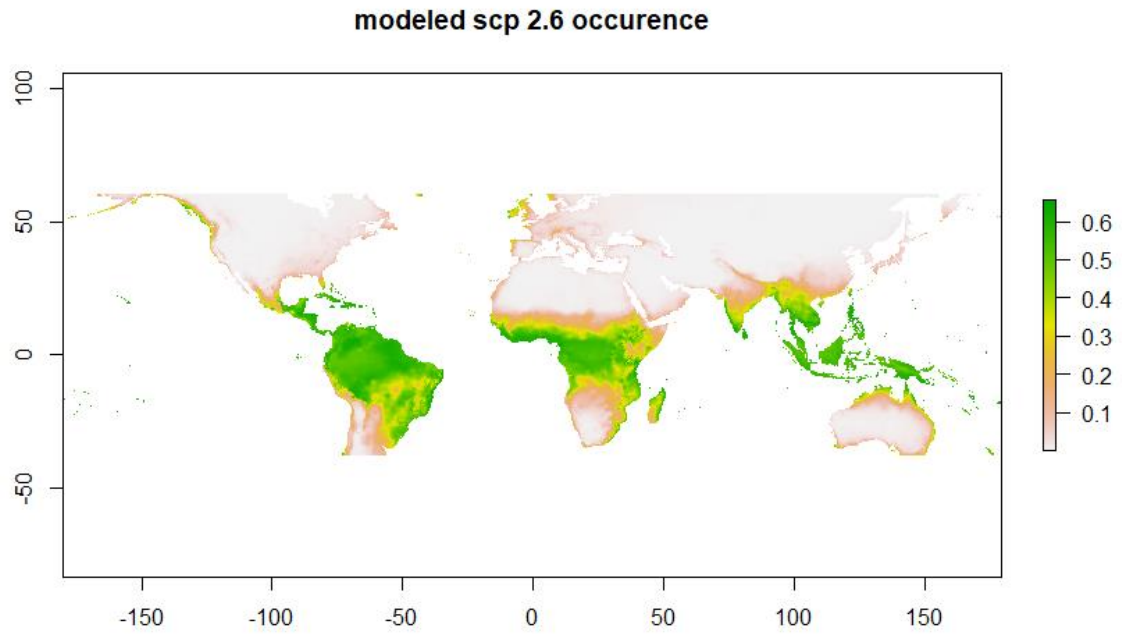


Figure 6: Modeled occurrence data for *M. esculenta* under RCP scenario 2.6. Greener areas indicate a higher probability of occurrence. Redder areas indicate a lower chance of occurrence. White areas indicate no chance of occurrence.

