

Global Forum on Biological Control and Training Workshop on Biological Control

Nairobi, Kenya 26-30 June 2023

Ecological Niche Modelling using MaxEnt software

Emily Kimathi

Co-organized by



Food and Agriculture
Organization of the
United Nations



icipe

Supported by



EUROPEAN UNION



USAID
FROM THE AMERICAN PEOPLE



Implemented by

giz Deutsche Gesellschaft
für Internationale
Zusammenarbeit (GIZ) GmbH

Background



- MaxEnt program for maximum entropy modelling of species' geographic distributions, was written by Steven Phillips, Miro Dudik and Rob Schapire, with support from AT&T Labs-Research, Princeton University, and the Center for Biodiversity and Conservation, American Museum of Natural History.

Steven J. Phillips, Robert P. Anderson and Robert E. Schapire, **Maximum entropy modeling of species geographic distributions**. Ecological Modelling, Vol 190/3-4 pp 231-259, 2006.

- From a set of environmental (e.g., climatic) grids and georeferenced occurrence localities, the model expresses a probability distribution where each grid cell has a predicted suitability of conditions for the species.



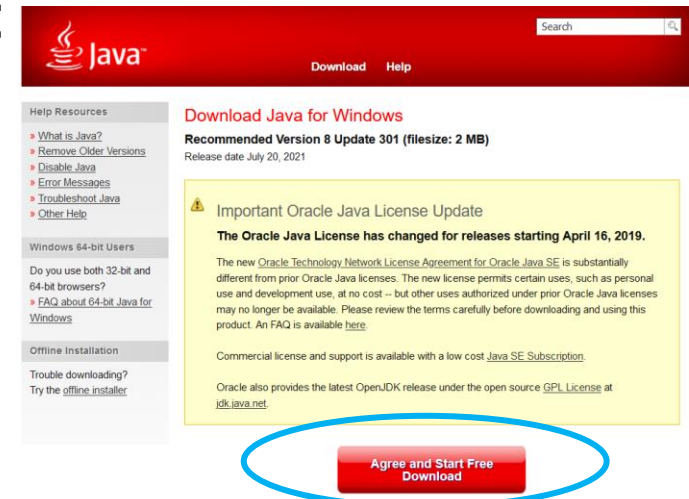
Food and Agriculture
Organization of the
United Nations



Software



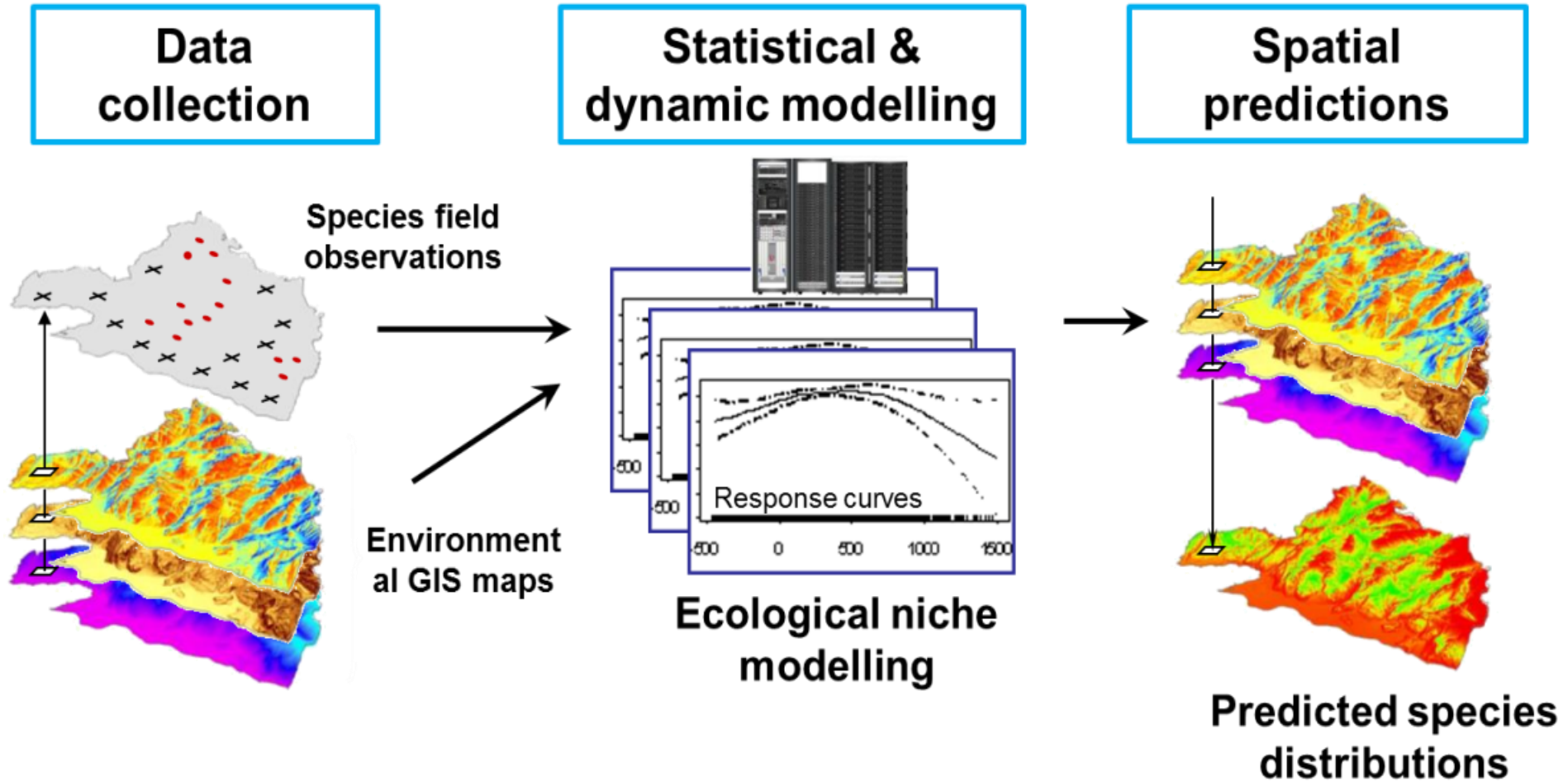
- The maxent software can be accessed at the Center for Biodiversity and Conservation at the American Museum of Natural History (AMNH) - https://biodiversityinformatics.amnh.org/open_source/maxent/
- The software consists of a jar file, maxent.jar, which can be used on any computer running Java version 1.4 or later.
- In case you do not have java, you can download it on this site: https://www.java.com/download/ie_manual.jsp



