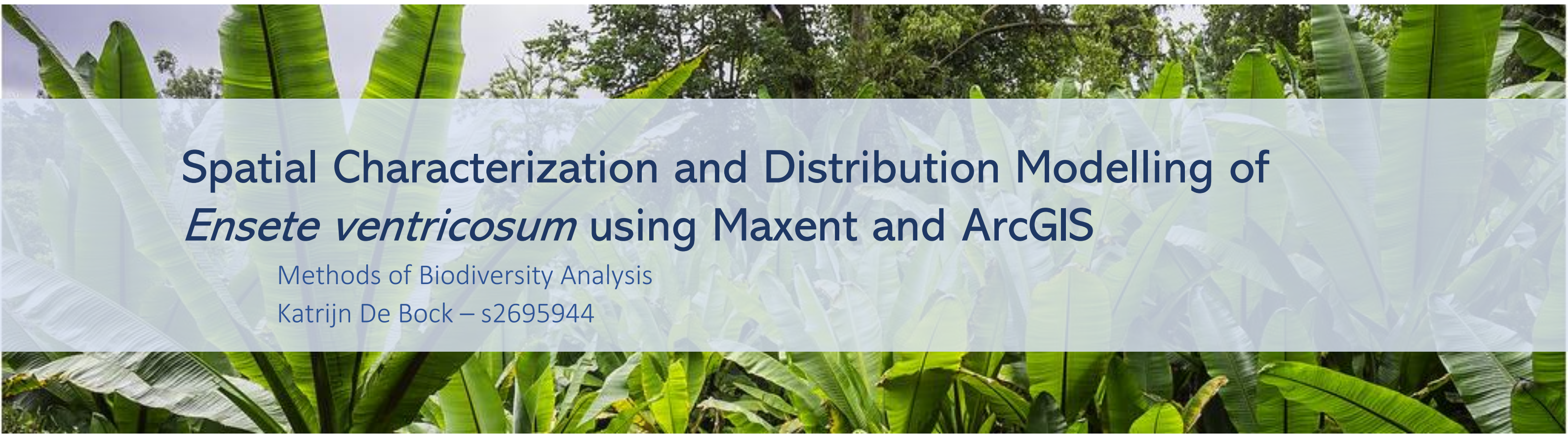

Spatial Characterization and Distribution Modelling of *Ensete ventricosum* using Maxent and ArcGIS

Methods of Biodiversity Analysis
Katrijn De Bock – s2695944

Introduction

Ensete ventricosum (Welw.) Cheesman, also known as enset, false banana, Ethiopian banana or Abyssinan banana, is a monocotyledon of the genus *Ensete* which belongs to the Musaceae (Yemataw *et al.* , 2018). The genus *Ensete* is closely related to the genus *Musa* which has the common banana (*M. balbisiana* and *M. acuminata*) as an important crop. However, the plant is not cultivated for their fruits since they contain very hard and large seeds (Karlsson, Tamado, Dalbato, & Mikias, 2014). Nevertheless, it is a multipurpose crop that ensures food ('Kocho' and 'Bulla' as pseudo-stem products and 'Amicho' as a root product), construction materials, medicines, fiber, cultural practices and animal forage. Especially in Ethiopia, which is a food insecure country, the productivity of enset is very high compared to other root and tuber crops because it has the advantage of year round food production (Yemataw *et al.*, 2017). As a result of this, more than 20 percent of the population is dependent on the enset-based farming system (Brandt *et al.*, 1997).

Wild *E. ventricosum* grows best at elevations between 2,000 and 2,750 meters (open mountain forests) with a mean annual rainfall of about 1,000 to 1,500 millimeters of which the majority falls between March and September. Cultivated enset grows at altitudes ranging from 1,200 to 3,100m (Eshetae *et al.*, 2019). The preferred daytime temperature is within a range of 16 to 24°C, but it can tolerate 8 to 30°C (Fern *et al.*, 2019**)**. The higher *E. ventricosum* grows/is being cultivated the higher the chance for frost damage since they can't tolerate temperatures equal to or lower than -2°C. When cultivating the crop below 1,500 meters, the constraint for plant growth is probably more related to the availability of water than to the high temperature because of the increased amount of evaporation that is taking place compared to higher elevations (Brandt *et al.*, 1997). However, it is been found that it is a drought tolerant species (Masayoshi, 1989).

