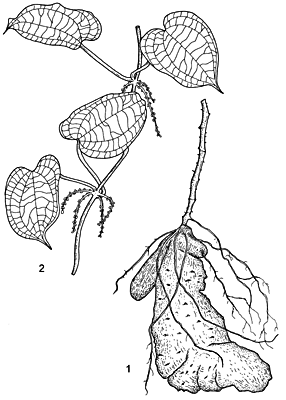
SDM report of *Dioscorea cayenensis*, Mees Hoffmanns, s1866923

**INTRODUCTION:**

The Guinea yellow Yam is an agricultural plant that can mostly be found in the tropical regions of West-Africa, like Nigeria, Ghana and Benin. Since it has originated through cultivation, and does not grow anywhere in the wild, it depends heavily on human aid to spread into new areas. The earliest report of this species was by Jean Baptiste de Lamarck in South America, in 1789. As of right now, only a few countries in South-America and Oceania have been introduced to it. It is not invasive, but might possibly get that status in Cuba in the near future (Oviedo et al 2012). It is a staple food crop in Africa due to its starch-filled tuber, and also has significant cultural value in Guinea (ornament).(PROTA, 2017)

An official physical description of the plant is listed below:

“Dioecious, exceptionally monoecious, glabrous herb with annual twining stem arising from tuber; tuber usually solitary, cylindrical to irregularly shaped, up to 10(-25) kg in weight, flesh whitish or yellowish, with or without prickly superficial roots, giving rise to one or few annually renewed tubercules; stem up to 12 m long, twining to the right, glabrous, spiny or not. Leaves usually alternate in basal part of stem and opposite in upper part, basal leaves often strongly reduced, simple; stipules absent; petiole 5-12 cm long; blade broadly ovate to suborbicular, 5.5-12 cm x 5-10 cm, broadly cordate at base, acuminate at apex, entire, 5-7-veined. Inflorescence an axillary unisexual spike, male 1-3 together and 4-6 cm long, female 1-2 together and 10-12 cm long. Flowers unisexual, regular, with 6-lobed perianth; male flowers small (1-2 mm in diameter), sessile or shortly pedicelled, with six stamens; female flowers with inferior, 3-celled ovary, styles 3, short. Fruit a capsule wider than long, 2-2.5 cm x 3-3.5 cm, opening by three valves, up to 6-seeded. Seeds 1-1.5 cm x 1-1.5 cm, with large circular wing.” (PROTA 2017)

a simple drawing of *cayenensis*.

A regular specimen of *Dioscorea cayenensis* requires 25-30 degrees Celcius for optimal growth. Beneath 20 degrees, vine growth is impaired, and above 35 degrees, sprouting is limited (PROTA 2017). Due to the short dry-season, *cayenensis* prefers the tropics of Africa instead of the subtropic savanna-climate. 600mm of precipitation a year is adequate for small individuals; 1500mm is the optimum for the average specimen (PROTA 2017). Drought-resistance is at its peak performance in the earliest phases of *cayenensis’s*  growth period of 11 months. It benefits greatly from fertile soil, and a pH-value between 5.5-6.5. It needs a sufficient amount of light for a proper tuber/vine ratio that results in higher yield. The highest altitude, at which an individual *cayenensis* was spotted, was 1100m (PROTA 2017).

Since climate change has gained an increasing impact on patterns of temperature and precipitation over the world in the last decades, chances are substantial that it will impact the future distribution of agricultural crops because of this as well. Since *cayenensis* does not grow in the wild, it can’t adapt on its own by moving to a more suitable landscape in terms of climatic factors. How exactly the climate change will impact specific abiotic parameters, that determine the spread of *cayenensis*, is what we try to uncover with this research.

With the aid of species occurrence data belonging to *Dioscorea cayenensis* from GBIF, and climate factor information from Worldclim, we seek to overlap these two datasets and reveal what climate variables mostly determine the distribution of *cayenensis* in West-Africa. Afterwards, we compare this with a future data-set from 2050, to see how the pattern of climate variables changes, and speculate how this might impact the distribution of the crop in regards to how much habitable space will remain in the future scenario. We use the programs MaxEnt and R to accomplish this portrayal of Species Distribution Models.

Research Question: To what degree will climate change impact the spatial distribution of *Dioscorea cayenesis* in Africa, looking at patterns of distinct climate variables?

*Hypothesis 1*: The mean temperature of the coldest month is significantly different in 2050 compared to 2019, in areas occupied by the Guinea yellow Yam; if so, the optimal climate conditions for the crop will shift northward.

*Hypothesis* 2: De mean temperature of the coldest quarter is significantly different in 2050 compared to 2019, in areas occupied by the Guinea yellow Yam; if so, the optimal climate conditions for the crop will shift northward.

*Hypothesis* 3: The annual precipitation is significantly different in 2050 compared to 2019, in areas occupied by the Guinea yellow Yam; if so, these areas will show a stronger degree of drought.