Food and Agriculture
Organization of the
United Nations



Methodology

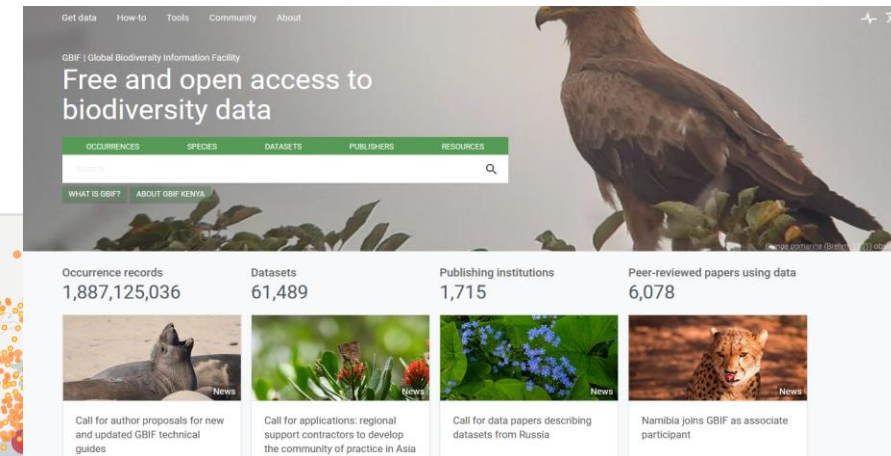
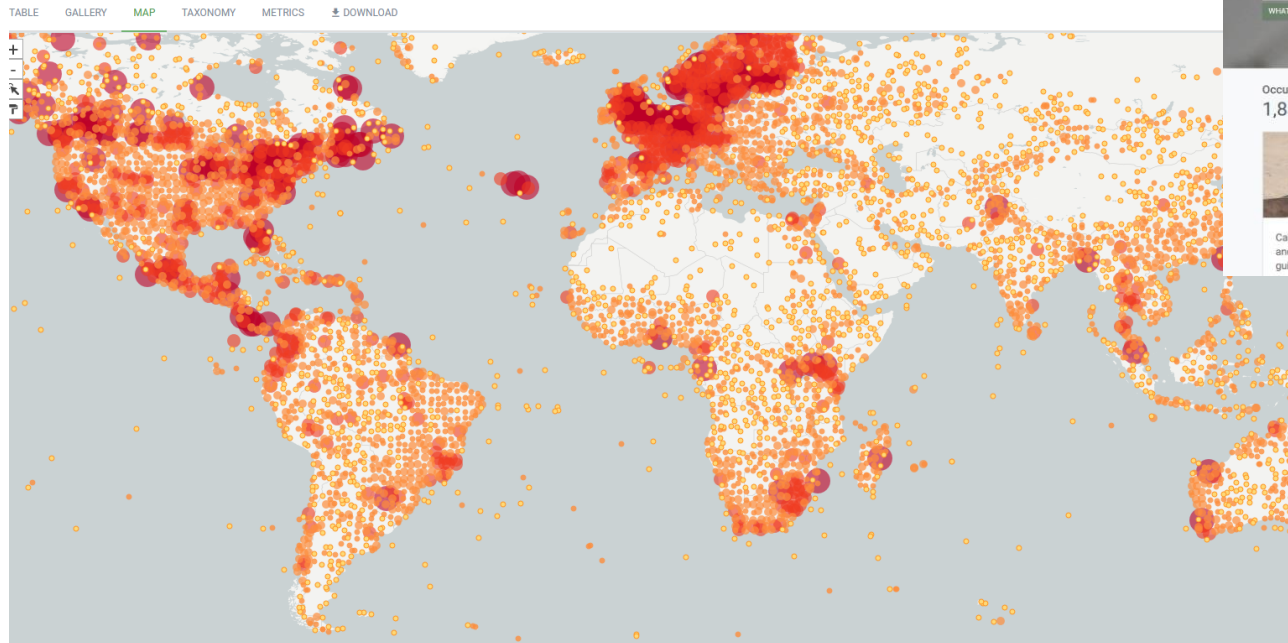
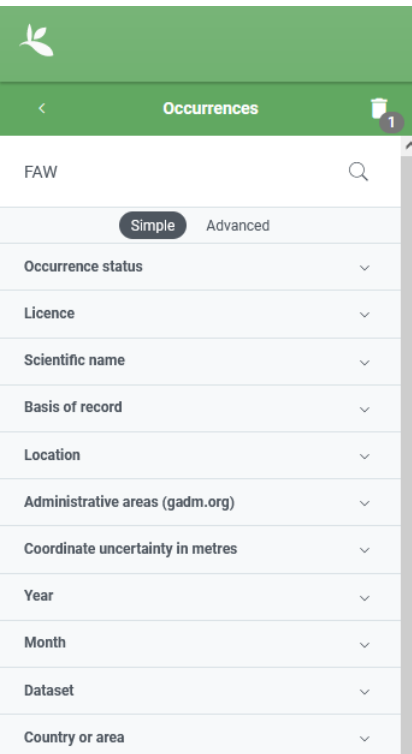


Source: Rpubs

Data sources – occurrence data



- Geo-referenced data can be collected in field using a GPS device.
- Secondary data can also be sourced from species data repositories.
- One of the largest species is the Global Biodiversity Information Facility (GBIF) - <https://www.gbif.org/>



Food and Agriculture
Organization of the
United Nations



Data sources – Environmental variables



- Predictor variables can be sourced from various sites depending on the parameters you want to use;