Several factors, both biotic (e.g. fungi and vertebrate pests) and abiotic (e.g. soil type, soil fertility, rainfall and altitude), threaten the growth and distribution of both the crop and wild varieties (Eshetae, *et al.*, 2019). Additionally, climate change will make droughts and other environmental events even more unpredictable (Mohammed *et al.*, 2013). Therefore, measures should be taken to determine the future impact of these threats on the enset populations (Eshetae *et al.*, 2019). Nevertheless, little information has been documented on the characteristics of enset in Ethiopia in terms of environmental indications and species distribution. This study aims to characterize enset based on data of occurrences from the species downloaded from GBIF and use this to identify potentially suitable areas for wild and cultivated enset with Maxent modelling.

Being a drought tolerant species and also having a wide range of elevation possibilities where it can grow. It is hypothesized that *E. ventricosum* might respond to climate change in different ways. First of all, since elevation is an important factor for plant growth, their distribution might be restricted to specific mountain areas (such as for example the Ethiopian Highlands) because dispersal towards other mountains will be limited. According to this hypothesis, the distribution of the crop will thus decrease since thing 1. might need to move up onto the mountain if they want to stay in their suitable conditions and thus a decrease in the habitat suitability might take place. Elaborating on this, it might be possible as well that other suitable mountain areas in other parts of the world will exist where the crop is not present yet. Secondly, it has been shown they can grow below 1,500m since they are drought-tolerant species. Therefore we can hypothesize that the distribution of *E. ventricosum* might expand to other ranges where climatic conditions won't be more suitable but where it can survive due to its tolerance.

Methodology

Flowchart

In the beneath figure (Figure 1), a general overview of the main steps being done in this study is shown.

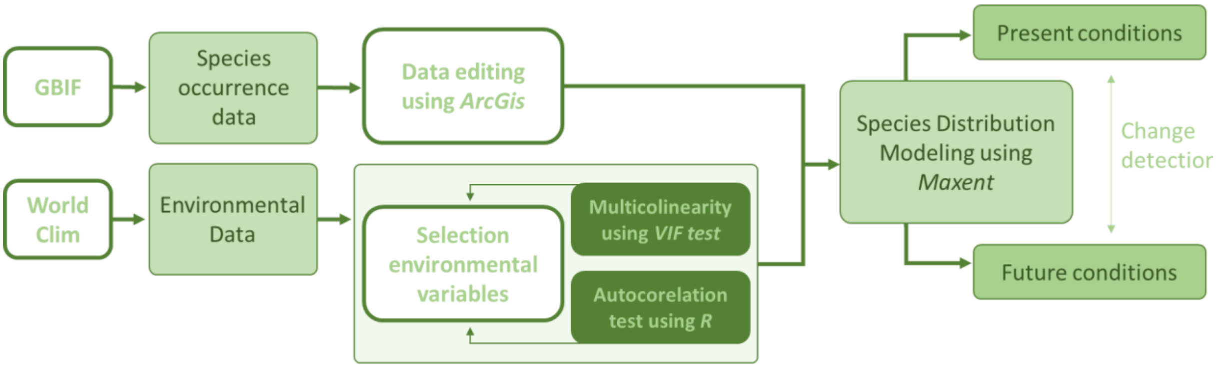

Figure 1 Flow chart of the different steps that are done in this research study.

Species occurrence data

Maxent modelling requires species occurrence records and environmental variables as input data. Using GBIF (Global Biodiversity Information Facility – https://www.gbif.org/), the coordinates (longitude and latitude) of 167 species of *E. ventricosum* were selected and used as species occurrence data points. By converting the data in ArcGIS using ArcMap we could display the collection localities to remove obvious errors. This implemented strange locations in the sea, different sampling strategies between different areas,… For our dataset 20 datapoints have been removed resulting in a complete dataset of 147 well-defined species. In Figure 2 we see the selected *E. ventricosum* species have a pantropic distribution. According to the ecological niche characteristics mentioned earlier, the distribution thus properly reflects our description.