*Hypothesis* 4: De precipitation in the driest month is significantly different in 2050 compared to 2019, in the areas occupied by the Guinea yellow Yam; if so, these areas will show a stronger degree of drought.

*Hypothesis* 5: De precipitation in the coldest quarter is significantly different in 2050 compared to 2019, in the areas occupied by the Guinea yellow Yam; if so, these areas will show a stronger degree of drought.

Species Distribution Models can aid farmers with taking the right course of action for future cultivation, before they invest precious time and resources into crops and locations that are unlikely to produce sustainable yield, and continue to do so in coming years. Depending on the output of the SDM, farmers might have to move their crops much further away from their original site in the decades to come, or stop cultivating this plant altogether and switch to crops that are better suited to grow in these new climatic conditions.

METHODOLOGY

We gather occurrence data of *Dioscorea cayenensis* on gbif.org, selecting occurrence data from African countries only(because of the large-scale cultivation of the crop there) and download it-> We upload the occurrence data into excel(as csv-file) and modify the table for just longitude and latitude as variables to remain-> We download climate variables from worldclim.org; version 1.4; The manual provided by Rutger Vos shows this pathway for selection of the variables: Future, 5min, HadGEM2-ES, rcp85, bi ->the selected environmental variables are loaded into R and renamed to easily distinguish them -> We crop the Area of Interest in both the Present and Future scenario layers according to the manual by Rutger Vos, creating an Area of Interest->We plot the species occurrence data over all the scenario layers->We save all the resulting outputs-> We create an autocorrelation table in Excel with the environmental variables through R->We remove all the irrelevant variables in this table, based on the amount of autocorrelation and their initial significance->We execute multiple VIF-tests to confirm the amount of correlation amounts in the table, each time a variable is removed->We upload the remaining variables into Maxent(these being 6, 11, 12, 14, 19 for this particular experiment), along with the occurrence data file->We run the maximum entropy algorithm in MaxEnt 3 times->We check if the AUC of the ROC for each run is above 0.5 to ensure that the model is reliable->We calculate the average Maximum training sensitivity plus specificity threshold from the multiple runs->We use this value to convert habitat suitability maps to thresholded binary maps in R(using the script provided by Rutger Vos)-> We calculate differences in habitat suitability at global scale and occurrence scale in R(using the script provided by Rutger vos) .

The occurrence data is a csv file, consisting of three separate columns: Name of the species, Longitude coordinates, Latitude coordinates. The data reflects the ecological niche of *cayenensis,* in the way that the species prefers tropical ecosystems with sufficient moisture of >600 mm p/y and a temperature-interval between 25-30 degrees; all the datapoints are located in the tropical zone, that is known to possess these conditions. All other, non-relevant columns were removed.

In regards to the environmental data, the remaining selected variables are: Minimal temperature of the coldest month (6), Mean temperature of the coldest quarter (11), annual precipitation (12), Precipitation of the Driest Month (14), Precipitation of the coldest quarter (19). Temperature and precipitation are measured in degrees Celcius and mm/year, respectively.

The MaxEnt settings were:

- Create Response Curves,

- Make pictures of predictions,

- Do Jackknife to measure variable importance,

*Basic:*

-Random Seed,

- Give visual warnings

- Show tooltips,

-Ask before overwriting,

-remove duplicate presence records,

-write clamp grid when projecting,

-do MESS analysis when projecting, random test percentage = 0,

-regularization multiplier = 1,

-max number of background points = 10000,

-replicates = 3,

-replicated run type = crossvalidate,

- add samples

*Advanced:*

-Add all samples to background,

-extrapolate, do clamping,

-write output grids,

-write plots,

- cache ascii files,

-Maximum iterations = 500,

-convergence threshold = 0.00001,

-adjust sample radius = 0,

-log file = maxent.log,

-default prevalence = 0.5,

*Experimental:*

-Logscale raw/cumulative pictures,

-Threads = 1,

-Lq to lqp threshold = 80,

-Linear to lq threshold = 10,

-Hinge threshold = 15,

-Beta threshold =-1,

-Beta categorical = -1,

-Beta lqp =-1,

-Beta hinge = -1,

-Default nodata value = -9999

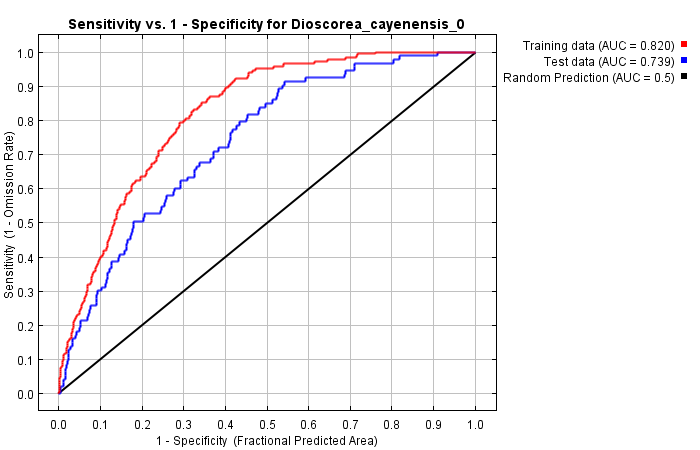
**RESULTS:**

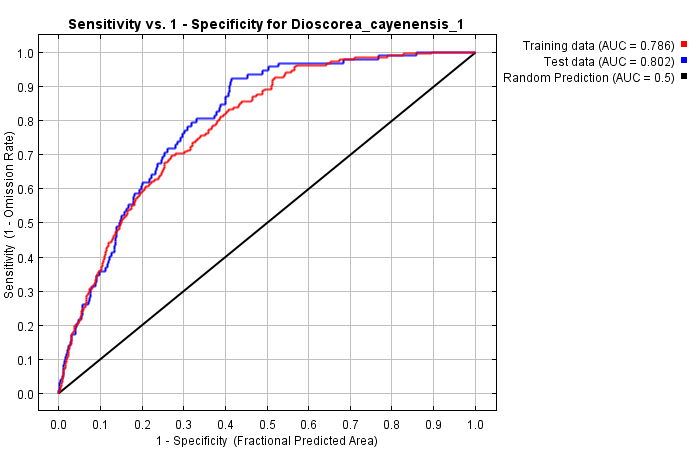
*Autocorrelation results:*

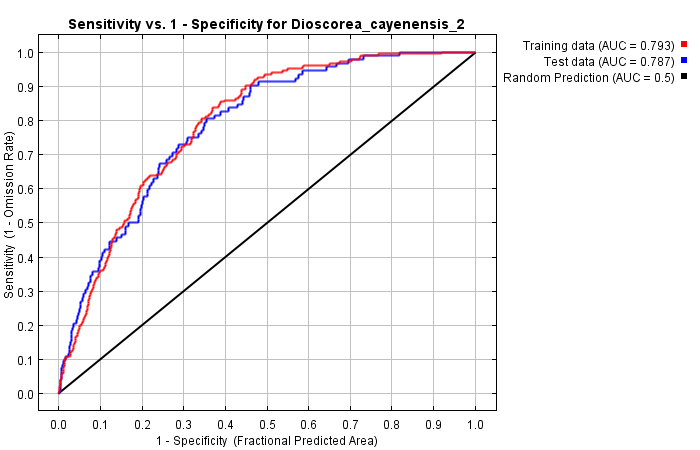
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Bio01 | Bio06 | Bio09 | Bio11 | Bio12 | Bio14 | Bio17 | Bio19 |
| Bio01 | 1 | 0.262487 | 0.798706 | 0.907111 | -0.44258 | -0.31814 | -0.35854 | -0.14494 |
| Bio06 | 0.262487 | 1 | 0.702486 | 0.522636 | 0.473924 | 0.435878 | 0.493938 | 0.497539 |
| Bio09 | 0.798706 | 0.702486 | 1 | 0.888976 | -0.06233 | -0.03795 | -0.05031 | 0.292897 |
| Bio11 | 0.907111 | 0.522636 | 0.888976 | 1 | -0.19483 | -0.15091 | -0.17259 | 0.07152 |
| Bio12 | -0.44258 | 0.473924 | -0.06233 | -0.19483 | 1 | 0.696583 | 0.730604 | 0.632585 |
| Bio14 | -0.31814 | 0.435878 | -0.03795 | -0.15091 | 0.696583 | 1 | 0.966958 | 0.267496 |
| Bio17 | -0.35854 | 0.493938 | -0.05031 | -0.17259 | 0.730604 | 0.966958 | 1 | 0.268001 |
| Bio19 | -0.14494 | 0.497539 | 0.292897 | 0.07152 | 0.632585 | 0.267496 | 0.268001 | 1 |
|  |  |  |  |  |  |  |  |  |

Correlation higher than 0.7 or lower than -0.7 is deemed as an unreliable variable, depending on how many times this value occurs.

Multicolinearity test resulted in removal of Bio06, and Bio17, because of VIF values higher than 10(unfortunately, I was unable to present the exact values here due to system errors)

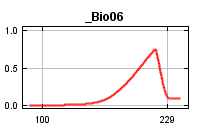
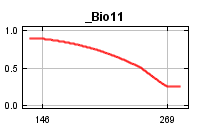
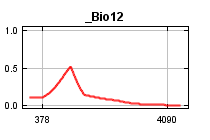
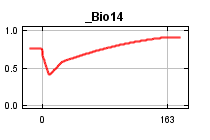
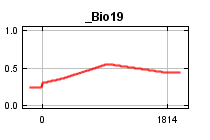