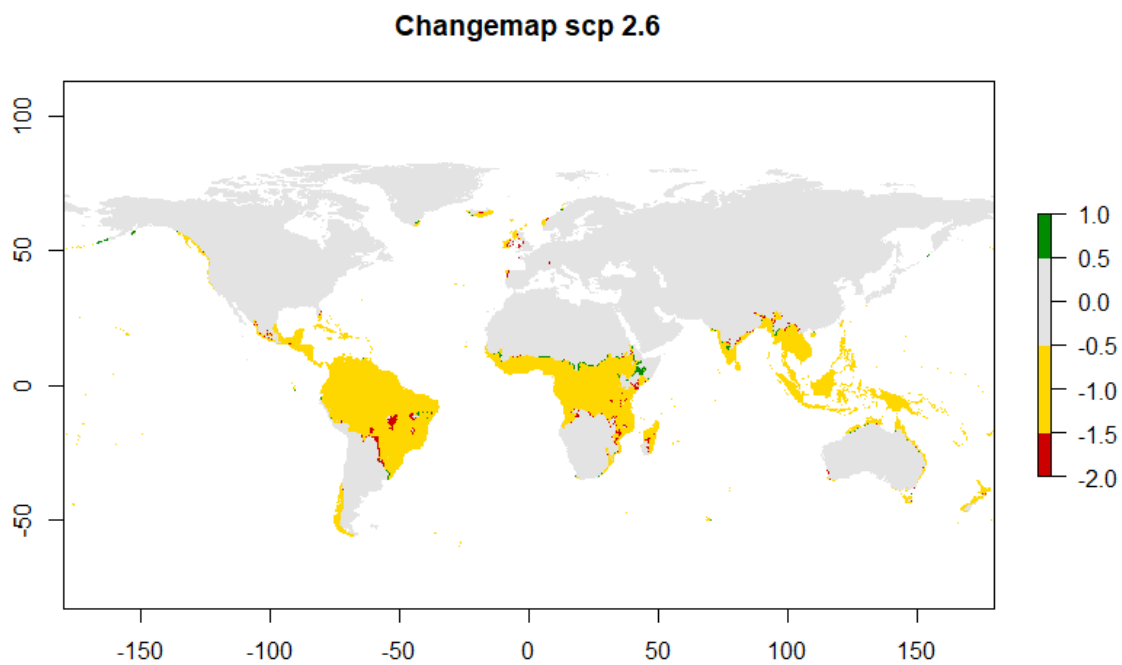


Figure 7: Changemap for occurrence data of *M. esculenta* under RCP Scenario 2.6. Yellow areas indicate no change. Red areas indicate a loss of occurrence area. Green areas indicate a gain in occurrence area.

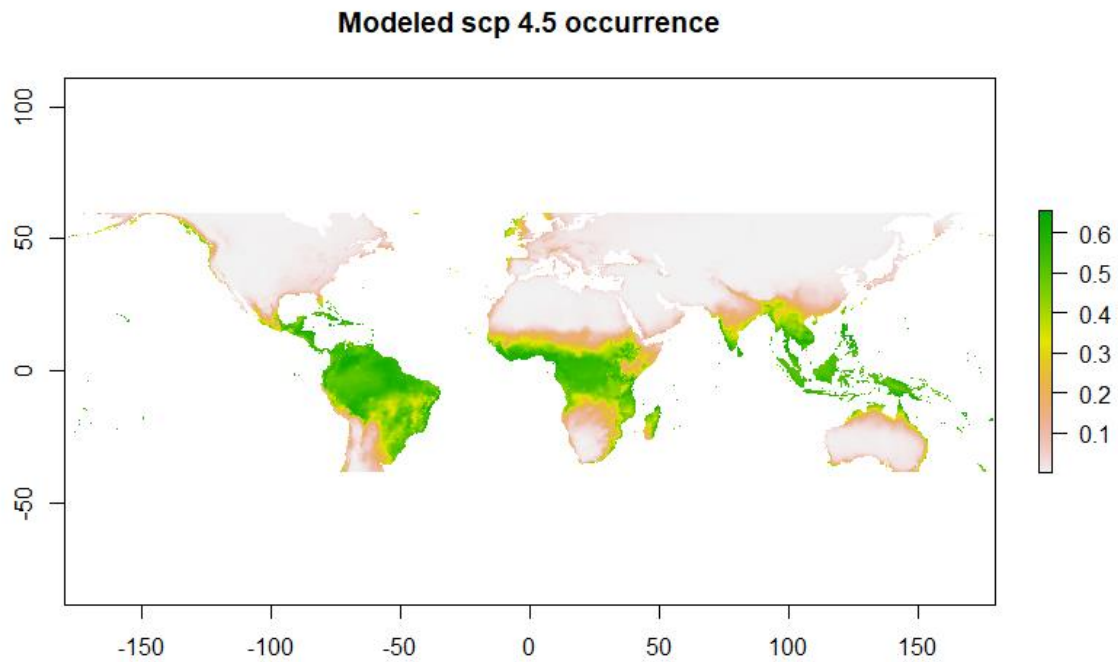


Figure 8: Modeled occurrence data for *M. esculenta* under RCP scenario 4.5. Greener areas indicate a higher probability of occurrence. Redder areas indicate a lower chance. White areas indicate no chance of occurrence.