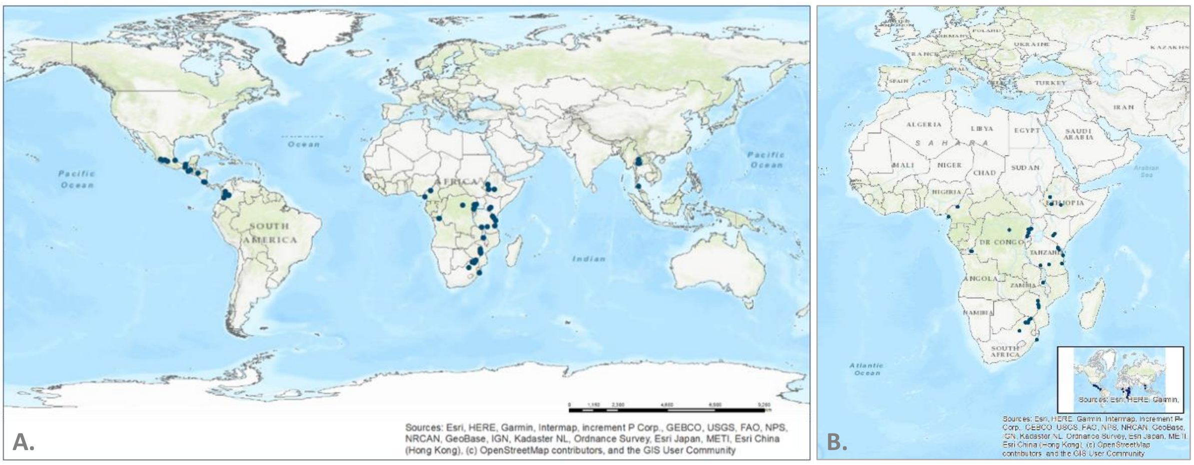

Figure 2 A. Map of the locations of the selected species of *Ensete ventricosum* for this study across the world; B. A more detailed map of the selected species in Africa where it clearly is been seen that Ethiopia is one of the countries where the species occurs.

Environmental data

Using WorldClim (Global Climate Data - http://www.worldclim.org/), 6 from the 19 bioclimatic variables were selected to use in our model. These bioclimate variables, which are derived from the monthly temperature and rainfall values, implement bio 4, bio 6, bio 8, bio 13, bio 15 and bio 16 (Table 1). Those data have been chosen since they are assumed to be most ecologically significant for the species. Precipitation in the wettest month together with cold temperatures are to conditions essential for growth of the species. Both present conditions and future conditions from the rcp4.5 scenario were selected for extrapolations of the distribution of our selected species in the future. To reduce model over parameterization, highly correlated environmental variables (r < 0.7 ; see Appendix 1 and VIF > 10,0) were excluded from the analysis. This resulted in bio 13 being removed from the analysis (Table 1).

Table 1 List of the bioclimatic variables used in this study and their Variance inflation factor.

Name	meaning	VIF (with bio 13)	VIF (without bio 13)
Bio 4	Temperate Seasonality	4.686347	4.612433
Bio 6	Min Temperature of Coldest Month	5.801555	5.697032
Bio 8	Mean Temperature of Wettest Quarter	3.068442	3.025638
Bio 13	Precipitation of Wettest month	50.152458	-
Bio 15	Precipitation Seasonality	1.439561	1.271662
Bio 16	Precipitation of Wettest Quarter	52.317427	1.936726

Model settings

Some settings were made to make run the model properly. Response curves where activated to see the response of different variables on each other and separately. Jackknife was activated to measure the variable importance. The output format was logistic. Auto features where also activated. The model has been trained 5 times (5 replicates). 63 presence records used for training, 16 for testing. 10061 points used to determine the Maxent distribution (background points and presence points).

Results

Model performance evaluation

To validate the performance of the Maxent model, the AUC (Area Under the Curve) of the ROC (Receiver Operating Curve) was validated in which the true positive rates (TPR), also called sensitivity, was plotted against false positive rates, or 1- specificity. The AUC values range from 0 to 1. The closer the AUC encounters 1, the better the model fits and thus the better the model the distribution of *Ensete ventricosum* predicts (Swets, 1998; Eshetae *et al.*, 2019). The average AUC test for each replication (five replications are accomplished) was 0.930, with a standard deviation (SD) of 0.027 (Figure 4). The model thus performs well.