1. Climate data

<https://worldclim.org/data/bioclim.html>

<https://www.climond.org/>

<http://www.climatologylab.org/terraclimate.html>

<https://power.larc.nasa.gov/>

<https://data.chc.ucsb.edu/products/>

2. Soil data

<https://data.isric.org>

3. LULC data

<http://2016africalandcover20m.esrin.esa.int/>



Food and Agriculture
Organization of the
United Nations



Data preparation



- Occurrence data is saved in .csv format
- It should comprise of three columns: ID, Longitude and Latitude

	A	B	C
1	ID	longitude	latitude
2	FAW	35.05792	0.963873
3	FAW	34.83	0.78
4	FAW	37.7475	-1.8079
5	FAW	36.06968	-0.28524
6	FAW	35.06041	0.996134
7	FAW	35.0268	0.990905
8	FAW	36.23478	0.0759
9	FAW	34.36222	0.14047
10	FAW	37.66308	-0.08699
11	FAW	35.06504	0.986421
12	FAW	36.23128	0.118339
13	FAW	36.78481	-1.16892
14	FAW	36.87578	-1.17689
15	FAW	36.80074	-1.14803

- Predictor/environmental variables should be a series of a gridded dataset.
- The gridded dataset should be in .ascii format and with the same extent, resolution and CRS.



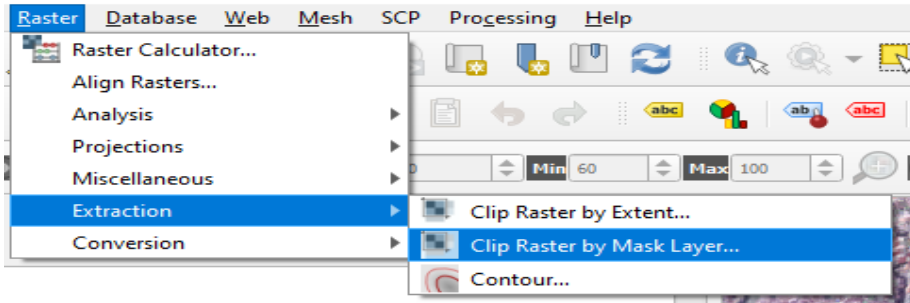
Food and Agriculture
Organization of the
United Nations



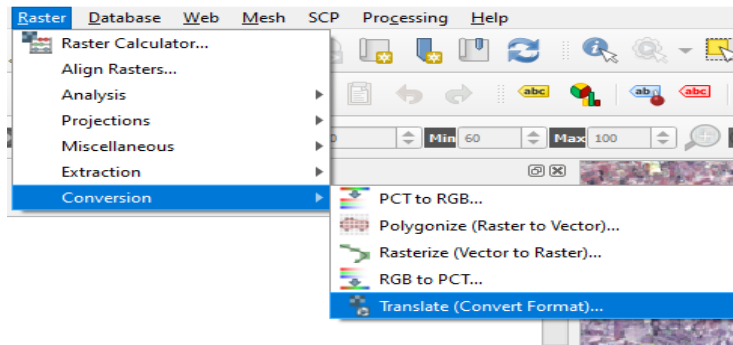
Data pre-processing



- To ensure variables have the same properties, clip the layers to the same extent.



- Once the layers are clipped they need to be resampled to the same resolution (columns/rows) of the reference layer.
- The last step is converting the raster files to .ascii format



Food and Agriculture
Organization of the
United Nations



Maxent user interface

- Navigate to ...\\FAW Global forum training 2023\\Modelling\\Software\\maxent

Maximum Entropy Species Distribution Modeling, Version 3.4.1

Samples
File Browse

Environmental layers
Directory/File Browse

Occurrence data

Predictor variables

Features

- ☒ Linear features
- ☒ Quadratic features
- ☒ Product features
- ☐ Threshold features
- ☒ Hinge features
- ☒ Auto features

Create response curves ☐

Make pictures of predictions ☒

Do jackknife to measure variable importance ☐

Output format: Cloglog

Output file type: asc

Output directory Browse

Projection layers directory/file Browse

Run Settings Help

annual_precip

190.59 1703.44

50000

Jackknife of AUC for Elephant

Environmental Variable	Without variable	With only variable	With all variables
Distance_to_fields	0.44	0.58	0.62
Distance_to_water	0.44	0.58	0.62
NDVI	0.44	0.58	0.62
Slope	0.44	0.58	0.62

Environmental variables

Bioclimatic Variables



Variable	Description	Units
Bio1	Annual Mean Temperature	°C
Bio2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	°C
Bio3	Isothermality (BIO2/BIO7) (×100)	-
Bio4	Temperature Seasonality (standard deviation ×100)	-
Bio5	Max Temperature of Warmest Month	°C
Bio6	Min Temperature of Coldest Month	°C
Bio7	Temperature Annual Range (BIO5-BIO6)	°C
Bio8	Mean Temperature of Wettest Quarter	°C
Bio9	Mean Temperature of Driest Quarter	°C
Bio10	Mean Temperature of Warmest Quarter	°C
Bio11	Mean Temperature of Coldest Quarter	°C
Bio12	Annual Precipitation	mm
Bio13	Precipitation of Wettest Month	mm
Bio14	Precipitation of Driest Month	mm
Bio15	Precipitation Seasonality (Coefficient of Variation)	-
Bio16	Precipitation of Wettest Quarter	mm
Bio17	Precipitation of Driest Quarter	mm
Bio18	Precipitation of Warmest Quarter	mm
Bio19	Precipitation of Coldest Quarter	mm

