**Predicting possible future habitats and farm locations for *Cocos nucifera* with a species distribution model**

*Simon Pieksma*

*s1676989*

*11-12-2019*

**Introduction**

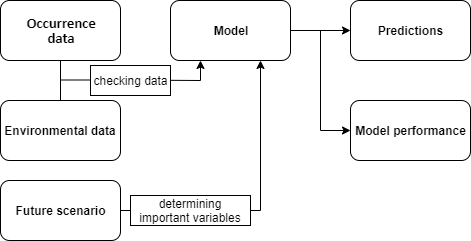
The coconut palm (*Cocos nucifera*) is a cultivated palm species. Every part of the tree is of some use but the palm is best known for its fruit. The insides of the ripe coconut fruit is edible and contains coconut water. Due to its useful properties many communities in the tropics rely heavily on the tree [1]. However, the coconut itself is actually not a nut but a drupe [2]. The palm tree has two known distinct varieties: The “talls” and the “dwarfs”. The “talls” being larger than their counter part and more commonly used in agriculture. Both varieties also have different subtypes. One subtype of the “talls” is believed to be the wild type or at least to be closely related to the original wild type. It should be noted that interestingly enough that the coconut tree does not have any closely related wild relatives, unlike many other cultivates crops [3].

The coconut is thought to have originated in the Indo-Malayan region and has subsequently spread after being cultivated and introduced to new locations by humans. Besides being spread by humans the coconut can also disperse via its seeds, which float in water and can thus travel long distances by sea. The seeds of the coconut have reported to be found as far away as Norway [3]. In the present the species is found in and around the tropics, having a pantropical distribution. Even though sandy beaches are its natural habitat, most farmers choose to grow the tree in deep soils, since this ensures the best yields [4].The tree can be found outside of the tropics as well, were it will flower but will struggle to develop normal seeds. The main reason that the coconut can be found outside the tropics can be attributed to its ability to withstand temperatures of around 4 °C and can also endure periods of drought lasting up to three months.For optimal growth of the coconut palm a year round warm and humid climate is required [4].

The coconut has no big invasive potential since its spread can easily be controlled by humans, due to its relatively slow growth rate and low amount of seeds produced per tree. Therefore a significant change in distribution due to climate change is unlikely. However due to the recent increase in global temperatures due to the emission of greenhouse gasses it can be hypothesised that more areas will become suitable to farm coconuts. This means that many areas that previously were too cold to grow the tree could potentially harbour successful plantations in the future. However climate change could also cause previous suitable locations to become unfavourable for coconut farming.

Since many communities rely on the coconut tree, it is essential to have information about the possible impacts climate change can have on the farming of coconuts. Therefore the aim of this research was to predict the change of the locations where the coconut can be farmed on a global scale. A species distribution model, based on present occurrences of the coconut tree, was used to gain insight into the environmental preferences of the coconut palm and to predict the distribution using a future scenario. It is hypothesised that in the future some locations with low temperatures, such locations closer to Europe, will become more suitable for farming. However some areas might also become unfavourable in the future.

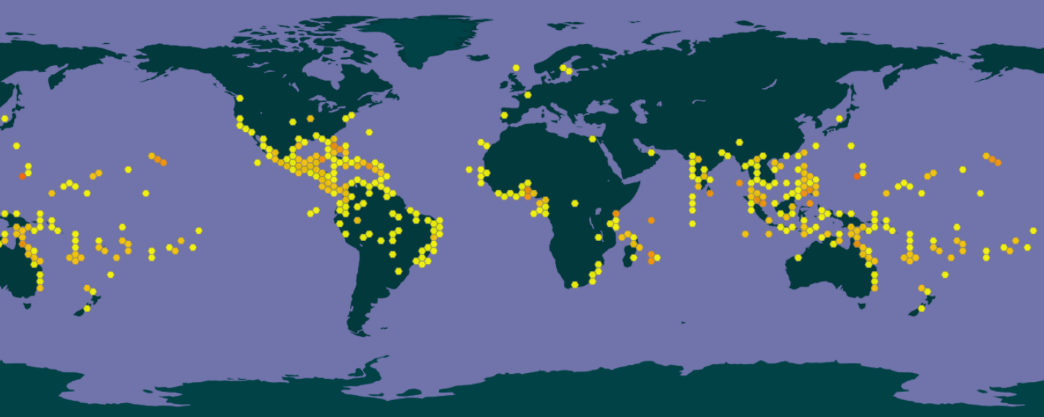
**Methodology and results**



*Fig 1. A flow chart of the steps taken to build the model and test the research hypothesis.*