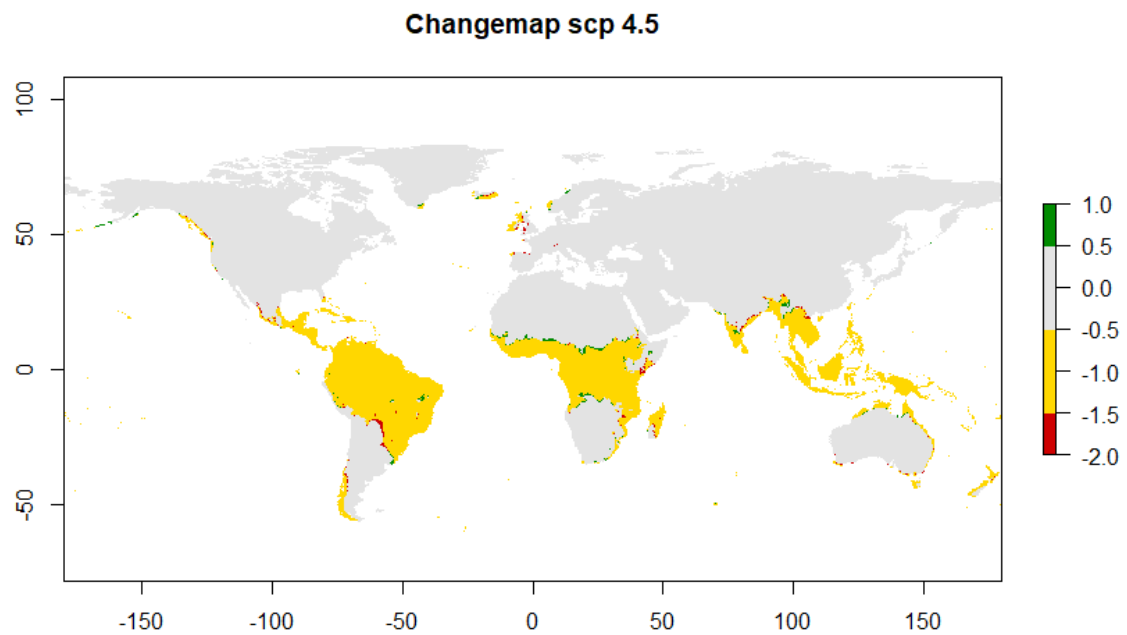


Figure 9: Changemap for occurrence data of *M. esculenta* under RCP Scenario 4.5. Yellow areas indicate no change. Red areas indicate a loss of occurrence area. Green areas indicate a gain in occurrence area.

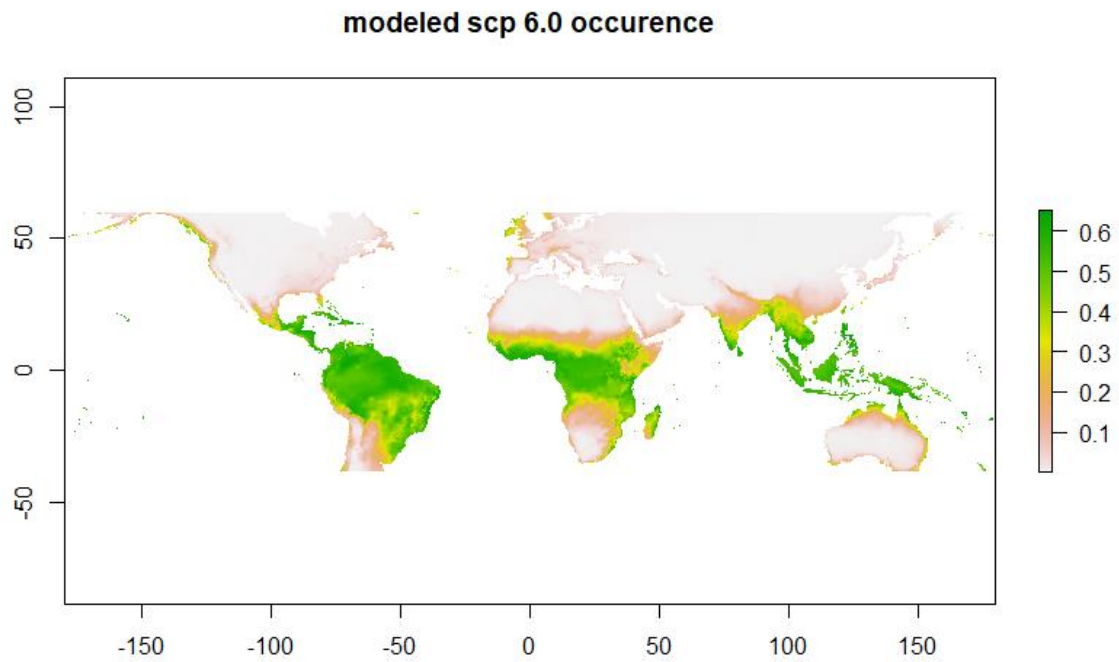


Figure 10: Modeled occurrence data for *M. esculenta* under RCP scenario 6.0. Greener areas indicate a higher probability of occurrence. Redder areas indicate a lower chance. White areas indicate no chance of occurrence.