<https://worldclim.org/data/index.html>



Food and Agriculture
Organization of the
United Nations

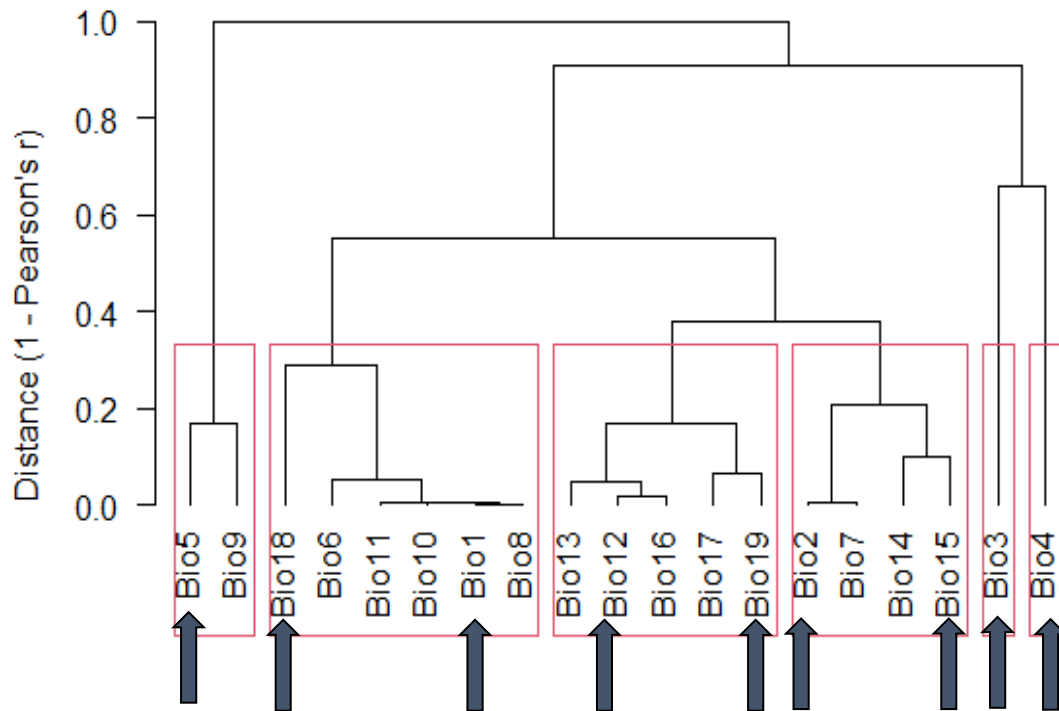


Multicollinearity



In statistics, multicollinearity is a phenomenon in which one predictor variable can be linearly predicted from the others with a substantial degree of accuracy.

Groups of intercorrelated variables at cutoff 0.7

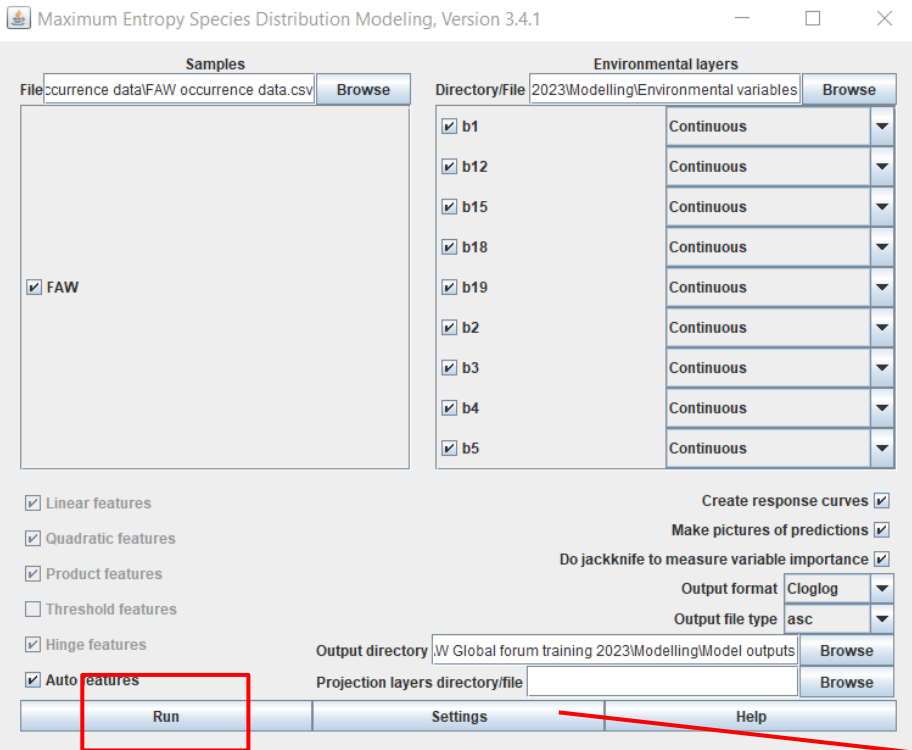


Select variables that are not collinear from the dendrogram.

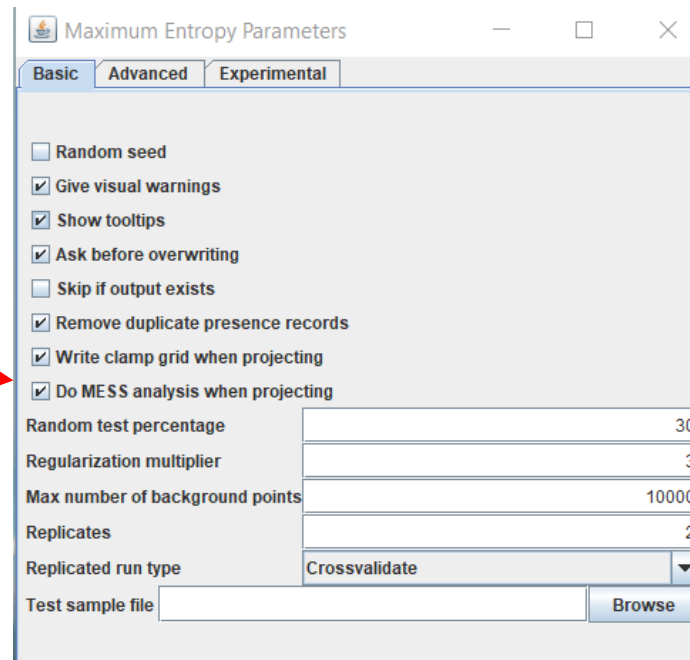
Methods

1. Pearson's
2. Spearman
3. Variance Inflation factor

Maxent user interface



- Supply a file containing presence geolocations, a directory containing environmental variables, and an output directory.
- Give a test percentage of **30** and regularization multiplier of **3**.
- Replicates = 2