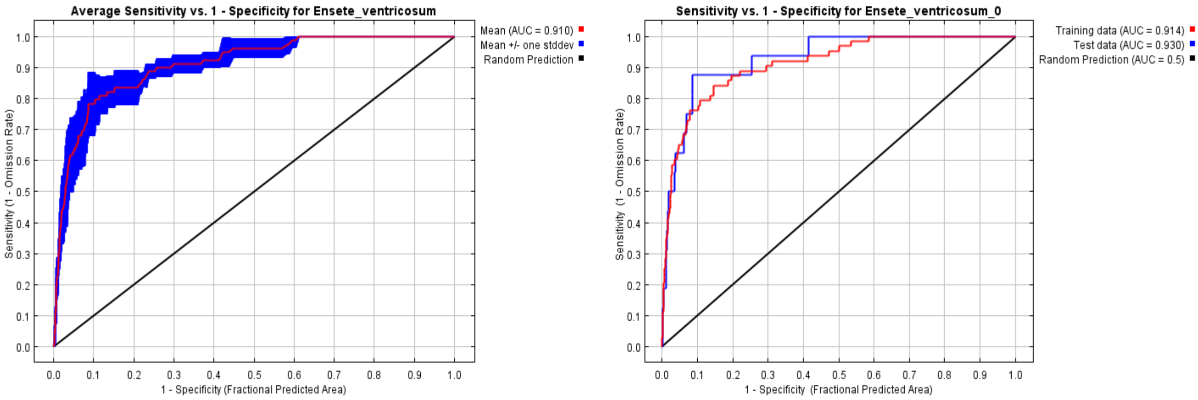

Figure 3 ROC curve and the AUC for the enset species of the model.

The response curves of the model output (Figure 4) show how each environmental variable affects the Maxent prediction. The Y axis shows the probability of the habitat suitability where the x axis explains the variance of temperature, in the case of Bio01 till Bio11, or precipitation for Bio12 till the end. The curves in Figure 4A. - 4E. show how the predicted probability of presence changes as each environmental variable is varied, keeping all other environmental variables at their average value. We see that Bioclimatic variable 04 (Temperature Seasonality) shows a higher probability of habitat suitability when the value is low. This means that low variation in temperature seasonality benefits the habitat suitability for the crop according to the model. Figure 4B. shows that a mediate temperature benefits habitat suitability. This is the same case for the mean temperature of the Wettest Quarter (Figure 4C.). Furthermore, a low variation in precipitation seasonality (Figure 4D.) is also beneficial and a minimum amount of precipitation, according to the model, is required for the habitat to be suitable for the species (Figure 4E.). These results correspond to the ecological description explained in the introduction.

The curves in Figure 4D. – 4E. reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other variables. As we see majority of these curves look quite similar to the curves from Figure 4A. - 4E.. In Figure 4J. we see a decrease after a peak while Figure 4E. shows a stagnation. This means that too much precipitation only decreases the probability of habitat suitability when this variable increases according to the model outputs. In Ecological context this seems reasonable. When considering other variables as well (as shown in curve 4E.), for example temperature, the combination of these variables together still have a positive effect on the habitat suitability.

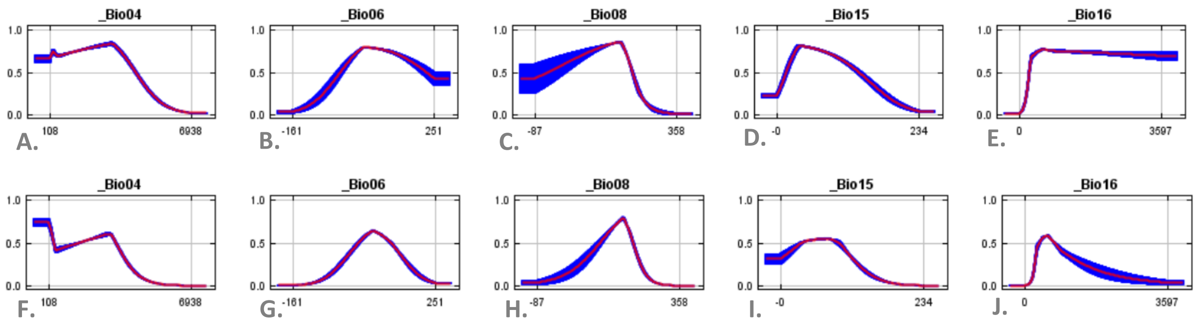

Figure 4 Response variables of the 5 bioclimatic variables being used in the model.