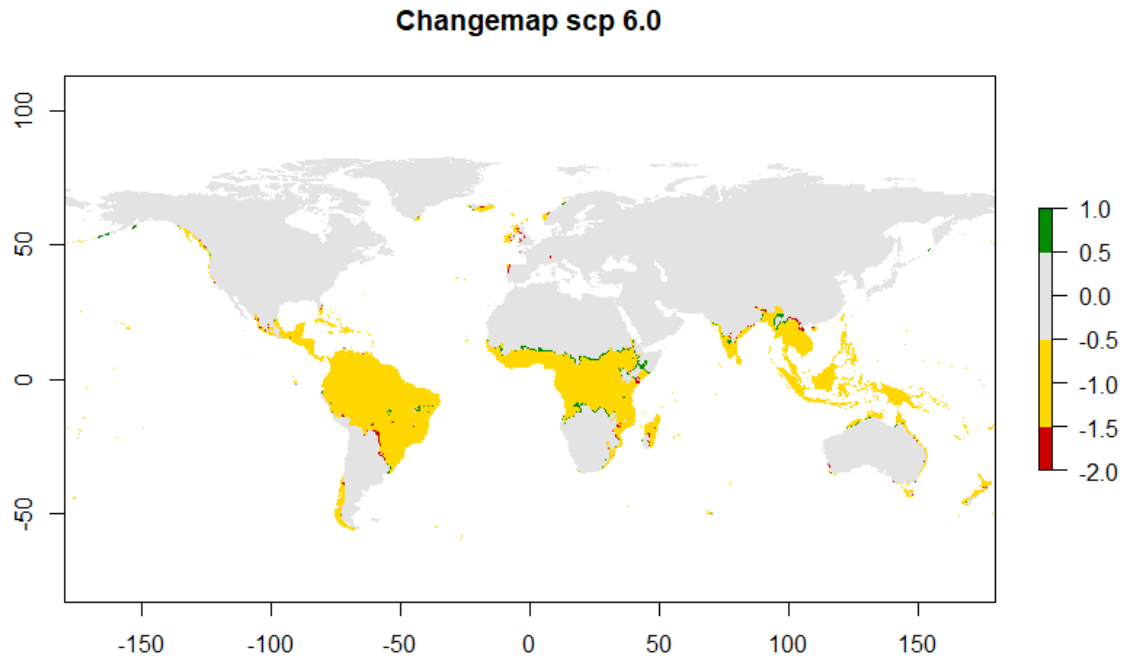


Figure 11: Changemap for occurrence data of *M. esculenta* under RCP Scenario 6.0. Yellow areas indicate no change. Red areas indicate a loss of occurrence area. Green areas indicate a gain in occurrence area.



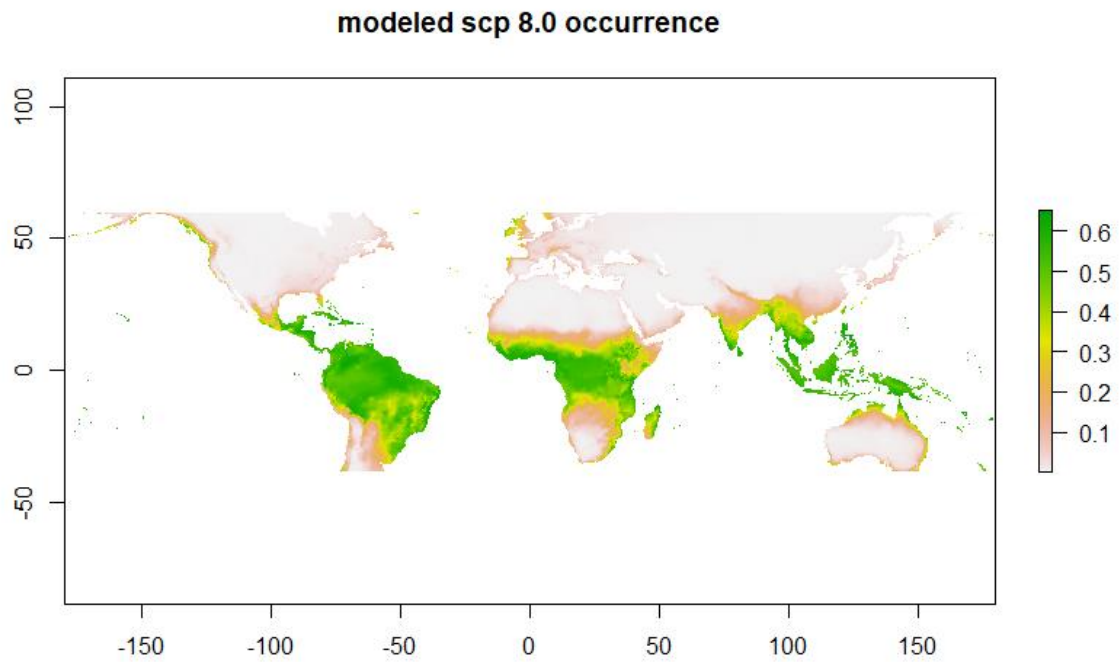


Figure 12: Modeled occurrence data for *M. esculenta* under RCP scenario 8.0. Greener areas indicate a higher probability of occurrence. Redder areas indicate a lower chance. White areas indicate no chance of occurrence.