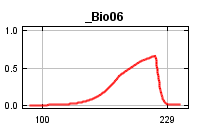
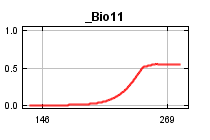
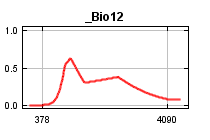
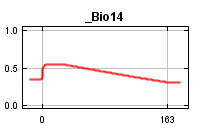
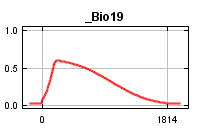


*Graphs of ROC for the three MaxEnt replicates.*

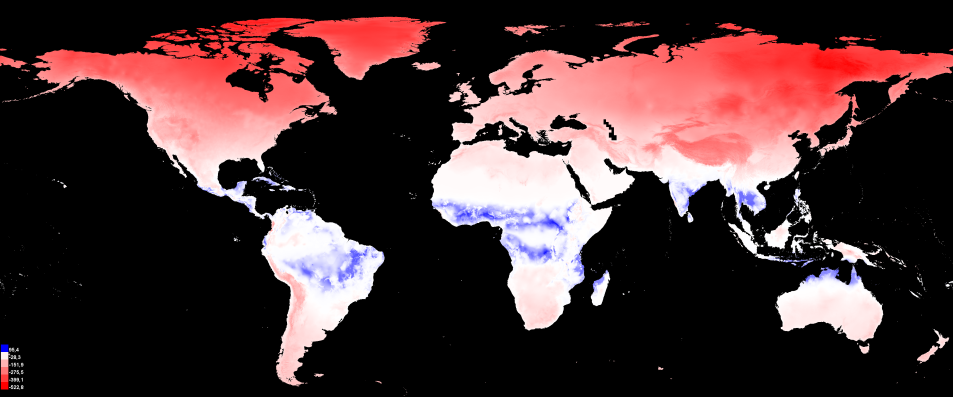
Alle three replicates have an AUC between 0.73-0.82; this is consistently higher than the value of the random prediction line with AUC=0.5, thus the model is deemed reliable.

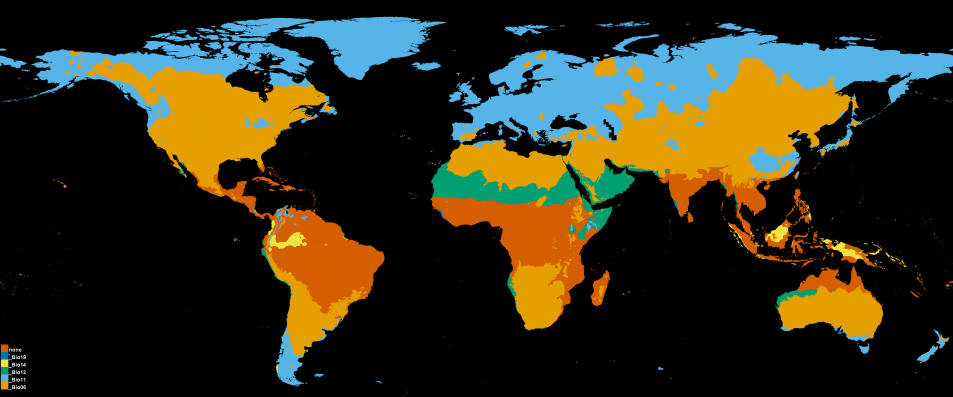
*Below: Graphs with curves of environmental variables against predicted probability of presence of cayenensis*

These curves show how the predicted probability of presence changes as each environmental variable is varied, keeping all other environmental variables at their average sample value.    
  
[](file:///C:\Practical\Maxent\Results\plots\Dioscorea_cayenensis_0__Bio06.png) [](file:///C:\Practical\Maxent\Results\plots\Dioscorea_cayenensis_0__Bio11.png) [](file:///C:\Practical\Maxent\Results\plots\Dioscorea_cayenensis_0__Bio12.png) [](file:///C:\Practical\Maxent\Results\plots\Dioscorea_cayenensis_0__Bio14.png) [](file:///C:\Practical\Maxent\Results\plots\Dioscorea_cayenensis_0__Bio19.png)

In contrast to the above marginal response curves, each of the following curves represents a different model, namely a MaxEnt model created using only the corresponding variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other variables.  Looking at the difference between some curves of the same variable, like 11, 14 and 19, influence from other variables is likely. [](file:///C:\Practical\Maxent\Results\plots\Dioscorea_cayenensis_0__Bio06_only.png) [](file:///C:\Practical\Maxent\Results\plots\Dioscorea_cayenensis_0__Bio11_only.png) [](file:///C:\Practical\Maxent\Results\plots\Dioscorea_cayenensis_0__Bio12_only.png) [](file:///C:\Practical\Maxent\Results\plots\Dioscorea_cayenensis_0__Bio14_only.png) [](file:///C:\Practical\Maxent\Results\plots\Dioscorea_cayenensis_0__Bio19_only.png)