Output Species Distribution Modelling: probabilistic distribution

Figure 5 describes the probabilistic distribution of *E. ventricosum* over the world produced by Maxent for both the present situation and under the rcp4.5 scenario for the future based on climatic variables.

When comparing the outcome of the present probabilistic distribution (Figure 5A.) with the actual occurrence of the species (Figure 2), we see there are a few additional locations that are suitable as well for the crop to grow. This is especially the case for South America where the species doesn't yet occur. This is because of the fact Maxent doesn't take into account dispersion of species and thus dispersion limitation. However, when comparing the output results of Maxent with a map of world mountains, developed by Sayre *et al.* (2018) (Figure 6), we see that the predicted distribution particularly coincides with mountain areas which are suitable areas for these species as mentioned in the introduction.

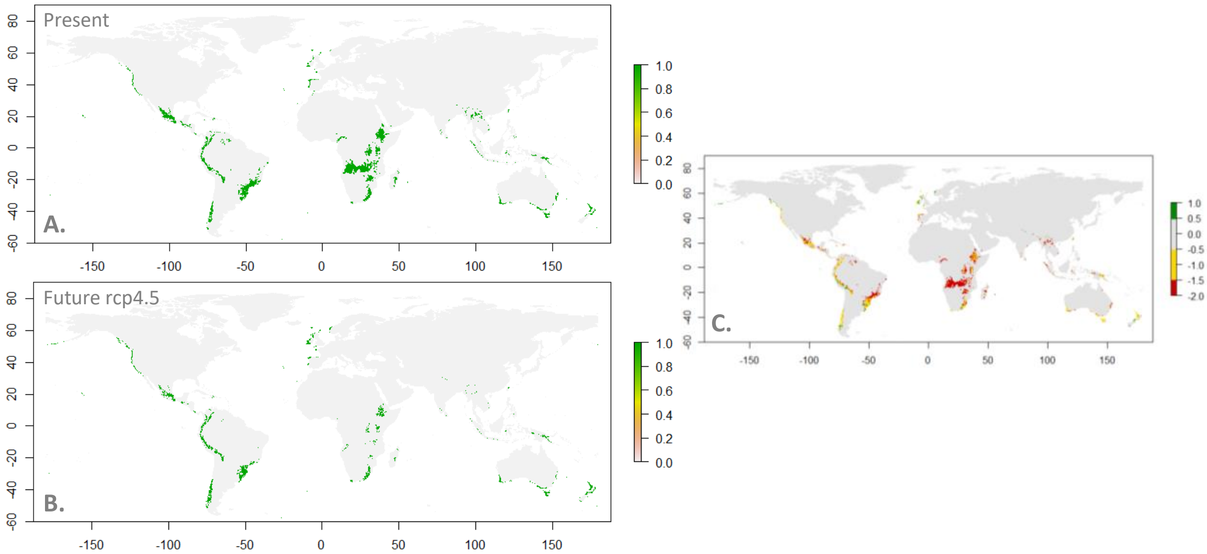

Figure 5 Probabilistic distribution of *E. ventricosum*. A. Probabilistic distribution of *E. ventricosum* for present conditions; B. Probabilistic distribution of *E. ventricosum* future conditions, the rcp4.5 scenario; C. The change in suitable climate conditions between both scenarios. In A. and B.: Green: presence of suitable areas. For C.: Yellow: remaining suitable areas, red: lost suitable areas, green: gained suitable areas (grey is never suitable).