- **Occurrence data** - ... \FAW Global forum training 2023\Modelling\Occurrence data\FAW occurrence data.csv
- **Predictor variables** - \FAW Global forum training 2023\Modelling\Environmental variables



Food and Agriculture
Organization of the
United Nations



Model output



Name	Date modified	Type	Size
plots	29/06/2023 14:42	File folder	
FAW	29/06/2023 14:42	Firefox HTML Document	8 KB
FAW_0.asc	29/06/2023 14:40	ASC File	9,536 KB
FAW_0	29/06/2023 14:40	Firefox HTML Document	12 KB
FAW_0.lambdas	29/06/2023 14:40	LAMBDA File	2 KB
FAW_0_omission	29/06/2023 14:40	Microsoft Excel Comma...	30 KB
FAW_0_sampleAverages	29/06/2023 14:40	Microsoft Excel Comma...	1 KB
FAW_0_samplePredictions	29/06/2023 14:40	Microsoft Excel Comma...	13 KB
FAW_1.asc	29/06/2023 14:40	ASC File	9,535 KB
FAW_1	29/06/2023 14:41	Firefox HTML Document	12 KB
FAW_1.lambdas	29/06/2023 14:40	LAMBDA File	2 KB
FAW_1_omission	29/06/2023 14:40	Microsoft Excel Comma...	30 KB
FAW_1_sampleAverages	29/06/2023 14:40	Microsoft Excel Comma...	1 KB
FAW_1_samplePredictions	29/06/2023 14:40	Microsoft Excel Comma...	13 KB
FAW_2.asc	29/06/2023 14:41	ASC File	9,536 KB
FAW_2	29/06/2023 14:41	Firefox HTML Document	12 KB
FAW_2.lambdas	29/06/2023 14:41	LAMBDA File	2 KB
FAW_2_omission	29/06/2023 14:41	Microsoft Excel Comma...	30 KB
FAW_2_sampleAverages	29/06/2023 14:41	Microsoft Excel Comma...	1 KB
FAW_2_samplePredictions	29/06/2023 14:41	Microsoft Excel Comma...	13 KB
FAW_avg.asc	29/06/2023 14:42	ASC File	9,536 KB
FAW_max.asc	29/06/2023 14:42	ASC File	9,536 KB
FAW_median.asc	29/06/2023 14:42	ASC File	9,536 KB
FAW_min.asc	29/06/2023 14:42	ASC File	9,536 KB
FAW_stddev.asc	29/06/2023 14:42	ASC File	9,536 KB
maxent	29/06/2023 14:42	Text Document	78 KB
maxentResults	29/06/2023 14:42	Microsoft Excel Comma...	10 KB