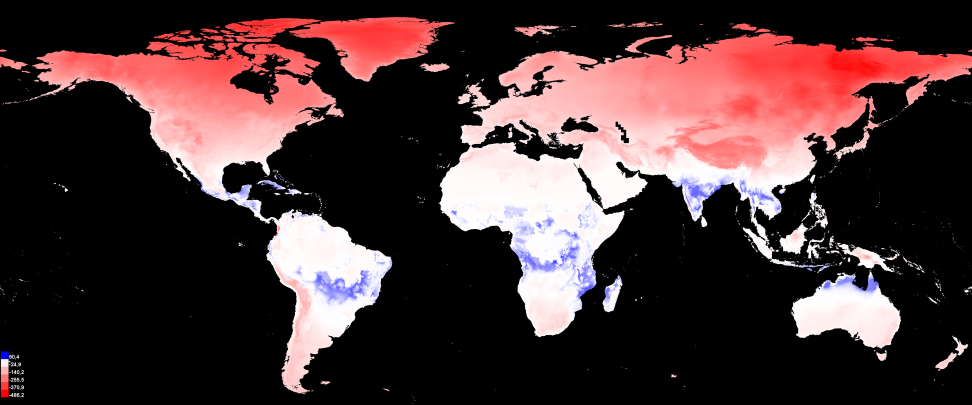
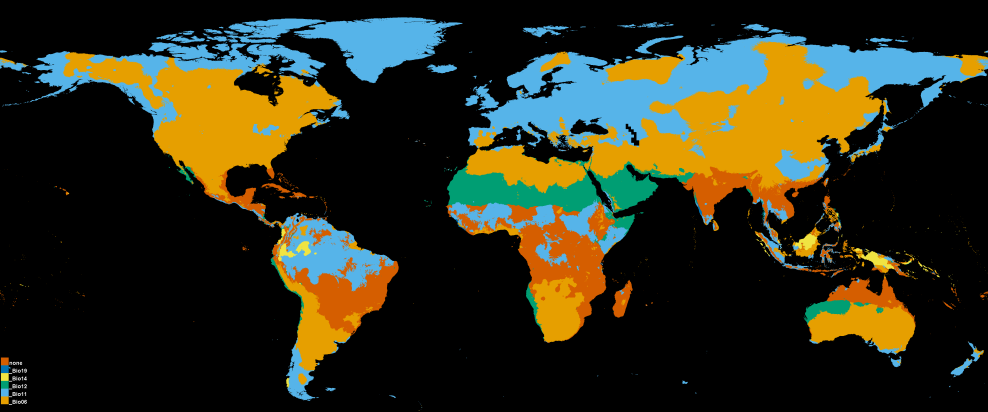
*Maps of the Present World dataset, reflecting dissimilarity of the set perimeter of the training dataset for environmental variables.*





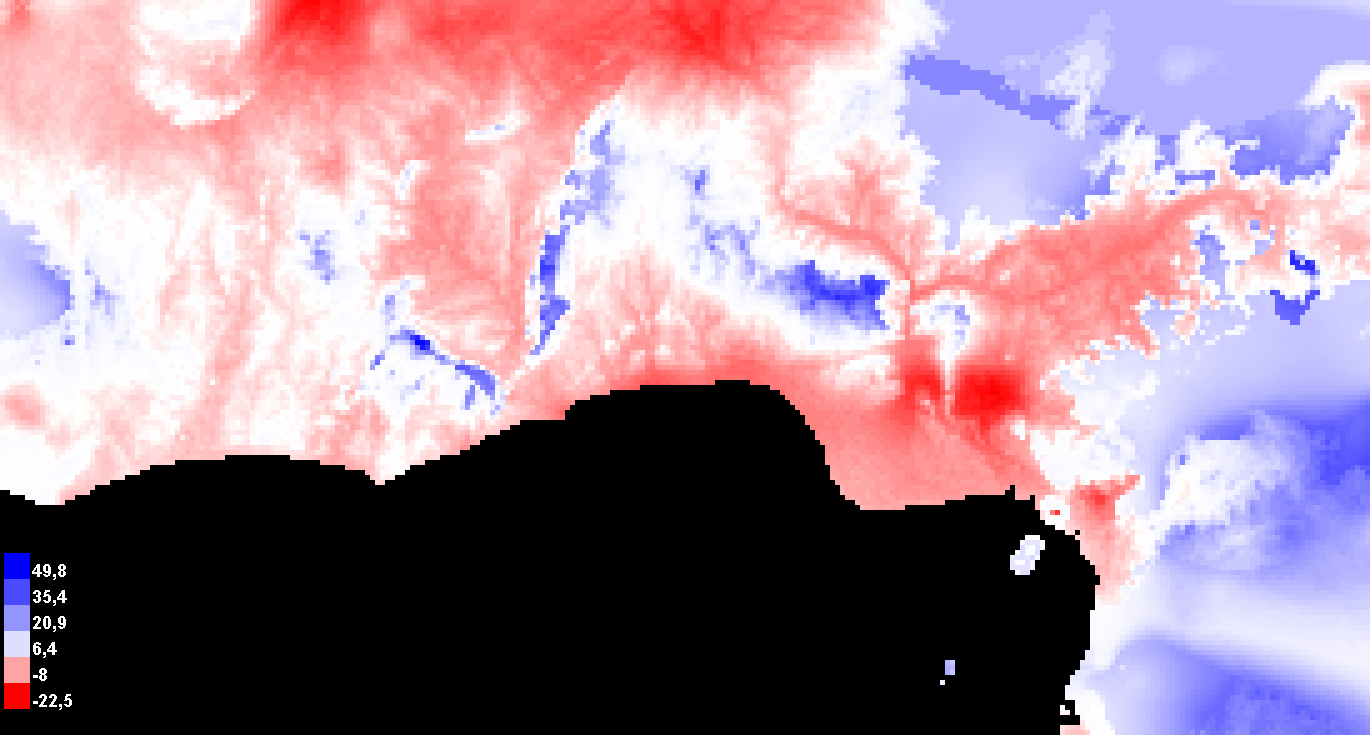
Looking at the Present data output, the entire region in which the species under investigation occurs, has next to no values/variables outside the perimeter of the training dataset (low dissimilarity is indicated by a blue color and brown indicates none of the selected variables). This output is reliable for further interpretation.

