**Distribution modelling for *Colocasia esculenta.***

**Cynthia Meijs**

**1239953**

**Introduction**

The taro (*Colocasia esculenta*) is a starchy root crop that is mainly distributed in southeast Asia. It is one of the oldest crops occurring from approximately 28.000 years ago. The taro grows best in humid soils and warm temperatures, it is very sensitive to frost.

**Methodology**

For the future model I selected the GISS-E2-R model with rcp60 for 2070. I selected 2070, because it is more likely to show change in distribution. I chose for rcp60 because I also expect it to result in more change. The rcp85 would probably show more change, but that would mean a worst case scenario and I like to think we won’t let climate change go that far.

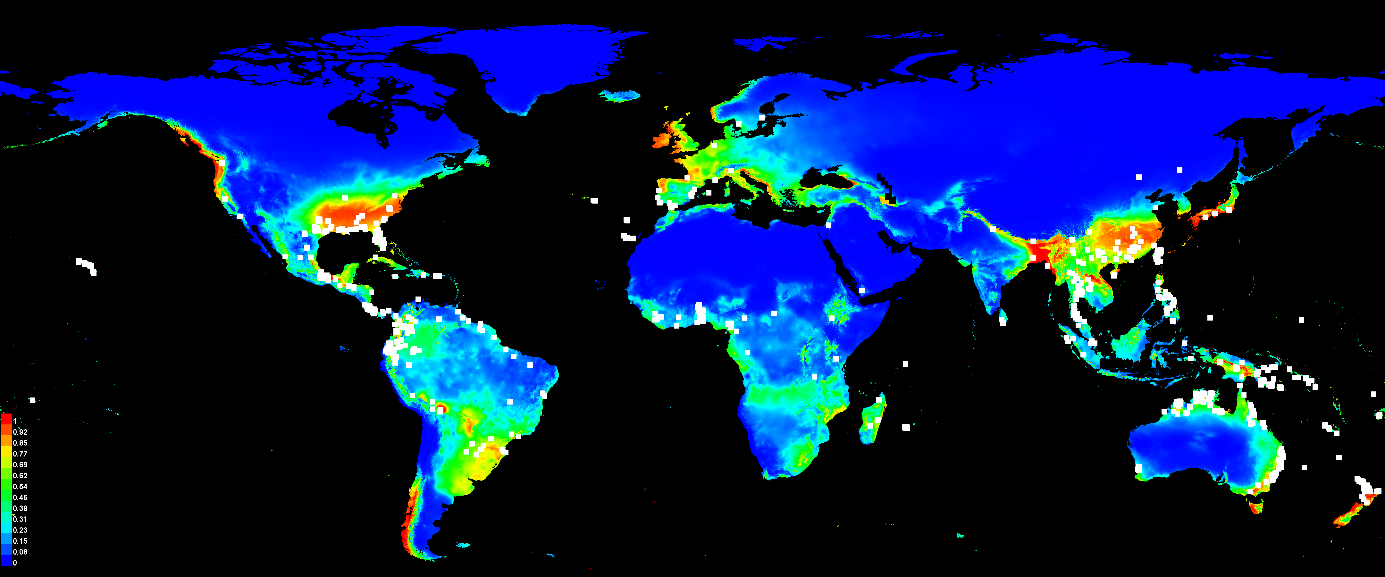
First I tried to run the model with data from the Solomon islands, according to that prediction the taro should mainly occur in Greenland. This obviously doesn’t make sense for this species and it was mainly caused by the variable bio15, precipitation seasonality.

I decided to run the model again with global data and with the variable selection I discarded bio15. When selecting the variables I started with correlations over 0.9 and I discarded the variable that I believed was less important for the taro. I repeated this several times, each time taking out the highest correlations and I ended up with bio2 (mean diurnal range), bio5 (max temperature of warmest month), bio6 (min temperature of coldest month), bio12 (annual precipitation) and bio14 (precipitation of driest month).

For the model in maxent I kept the automatic settings, because it uses different ways to make a prediction.