Food and Agriculture
Organization of the
United Nations

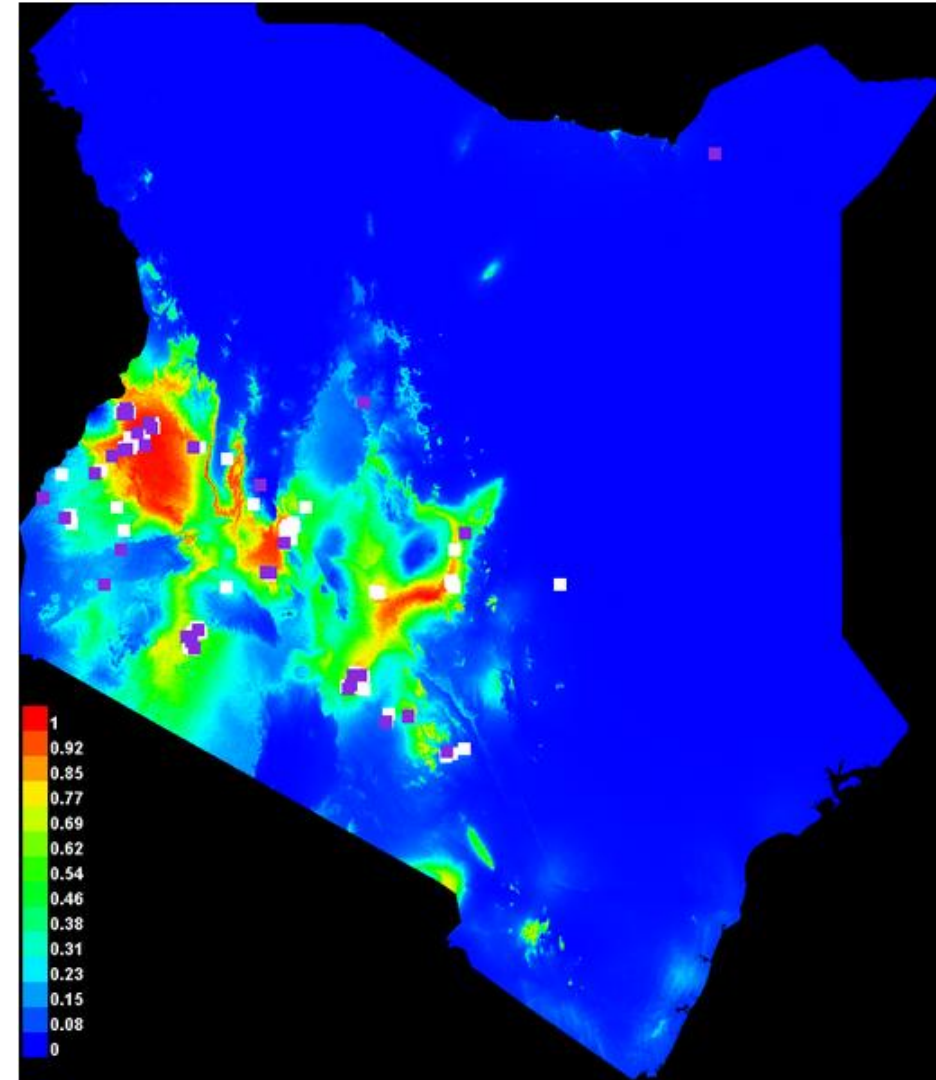


Model prediction

Click on the html file see a report of the results

plots	29/06/2023 14:42
FAW	29/06/2023 14:42
FAW_0.asc	29/06/2023 14:40
FAW_0	29/06/2023 14:40
FAW_0.lambdas	29/06/2023 14:40
FAW_0_omission	29/06/2023 14:40
FAW_0_sampleAverages	29/06/2023 14:40
FAW_0_samplePredictions	29/06/2023 14:40
FAW_1.asc	29/06/2023 14:40
FAW_1	29/06/2023 14:41
FAW_1.lambdas	29/06/2023 14:40
FAW_1_omission	29/06/2023 14:40
FAW_1_sampleAverages	29/06/2023 14:40
FAW_1_samplePredictions	29/06/2023 14:40

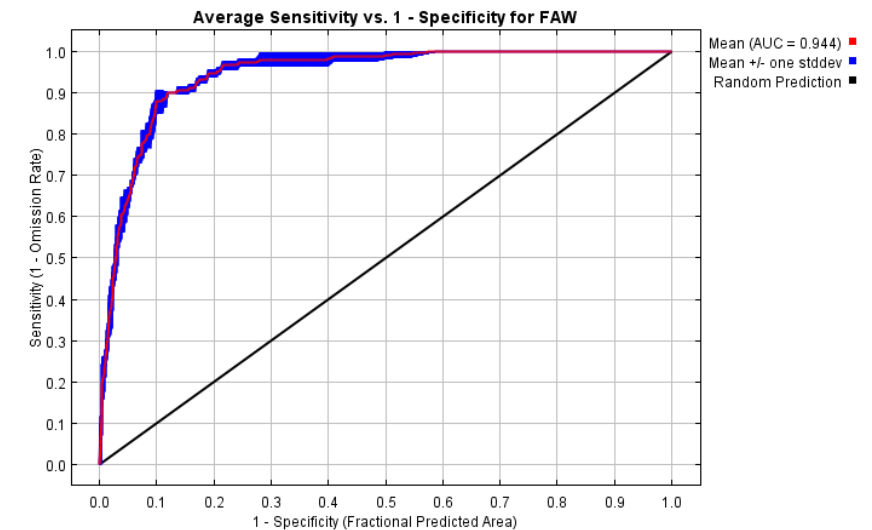
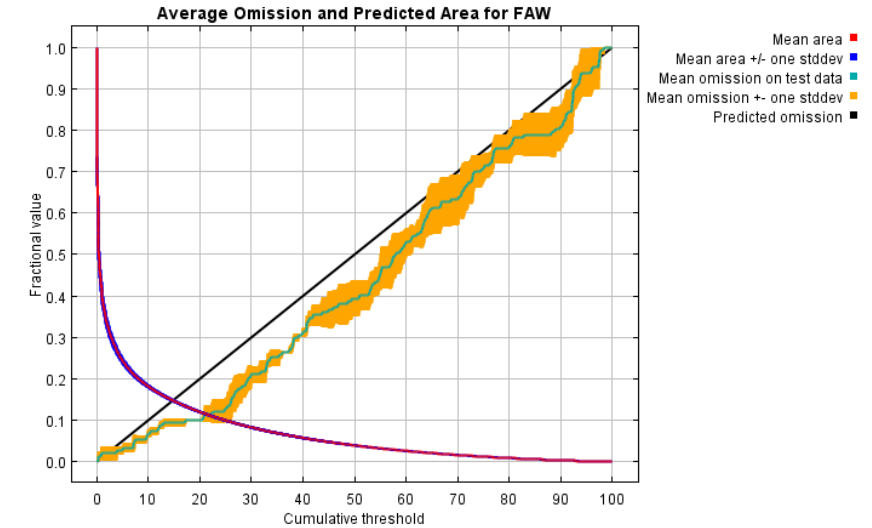
The image uses colors to indicate predicted probability with red indicating high probability of suitable conditions for the species, green indicating moderate, and lighter shades of blue indicating low predicted probability of suitable conditions.



Model validation



- The test data (30%) set aside is used to validate the model.
 - The first plot shows how testing and training omission and predicted area vary with the choice of cumulative threshold
 - The omission rate should be close to the predicted omission
-
- The second plot shows the area under the ROC curve (AUC).
 - This plot illustrates the relationship between Sensitivity (omission) and Specificity (commission).
 - AUC close to 1 shows high model's goodness of fit.



Variable contribution



- Analysis of variable contribution gives estimates of relative contributions of the environmental variables to the Maxent model.
- The right-hand column in the table shows permutation importance. This measure depends only on the final Maxent model, not the path used to obtain it.
- The contribution for each variable is determined by randomly permuting the values of that variable among the training points (both presence and background) and measuring the resulting decrease in training AUC.

Variable	Percent contribution	Permutation importance
b1	44.6	32
b12	35.9	28.8
b19	6.8	2.8
b2	5.5	13
b15	4.3	9.3
b3	1.2	4
b5	0.7	4.8
b4	0.5	3.5
b18	0.4	1.9

- A large decrease indicates that the model depends heavily on that variable. Values are normalized to give percentages