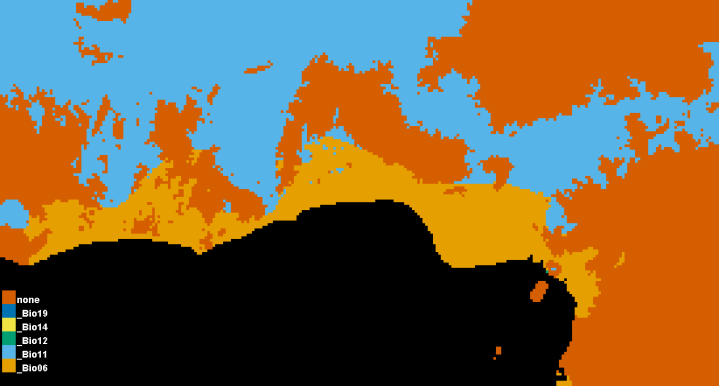
*Below: World Maps of the Future dataset, reflecting dissimilarity of the set perimeter of the training dataset for environmental variables.*

[](file:///C:\Practical\Maxent\Results\Dioscorea_cayenensis_0_he45bi50_WLD_novel.png)  
  
[](file:///C:\Practical\Maxent\Results\Dioscorea_cayenensis_0_he45bi50_WLD_novel_limiting.png)

In this Future data output however, visibly more variables exist in the region, outside of the established perimeter, and strong caution is advised when interpreting this data further.

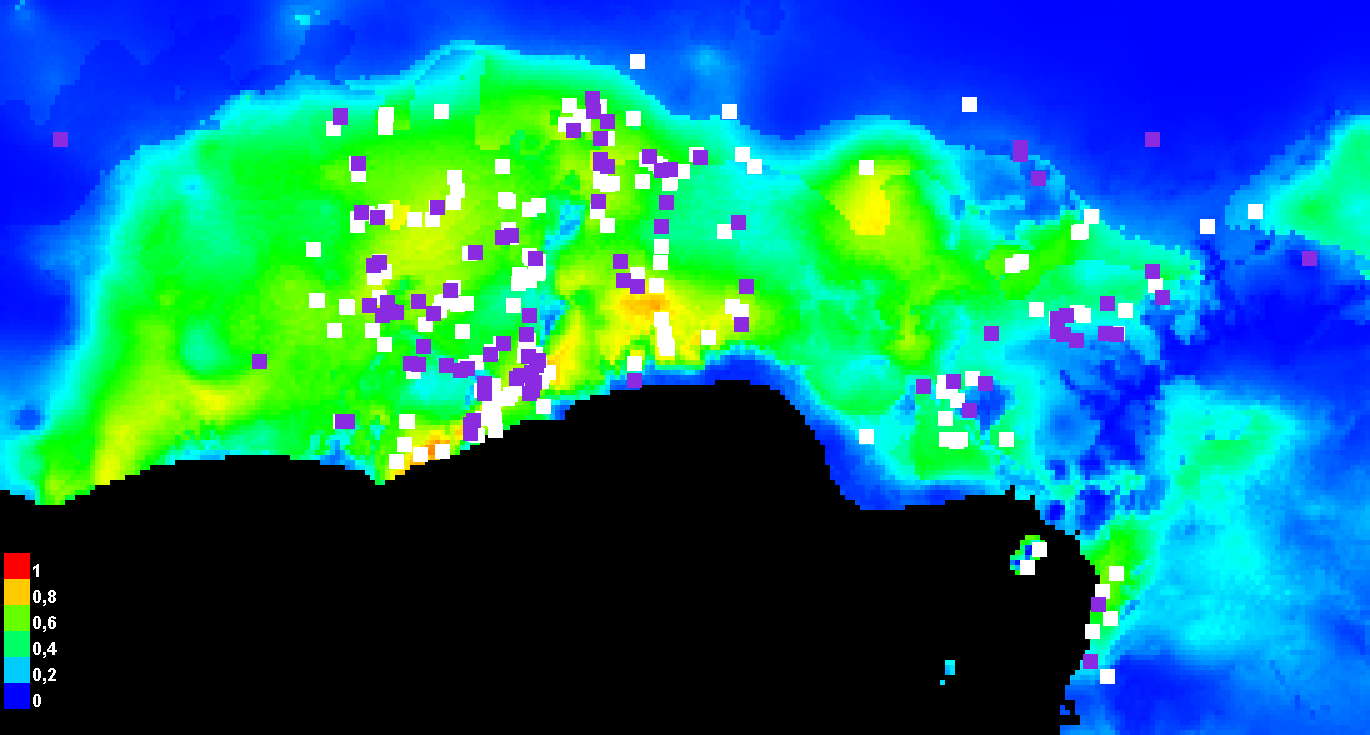
*Below: Cropped AOI maps, reflecting dissimilarity of the set perimeter of the training dataset for environmental variables.*