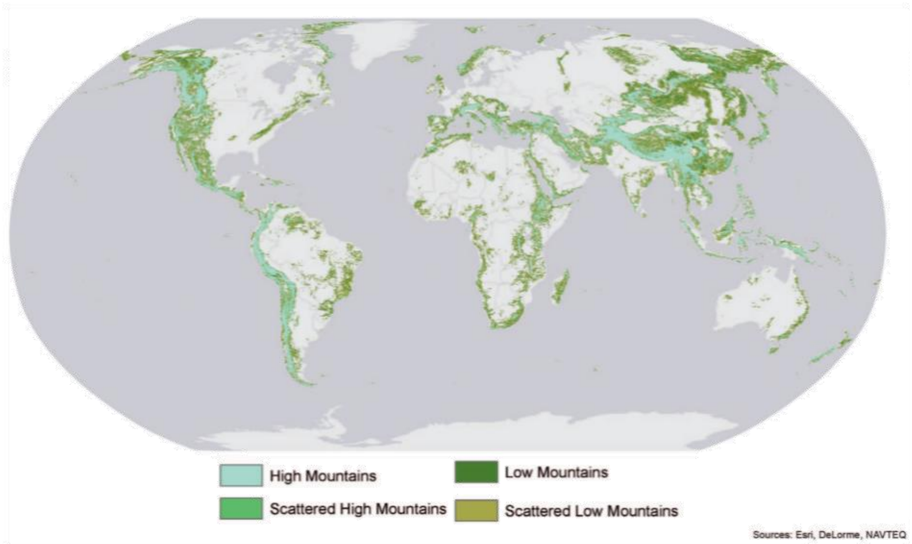

Figure 6 A High-Resolution Map of World Mountains, developed by Sayre *et al.* (2018), with 4 different classes of mountain areas: high, scattered high, low and scattered low.

When focusing on the predicted rcp4.5 scenario (Figure 5B.) and comparing this with the present distribution, we see a modest decrease in suitable habitats (Figure 5C.). This suggests that, under scenario 4.5 for 2050, climate change has an effect on the distribution of *E. ventricosum*. The decrease is especially striking in Africa, where the species is most common. Most mountain areas stay suitable, which is again also the case for the Andes in South America. Furthermore, new suitable conditions are gained in these South American (and also Mexican) mountain areas. Likewise, we have a rise of new suitable habitat areas in Europe, more specifically the United Kingdom and also a part of Norway.

Discussion

From the model output, we clearly see that *Ensete ventricosum* is a species that needs montane climatic conditions. First of all, the model also implemented the mountain areas from South America (Andes) as being suitable for the species to grow. However, since the fruits of *Ensete* have very hard seeds which are at maturity between 1 and 2 centimeters in size, the distribution capacity of the species is limited as being hypothesized (Bekele and Shigeta, 2011).

As explained in the results, there is a decrease in habitat suitability when looking at the rcp 4.5 scenario compared to the present suitable habitats. According to a study of Nwofor and Chineke (2015) the present precipitation and temperature patterns are expected to change. The precipitation patterns will show more periodic swings between wet and dry conditions which will result in aridification and rapid development of deserts. Despite being drought tolerant species, these conditions will thus probably be too severe for the species to survive in areas with an altitude lower than 1500 meters. A movement upwards the mountain will thus be most likely for the species to survive next to adapting to the new environmental conditions they will encounter.

However, the dataset used for the Maxent modelling was quite poor which might have caused some side effects in the output of the model. Furthermore, a better understanding of how these species react to climate change is essential in how to conserve the species. Taking this information into account when interpreting the output of species distribution modelling will make a clearer final report. Additionally to this, Maxent also doesn’t consider several factors such as stochastic events, human land use changes, biotic interactions,… It is important to take these factors into account since they also highly influence the actual niche of the species (Carneiro *et al.*, 2016).

Conclusion

In this study, we investigated the spatial characterization and species distribution of *Ensete ventricosum.* We can conclude that *E. ventricosum* will be confronted with less suitable conditions when the warming of the earth will become more apparent which will result in the decline of the populations if they won’t adapt fast enough to these novel conditions. However, using species distribution modelling will be a helpful tool to predict their potential ranges in response to climate change and may be helpful for making decisions to conserve and manage the species before it is too late.