*Occurrence data modifications*

Occurrence data was obtained from [www.gbif.org](http://www.gbif.org) [5]. The occurrences of the coconut palm are mainly centred around the equator. Many occurrences are on or close to a coastline. This is consistent with literature on the environmental preferences of the species. Based on the data it is apparent that the tree is also observed more inland in some continents. However absence of observation does not necessarily mean that the species does not occur at that location. Some locations, especially in the rainforests are hard to reach which results in less data on these areas. It should be noted that there are also occurrences in colder areas like Europe and north America. For a study on the natural habitat of the coconut palm these occurrences would have to be removed from the data set. However, since this study looked into all possible habitats and farm locations for the palm and according to the occurrence data the palm can grow in these areas after being introduced by humans, these data points were not removed. The occurrence data was edited so that inconclusive occurrences, with the latitude and/or longitude missing, were no longer included.



*Fig 2. The occurrence data as obtained from* [*www.worldclim.org*](http://www.worldclim.org) *[6]*

*Environmental data modifications*

The environmental data was downloaded from [www.worldclim.org](http://www.worldclim.org) [6]. Data from version 1.4 was used since this contained both present and future conditions. For both conditions a spatial resolution of 5 minutes was chosen. The datasets contained 19 initial bioclimatic variables which were mainly based on temperature and precipitation. For the environmental data of the future conditions the Representative Concentration Pathway 8.5 scenario of 2050 was chosen [7]. This scenario is one of the more extreme estimates of global greenhouse gas emissions and could thus potentially indicate whether or not this can influence the coconut distribution. The statistical programs R (version 3.6.1)and R studio [8] were used to crop the data and determine the importance of each bioclimatic variable. The variables were tested for autocorrelation and variables which had a high autocorrelation (r>0.7) were removed. Removal of these variables was also based on ecological knowledge. For example the coconut does not survive temperatures below 0 °C, therefore the lowest temperature of the coldest month was deemed an important variable and thus kept in the data set [4]. Afterwards the remaining variables were tested for multicollinearity. The temperature seasonality and temperature annual range variables both had an VIF score around 30. This suggested that one variable could be explained by the other. Therefore the temperature seasonality variable was subsequently removed since it had the highest VIF score. The variables left in the data set and their codes are shown in Table 1.

*Table 1. The codes used for the variables.*

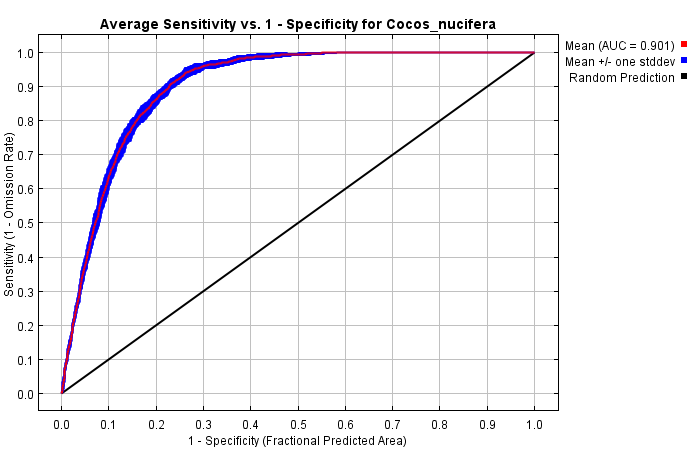
|  |  |
| --- | --- |
| Variable code | Variable name |
| Bio1 | Annual mean temperature |
| Bio2 | Mean diurnal range |
| Bio5 | Max temperature of warmest month |
| Bio6 | Min temperature of coldest month |
| Bio7 | Temperature annual range |
| Bio12 | Annual precipitation |
| Bio13 | Precipitation of wettest month |
| Bio14 | Precipitation of driest month |
| Bio15 | Precipitation seasonality |
| Bio 17 | Precipitation of driest quarter |

*Model settings*

The open source program Maxent (version 3.4.1) was used to make predictions based on the occurrence and environmental data [9]. The number of replicates to train the model was set to five. This resulted in a more accurate representation since any extremities of one replicated were averaged by the other replicates. The predictions were extrapolated, since it was assumed that areas close by an occurring coconut palm were also a suitable habitat. The extrapolation option was also turned on because the possibility that some observations were missing from certain areas close to current occurrences was kept in mind.