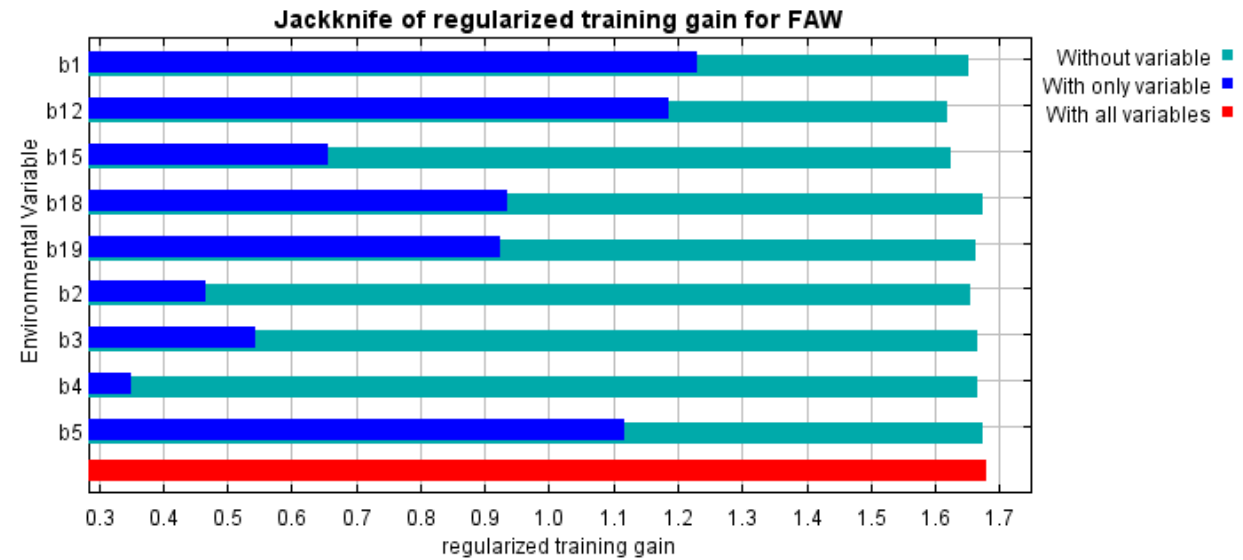
Food and Agriculture
Organization of the
United Nations



Jackknife test of variable importance

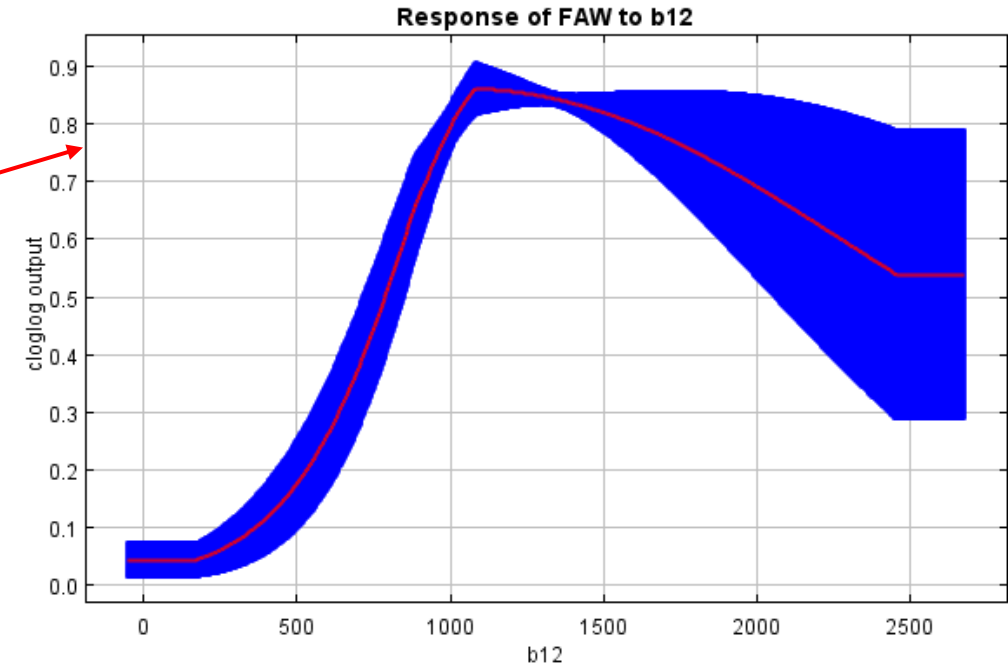
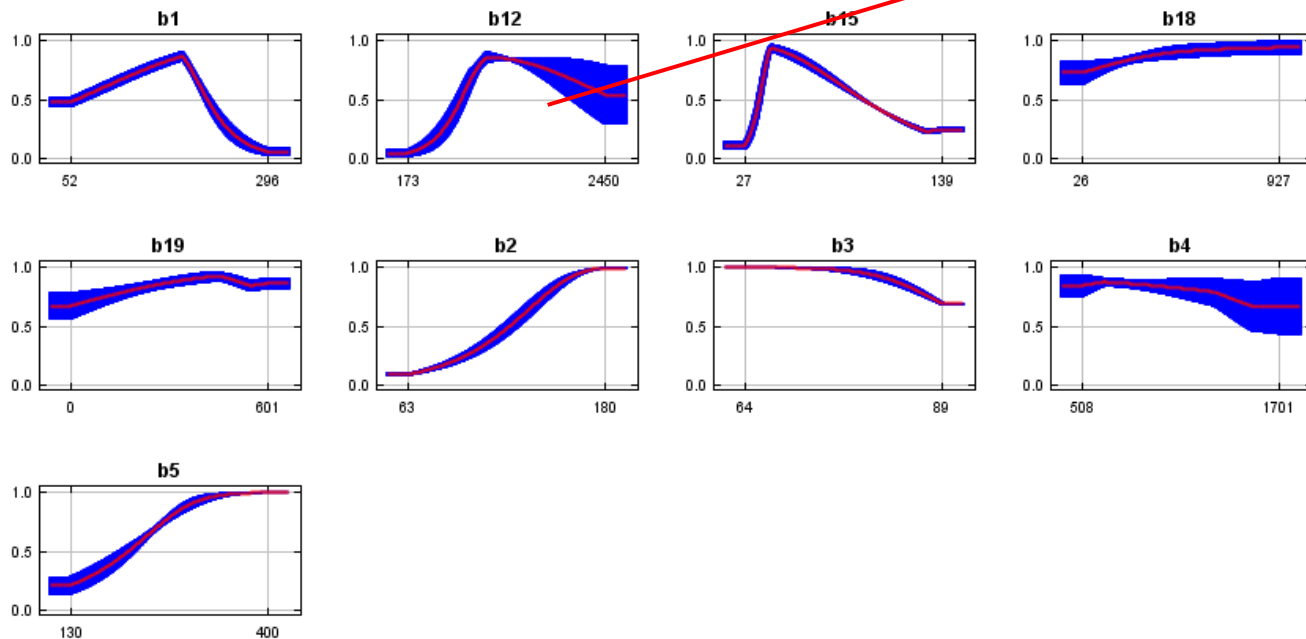


- Jack knife test provides the hierarchy of predictor variable importance when used in isolation and together with the rest of the variables.
- Each variable is excluded in turn, and a model created with the remaining variables.
- Then a model is created using each variable in isolation.
- In addition, a model is created using all variables, as before.