References

Brandt, S. a., Spring, A., Hiebsch, C., McCabe, J. T., Tabogie, E., Wolde-Michael, G., … Tesfaye, S. (1997). The "Tree Against Hunger": Enset-based Agricultural System in Ethiopia. *Advancement Of Science*, 56. Retrieved from http://www.aaas.org/international/africa/enset/

Eshetae, M. A., Hailu, B. T., & Demissew, S. (2019). Spatial characterization and distribution modelling of Ensete ventricosum (wild and cultivated) in Ethiopia. *Geocarto International*, *0*(0), 000. https://doi.org/10.1080/10106049.2019.1588392

Karlsson, L. M., Tamado, T., Dalbato, A. L., & Mikias, Y. (2014). Early growth and development of Ensete ventricosum (Musaceae) seedlings. *Journal of Plant Sciences*, *1*(1), 11–17. https://doi.org/10.11648/j.jps.20130101.13

Masayoshi, S. (1989). Folk In-Situ Conservation of Ensete [Ensete ventricosum (Welw.) E.E. Cheesman]: Toward the Interpretation of Indigenous Agricultural Science of the Ari, Southwestern Ethiopia. *African Study Monographs*, *10*(3), 93–107. Retrieved from http://ci.nii.ac.jp/naid/110000066028/en/

Mohammed, B., Martin, G., & Laila, M. K. (2013). Nutritive values of the drought tolerant food and fodder crop enset. *African Journal of Agricultural Research*, *8*(20), 2326–2333. https://doi.org/10.5897/ajar12.1296

Yemataw, Z., Bekele, A., Blomme, G., Muzemil, S., Tesfaye, K., & Jacobsen, K. (2018). A review of enset [Ensete ventricosum (Welw.) Cheesman] diversity and its use in Ethiopia. *Fruits*, *73*(6), 301–309. https://doi.org/10.17660/th2018/73.6.1

Yemataw, Zerihun, Chala, A., Ambachew, D., Studholme, D. J., Grant, M. R., & Tesfaye, K. (2017). Morphological variation and inter-relationships of quantitative traits in enset (Ensete ventricosum (welw.) cheesman) germplasm from south and south-western Ethiopia. *Plants*, *6*(4), 1–16. https://doi.org/10.3390/plants6040056