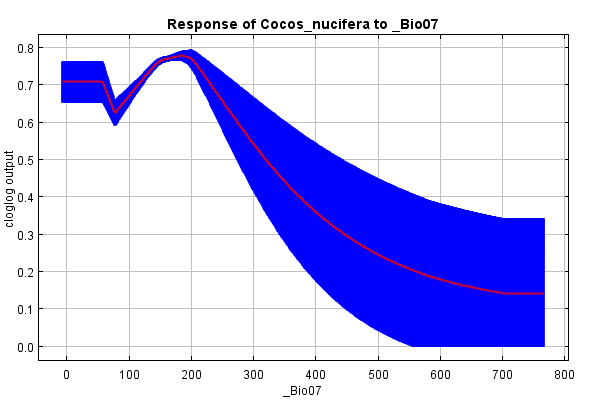
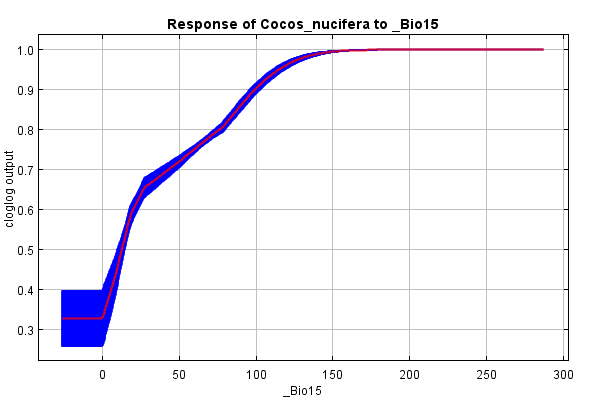
*Model diagnostics*

An receiver operating characteristic (ROC) curve was used as a measurement for performance of the model. The value of the area under the curve (AUC) gives an indication of its accuracy . The AUC of the model in this study was 0.901 which is close to 1.0. Meaning that the performance of the model is adequate to draw conclusions.



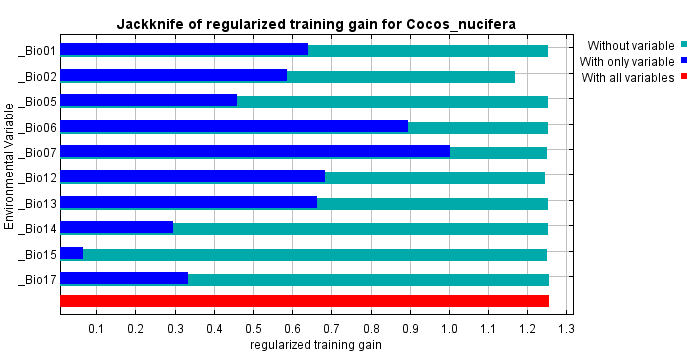
*Fig 3. The ROC curve averaged over the replicate runs.*

The response curves of the model shows how the probability of presence of the coconut tree changes when each variable is varied. Thus giving an indication of how well an variable can be used to predict the presence of the coconut tree and at which value the variable can reliably predict the presence of the coconut tree. The temperature annual range, as predicted by ecological information, is an important variable in predicting the occurrences (Fig 4A, Fig 5). The probability of presence declines with an increase in this variable. This means that in areas where the temperature fluctuates heavily between seasons, the coconut is less likely to occur. This corresponds with the known habitat of the coconut, which is in the tropics, where temperatures stay warm and relatively consistent year round.

 *A* *B*

*Fig 4. The response curve of two of the variables: The temperature annual range (A) and precipitation seasonality (B). The y axis shows the predicted probability and on the x axis are the different values of the variable.*