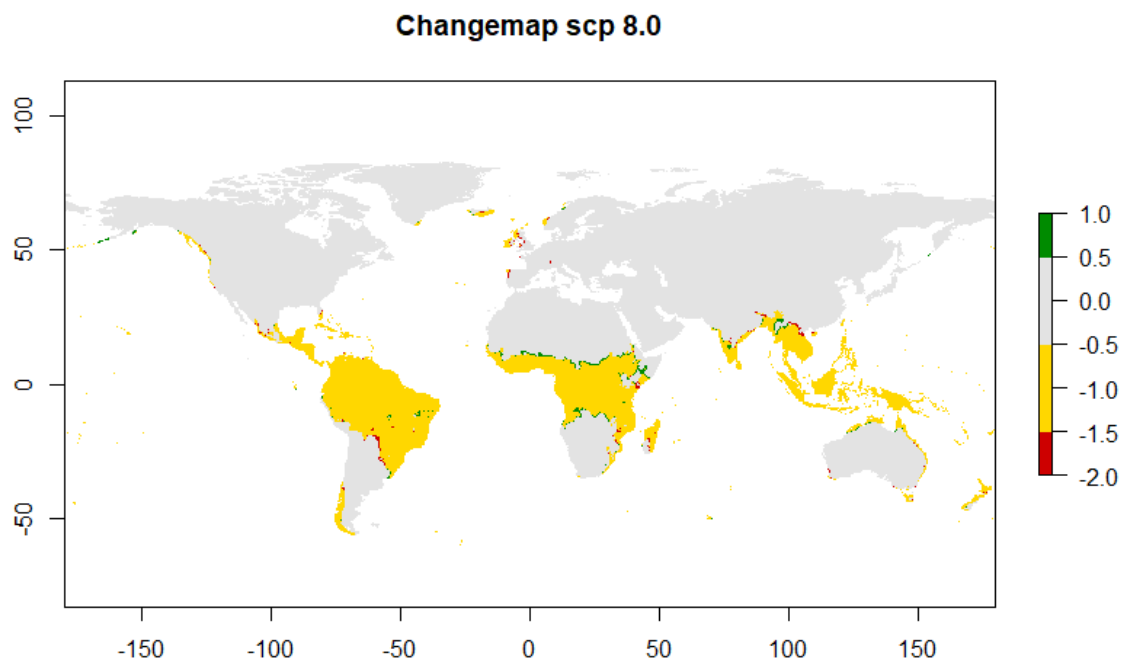


Figure 13: Changemap for occurrence data of *M. esculenta* under RCP Scenario 8.0. Yellow areas indicate no change. Red areas indicate a loss of occurrence area. Green areas indicate a gain in occurrence area.

## Discussion & conclusion

As can be inferred from the changemaps, there is little to no change expected in the occurrence of *M. esculenta* under any RCP scenario. This was already as expected- Cassava is a hardy crop able to withstand a wide range of temperatures and long periods of drought.

Something to consider in order to improve the modelling for this species, is the removal of botanical observations. It would be interesting to rerun the model without these observations. However, a specific bias might occur when doing so, since cassava is very hardy but is almost entirely cultivated around the equator. Whether the removal of these observations is beneficial or detrimental to the species distribution model is therefore up for debate.

## Bibliography

Connor, D. J., Cock, J. H. and Parra, G. E. (1981) 'Response of cassava to water shortage I. Growth and yield', *Field Crops Research*, 4(C), pp. 181–200. doi: 10.1016/0378-4290(81)90071-X.

El-sharkawy, B. M. A., Pilar, A. D. E. L. and Hershey, C. (1992) 'Yield Stability of Cassava during prolonged mid-season stress', 28, pp. 165–174.

El-Sharkawy, M. A. and Cadavid, L. F. (2002) 'Response of cassava to prolonged water stress imposed at different stages of growth', *Experimental Agriculture*. Universiteit Leiden / LUMC, 38(3), pp. 333–350. doi: 10.1017/S001447970200306X.

Keating, B. . and Everson, J. . (1979) 'Effect of Soil Temperature on Sprouting and Sprout', *Field Crops Research*, 2, pp. 241–251.

Stone, G. D. (2011) 'Fallacies in the Genetic-Modification Wars, Implications for Developing Countries, and Anthropological Perspectives', *Current Anthropology*, 43(4), pp. 611–630.