[](file:///C:\Practical\Maxent\Results\Dioscorea_cayenensis_0_he45bi50_AOI_novel.png)

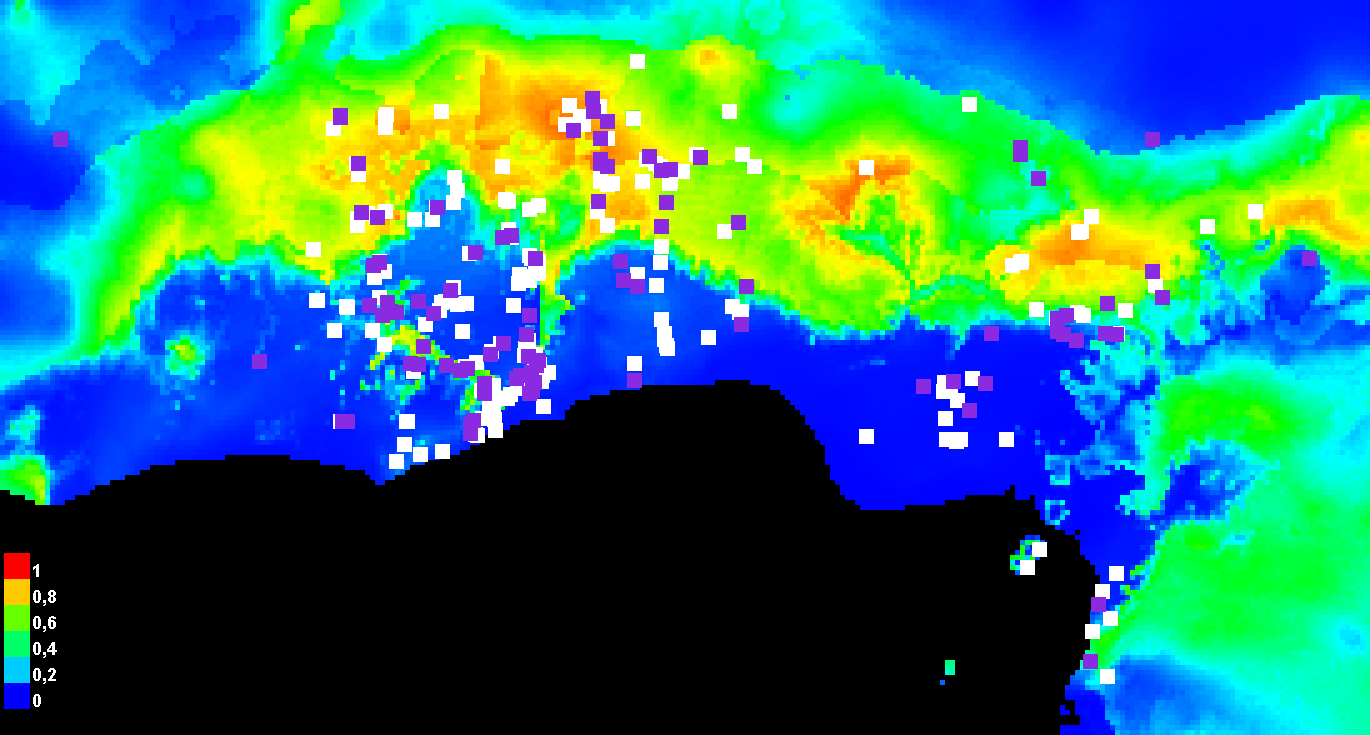
[](file:///C:\Practical\Maxent\Results\Dioscorea_cayenensis_0_he45bi50_AOI_novel_limiting.png)

As we zoom in on the Future data, the dissimilarity (indicated by red) seems to be related entirely to Bio11 en Bio06(being blue and yellow respectively), both variables revolving around temperature. The optimal temperature-intervals are likely to shift northward in the coming years (climate change being the main driver). When predicting the species-distribution combined with a suitable environment, it is thus important to remember that these regions can diverge significantly from the norm in terms of temperature, and as a result are unlikely to be classified as a suitable future habitat for the species.