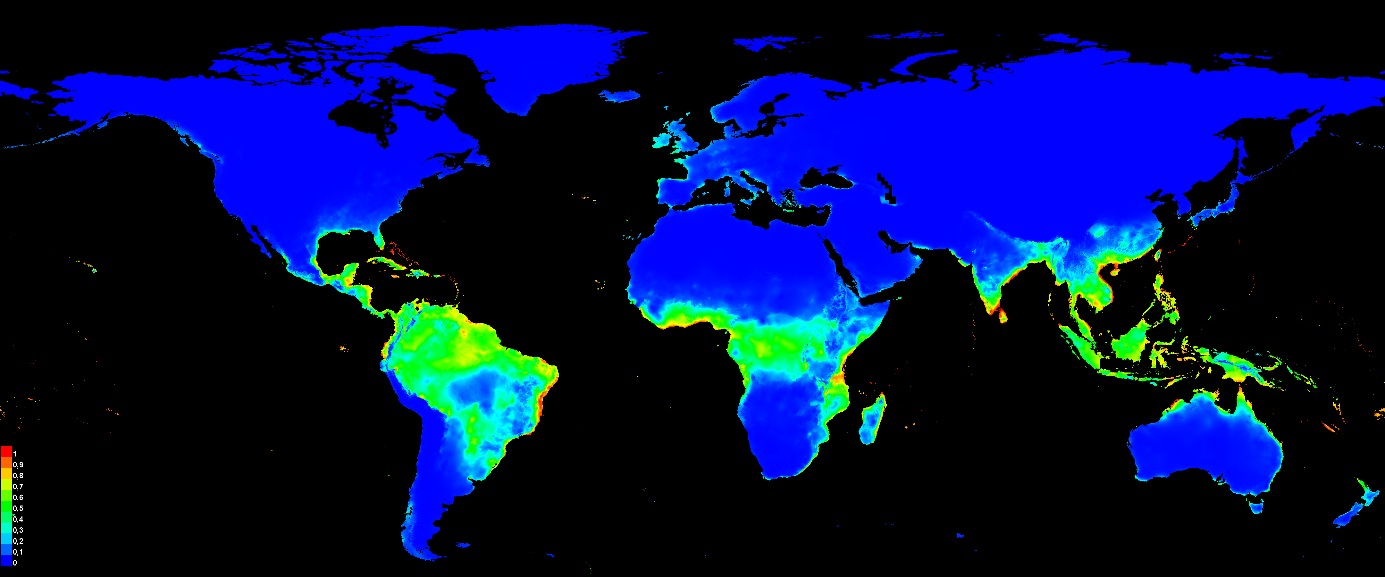
The precipitation seasonality variable contributed the least to the model (Fig 4B, Fig 5). This means that precipitation is less important for the coconut than temperature. With low precipitation the chance of presence is low however. Meaning that in very dry areas, like deserts, the coconut palm is not likely to occur.



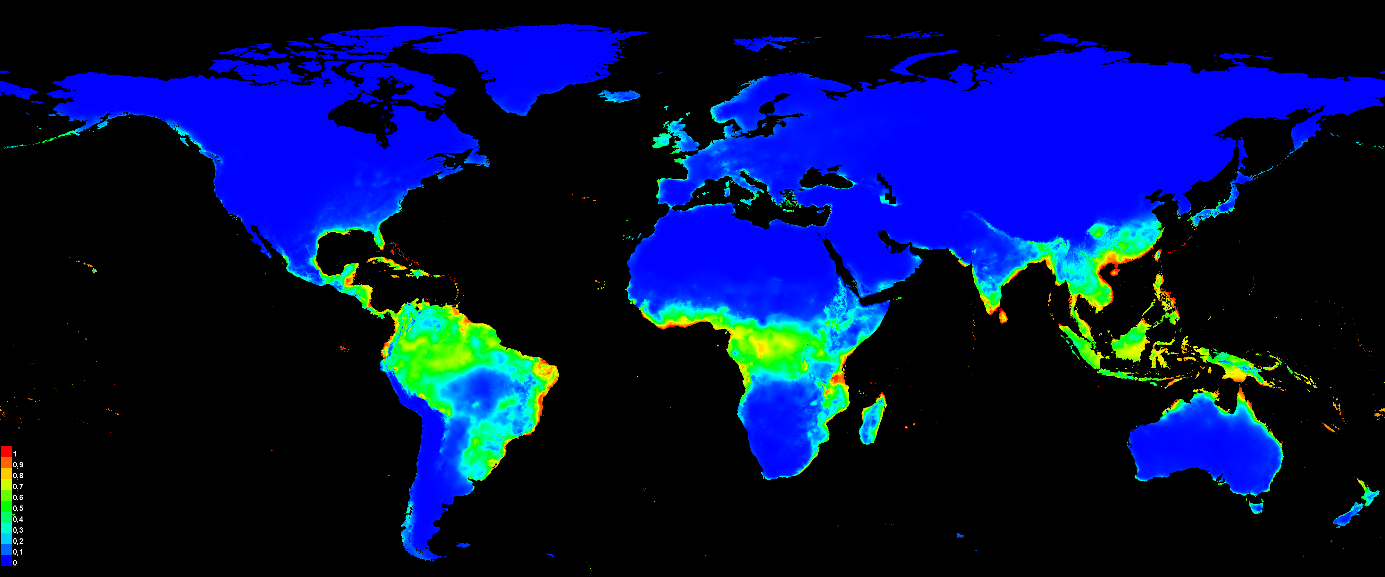
*Fig 5. The results of the jacknife test showing the importance of each variable. Bio7 appears to contain most the information and Bio2 contains the most information that is not present in other variables.*