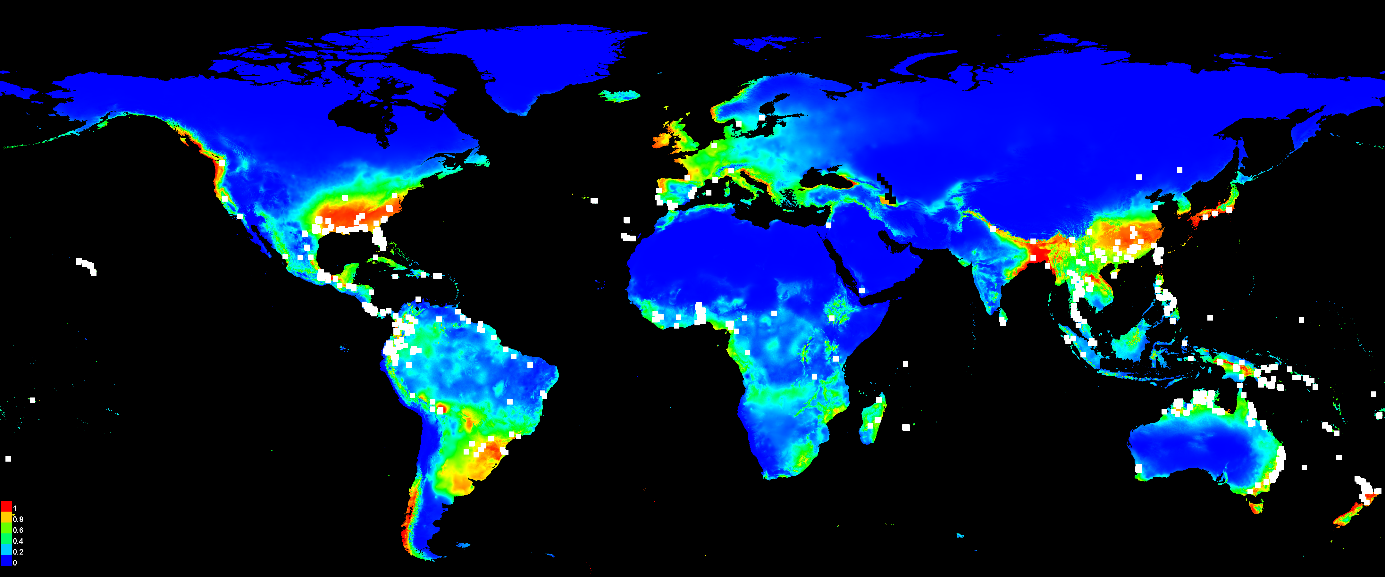
**Model output**

Fig. 1 shows the prediction of distribution of the taro in the present based on the climatic conditions. The distribution occurs mainly in southeast Asia, south Brasil, southern chili, southeast and western United States and western Europe.



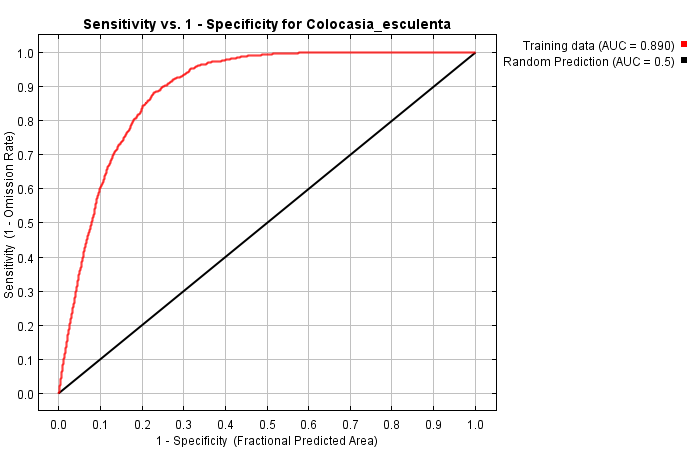
*Fig. 1: Predicted distribution of C. esculenta in the present. Warmer colors show better predicted conditions. White dots are locations used as training set.*

Fig. 2 shows the predictions for the distribution of *C. esculenta* for 2070. The distribution has not changed much in comparison to the present distribution.



*Fig. 2: Predicted distribution of C. esculenta in 2070. Warmer colors show better predicted conditions. White dots are locations used as training set.*

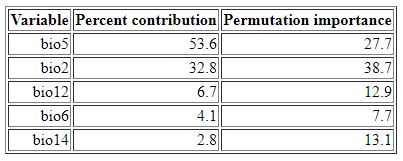
Fig. 3 shows a graph comparing the data from the training set against a random prediction, it gives a AUC value of 0.890 for the training set.



*Fig. 3: Graph calculating the AUC for the training set compared to a random prediction.*

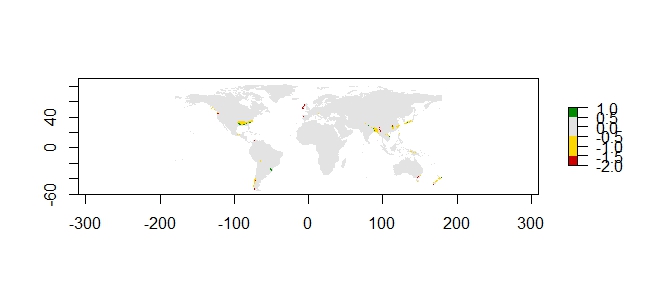
Table 1 shows the contribution table of the variables used for the prediction. Bio5 (maximum temperature of the warmest month) contributes most to the prediction with 53.6% and bio2 (mean diurnal range) contributes to the prediction with 32.8%.

*Table1: Variable importance. Shows the contributions of the variables in the predictions.*



**Response to future scenario**

In the distribution change map (Fig. 4) between the present and the future distributions, suitable areas remain in the same locations with small range shifts.



*Fig. 4: Distribution change map. Grey: never suitable. Yellow: remains suitable. Red: area is lost in the future. Green: area is gained in the future.*

**Biological interpretation**

According to the model the distribution of *C. esculenta* will not change much in 2070, but there will be a range shift. Because much of the habitat is expected to remain suitable, I do not expect much to change for the taro, it will probably follow the climate range shift.

The selected model performs well according to the AUC value of 0.890, but some distributions are a bit strange, like the distribution in Europe, southern chili and northeastern United States, since the taro mainly grows in humid environments around the equator.

I think this is partly because the dataset also includes museum specimens, for this the dataset should be cleaned. Another reason is that the variable that mainly contributed to the model was the maximum temperature of the warmest month, but for this species the minimum temperature is more important. Perhaps a different model would give more realistic results, or I could try a different variable selection.