*Below: Maps of the AOI, overlayed with present species distribution of Dioscorea cayenensis. Purple belongs to the test-data, white to the training-data.*



Present Scenario: most samples occur in a suitable environment, indicated by warmer colors, with the exception of an odd pattern at the east of Nigeria.



Future Scenario: de suitable conditions have shifted northward, and have shrunk in terms of their latitudinal length.

According to the SDM’s in R (not visible due to complications with the script), a decrease in the availability of suitable geographical space has become visible in the Future Scenario when comparing it with the Present Scenario. Due to this species not occurring in the wild, and the fact that close to every population has been cultivated, it is unlikely that they will be reach the suitable environment on their own.

**Discussion and Conclusion:**

It is remarkable that the most dissimilar variables are entirely related to temperature and not to precipitation amounts. Looking at the projection of the world, this is the case (albeit on small-scale) in South-America and South-East Asia. However the species has only recently arrived in some countries, and has not been cultivated yet on a large scale; further investigation in these locations is less relevant. Aside from that, a lot of samples are concentrated around Nigeria, where the conditions for *cayenensis* are deemed far from optimal.

It has become common knowledge now, that increasing temperatures and extreme rainfall or drought will become a consequence of climate change(Weber et al, 2018). Our results somewhat reflect this, in the sense that savanna-climate zones will resemble the tropics more and more in terms of temperature, while the tropics themselves become increasingly hotter. Precipitation patterns, on the other hand, do not seem to change at all. It is known that *cayenensis* is also strongly depending on a fertile and slightly acidic soil, but Nigeria has become known for years over their gross mismanagement of their soil due to poverty and pollution (Huising et al., 2017). Other overlooked properties of the environment might cause this odd pattern, like a difference in cultivation method that provides sufficient moisture and light. Because *cayenensis* has a narrow range of temperature-tolerance, the estimated shifting of temperature-optima will most likely have very negative consequenses for the crop’s cultivation yield as a whole. Farmers are likely to be forced to relocate or switch their cultivation crops, due to the fact that *cayenensis* can’t relocate and settle on its own in the wild.

Two out of five hypotheses can be accepted; There is a visible difference in the patterns of temperature between 2019 and 2050. The minimum temperature of the coldest quarter changes strongly up north, while the minimum temperature of the coldest month changes significantly in the south. The other three hypotheses, revolving around precipitation, can be rejected because their change is not observed at all between the two periods of time.

It must be taken into account that SDM’s can only work with present biotic data to map out the distribution of species. With regards to the future, we can only guess, given what we know about earlier distribution patterns.

Future research could aim at the potential shift in precipitation patterns in other tropical areas, and how this could influence the spread of *cayenensis*. For this goal, samples can be taken from Brazilia, Puerto Rico and Colombia, while considering that the plant is not wide-spread in these countries. In addition, several more crops from Nigeria with a similar temperature tolerance-range could be investigated in the same manner, to discover a solid explanation as to why crops under suboptimal conditions are currently still growing there.

Rutger Vos, 2019. Exercise: Model your chosen species’ habitat suitability under present and future climate conditions, *Manual from the course Methods in Biodiversity Analysis, Leiden University, Netherlands*

Oviedo Prieto R, Herrera Oliver P, Caluff MG, et al., 2012. National list of invasive and potentially invasive plants in the Republic of Cuba – 2011. (Lista nacional de especies de plantas invasoras y potencialmente invasoras en la República de Cuba – 2011). *Bissea: Boletin sobre Conservacion de Plantas del Jardin Botanico Nacional de Cuba*, 6(Special Issue 1):22-96

PROTA, 2017. PROTA4U web database. Grubben GJH, Denton OA, eds. Wageningen, Netherlands: Plant Resources of Tropical Africa. <https://uses.plantnet-project.org/en/Dioscorea_cayenensis_(PROTA)>

T, A. O., Huising, J., Ojo, A. O., Azeez, J., Are, K. ., Olakojo, S. ., … Ojeniyi, S. . (2017). Status of Integrated Soil Fertility Management (ISFM) in Southwestern Nigeria. *International Journal of Sustainable Agricultural Research*, *4*(2), 28–44.

Weber, T., Haensler, A., Rechid, D., Pfeifer, S., Eggert, B., & Jacob, D. (2018). Analyzing Regional Climate Change in Africa in a 1.5, 2, and 3°C Global Warming World. *Earth’s Future*, *6*(4), 643–655. <https://doi.org/10.1002/2017EF000714>