*Model predictions*

For the present and future conditions predictions and estimates of the potential habitats were made. These predictions can also be seen as sites were the coconut tree can be grown and farmed. An comparison between the present and future conditions shows both an increase and decrease of potential suitable habitats (Fig 5). The total suitable habitats for the species increased. However in the north and centre of South America a significant decline can be seen. In the areas where the coconut palm already occurs, such as Africa and Indonesia, there has been an increase in suitable areas in which the coconut can be grown. Notably, in Europe and the coast of Scandinavia there is also an increase in suitable locations. This increase in colder regions can be attributed to the predicted rise of global temperatures in the future which make these regions less harsh during the winter period. Since most new potential habitats will from relatively close to the current habitats and the coconut can spread its seeds via water it will be possible that palm trees will start growing there in the future. As for more remote areas and also Europe the introduction of the palm in these regions will rely on humans.



*A*



*B*

*Fig 6 The current distribution (A) and future predicted distribution by the model(B) of the coconut palm. The colour indicates the probability that the coconut palm can be present and thus grown at that location (blue being the lowest and red being the highest probability).*

**Discussion and conclusion**

*Summary*

Globally the amount of suitable habitats for the coconut palm will increase in the Representative Concentration Pathway 8.5 scenario in 250. In the north and centre of South America there will be a significant decline according to the model. Parts of Europe might become suitable for coconut farming. This is in accordance to the hypothesis that some colder areas will become suitable due to a rise in temperature. Precipitation seems to be a less important factor than temperature to determine if a palm tree can survive in a certain area. This is because coconut trees can withstand long periods of drought but will die under conditions which are too cold. It should be noted that this scenario is an extreme and only one of multiple possible climate scenarios for the future. Therefore the model only provides evidence for what will happen might this extreme scenario be a correct prediction.

*Model improvements and limits*

Limited knowledge on the distribution limits of a species can negatively influence the amount of correct predictions by a model [10]. Even though, humans are largely responsible for the spread of the coconut and its seeds can travel great distances via water, more knowledge on which locations the coconut can reach on its own will improve the model. For example it can be hypothesised that the seeds might not reach certain islands by water due to strong currents in the ocean. The model could be improved by adding more variables. The current environmental data contained mainly variables derived from the temperature and precipitation. For future studies different variables could be added. The type of soil and its properties in an area could be useful to predict the suitability for coconut trees. As mentioned before the natural habitat of the coconut consists on sandy beaches, therefore the soil type and salinity could be factors worth adding to the model. Taking these additions into account to produce better models and testing other future scenarios will give a better understanding on how climate change might affect e distribution and farming locations of the coconut tree.

**References**

1. Arunachalam, V. Coconut. in *Genomics of Cultivated Palms* (2012). doi:10.1016/b978-0-12-387736-9.00002-9

2. Siriphanich, J. *et al.* Coconut (Cocos nucifera L.). in *Postharvest Biology and Technology of Tropical and Subtropical Fruits: Cocona to Mango* (2011). doi:10.1533/9780857092885.8

3. Jerard, B. A., Damodaran, V., Jaisankar, I., Velmurugan, A. & Swarnam, T. P. Coconut Biodiversity – Nature’s Gift to the Tropical Islands. in *Biodiversity and Climate Change Adaptation in Tropical Islands* (2008). doi:10.1016/b978-0-12-813064-3.00006-5

4. Hocher, V., Verdeil, J. L. & Malaurie, B. Cocos nucifera coconut. *Biotechnol. fruit nut Crop.* 90–112 (2009). doi:10.1079/9780851996622.0090

5. GBIF.org (5 December 2019) GBIF Occurrence Download <https://doi.org/10.15468/dl.shz5kd>

6. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. WorldClim. *International Journal of Climatology* (2005).

7. van Vuuren, D. P. *et al.* The representative concentration pathways: An overview. *Clim. Change* (2011). doi:10.1007/s10584-011-0148-z

8. R Development Core Team, R. *R: A Language and Environment for Statistical Computing*. *R Foundation for Statistical Computing* (2011). doi:10.1007/978-3-540-74686-7

9. Jane Elith, Steven J. Phillips, Trevor Hastie, Miroslav Dudík, Yung En Chee, and Colin J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17:43-57.

10. Carneiro, L. R. D. A., Lima, A. P., Machado, R. B. & Magnusson, W. E. Limitations to the use of species-distribution models for environmental-impact assessments in the Amazon. *PLoS One* (2016). doi:10.1371/journal.pone.0146543