Appendices

	Bio01	Bio02	Bio03	Bio04	Bio05	Bio06	Bio07	Bio08	Bio09	Bio10	Bio11	Bio12	Bio13	Bio14	Bio15	Bio16	Bio17	Bio18	Bio19
Bio01	1	-0,243101	0,2180807	-0,289364	0,7404613	0,8133799	-0,257176	0,8395536	0,8618138	0,8748398	0,9046594	0,2116714	0,2813030	0,0479540	0,2843316	0,2663236	0,0582927	0,0362487	0,18623898
Bio02	-0,243101	1	-0,523755	0,6418282	0,3440780	-0,690813	0,8724866	0,0043189	-0,445592	0,0612983	-0,478199	-0,698266	-0,585068	-0,532823	0,4801901	-0,600982	-0,561251	-0,552876	-0,5478307
Bio03	0,2180807	-0,523755	1	-0,890077	-0,382291	0,6347896	-0,846176	-0,033287	0,4109662	-0,226845	0,5605517	0,6767188	0,5228890	0,5302997	-0,304936	0,5451462	0,5570945	0,5028253	0,57327083
Bio04	-0,289364	0,6418282	-0,890077	1	0,3695685	$\boxed{-0,720463}$	0,9171426	0,0446128	-0,497506	0,2039253	-0,666848	-0,704472	-0,666707	-0,375761	0,1927966	-0,680512	-0,406167	-0,507687	-0,4965018
Bio05	0,7404613	0,3440780	-0,382291	0,3695685	1	0,2511305	0,4391515	0,7736581	0,4882895	0,9474099	0,4165939	-0,319137	-0,192985	-0,296962	0,4724229	-0,219105	-0,308569	-0,410744	-0,1848461
Bio06	0,8133799	-0,690813	0,6347896	$\boxed{-0,720463}$	0,2511305	1	-0,759337	0,5063759	0,8836787	0,4782209	0,9552297	0,5917256	0,5651684	0,3549054	-0,076371	0,5651027	0,3816469	0,3415312	0,50485537
Bio07	-0,257176	0,8724866	-0,846176	0,9171426	0,4391515	-0,759337	1	0,0500875	-0,491945	0,1930227	-0,606552	-0,763751	-0,654297	-0,529038	0,3884666	-0,671796	-0,561661	-0,593113	-0,5928462
Bio08	0,8395536	0,0043189	-0,033287	0,0446128	0,7736581	0,5063759	0,0500875	1	0,5036241	0,8632849	0,6237727	-0,003988	0,0552944	-0,053152	0,3263407	0,0407045	-0,051649	-0,016645	-0,0292563
Bio09	0,8618138	-0,445592	0,4109662	-0,497506	0,4882895	0,8836787	-0,491945	0,5036241	1	0,6425231	0,8971602	0,3611575	0,3810768	0,1737390	0,0882958	0,3741877	0,1917345	0,0826394	0,37095213
Bio10	0,8748398	0,0612983	-0,226845	0,2039253	0,9474099	0,4782209	0,1930227	0,8632849	0,6425231	1	0,5929593	-0,132018	-0,040963	-0,135243	0,3664580	-0,064256	-0,139540	-0,233509	-0,0421872
Bio11	0,9046594	-0,478199	0,5605517	-0,666848	0,4165939	0,9552297	-0,606552	0,6237727	0,8971602	0,5929593	1	0,4786188	0,5161040	0,2057636	0,1216525	0,5101966	0,2273867	0,2432099	0,37504420
Bio12	0,2116714	-0,698266	0,6767188	-0,704472	-0,319137	0,5917256	-0,763751	-0,003988	0,3611575	-0,132018	0,4786188	1	0,8698478	0,7320972	-0,477985	0,9041648	0,7618555	0,7846251	0,73716094
Bio13	0,2813030	-0,585068	0,5228890	-0,666707	-0,192985	0,5651684	-0,654297	0,0552944	0,3810768	-0,040963	0,5161040	0,8698478	1	0,3956522	-0,160208	$\boxed{0,9884927}$	0,4228552	0,6374911	0,58144343
Bio14	0,0479540	-0,532823	0,5302997	-0,375761	-0,296962	0,3549054	-0,529038	-0,053152	0,1737390	-0,135243	0,2057636	0,7320972	0,3956522	1	-0,597551	0,4357827	0,9939507	0,6494583	0,63563774
Bio15	0,2843316	0,4801901	-0,304936	0,1927966	0,4724229	-0,076371	0,3884666	0,3263407	0,0882958	0,3664580	0,1216525	-0,477985	-0,160208	-0,597551	1	-0,211967	-0,619330	-0,439159	-0,4282765
Bio16	0,2663236	-0,600982	0,5451462	-0,680512	-0,219105	0,5651027	-0,671796	0,0407045	0,3741877	-0,064256	0,5101966	0,9041648	$\boxed{0,9884927}$	0,4357827	-0,211967	1	0,4636643	0,6726888	0,60629139
Bio17	0,0582927	-0,561251	0,5570945	-0,406167	-0,308569	0,3816469	-0,561661	-0,051649	0,1917345	-0,139540	0,2273867	0,7618555	0,4228552	0,9939507	-0,619330	0,4636643	1	0,6663904	0,66108575
Bio18	0,0362487	-0,552876	0,5028253	-0,507687	-0,410744	0,3415312	-0,593113	-0,016645	0,0826394	-0,233509	0,2432099	0,7846251	0,6374911	0,6494583	-0,439159	0,6726888	0,6663904	1	0,36760379
Bio19	0,1862389	-0,547830	0,5732708	-0,496501	-0,184846	0,5048553	-0,592846	-0,029256	0,3709521	-0,042187	0,3750442	0,7371609	0,5814434	0,6356377	-0,428276	0,6062913	0,6610857	0,3676037	1

Appendix 1 Results of Autocorrelation test done from the different bioclimatic variables. The numbers surrounded by a red box are variables having a high pairwise correlation. Therefore we deleted bioclimate variable Bio13 out in order to have VIF values lower than 10 and less then 2 high pairwise correlations.