Response curves

- Response 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.

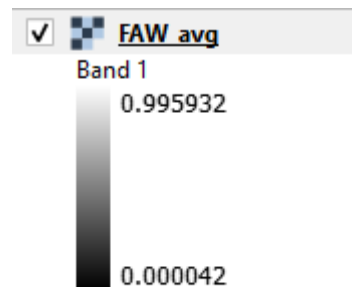


Model visualization and mapping



- Ascii output is the probability layer to be utilized for mapping

FAW_2_sampleAverages	29/06/2023 14:41	Microsoft Excel Comma
FAW_2_samplePredictions	29/06/2023 14:41	Microsoft Excel Comma
FAW_avg.asc	29/06/2023 14:42	ASC File
FAW_max.asc	29/06/2023 14:42	ASC File
FAW_median.asc	29/06/2023 14:42	ASC File
FAW_min.asc	29/06/2023 14:42	ASC File
FAW_stddev.asc	29/06/2023 14:42	ASC File
-----	29/06/2023 14:42	Text Document



- The raster layer values range between 0 – 1.
- Darker shades represent low probability of the species while lighter shades represents high probability.



Food and Agriculture
Organization of the
United Nations



Map development

- Create a color symbology for better visualization.



Layer Properties — FAW_avg — Symbology

Information
Source
Symbology
Transparency
Histogram
Rendering
Temporal
Pyramids
Elevation
Metadata
Legend
QGIS Server

Band Rendering

Render type: Singleband pseudocolor

Band: Band 1

Min: 0.0000417448 Max: 0.995932

► Min / Max Value Settings

Interpolation: Discrete

Color ramp: [Color ramp bar]

Label unit suffix:

Label precision: 1

Value <=	Color	Label
0.19921979584	[Dark Green]	<= 0.2
0.398397846880...	[Medium Green]	0.2 - 0.4
0.59757589792	[Light Green]	0.4 - 0.6
0.796753948960...	[Yellow]	0.6 - 0.8
inf	[Red]	> 0.8

Mode: Equal Interval

Classes: 5

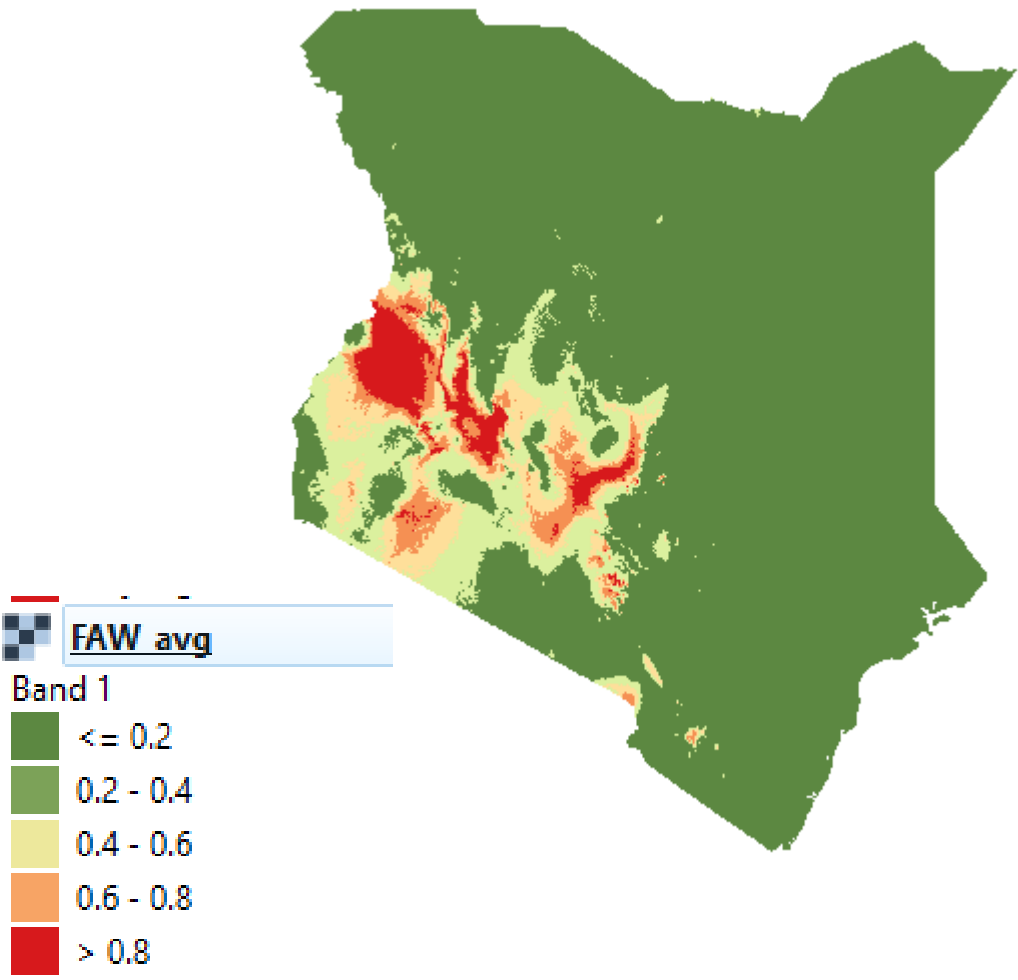
Buttons: Classify, Add, Remove, Refresh, Save, Legend Settings...

☐ Clip out of range values

Layer Rendering

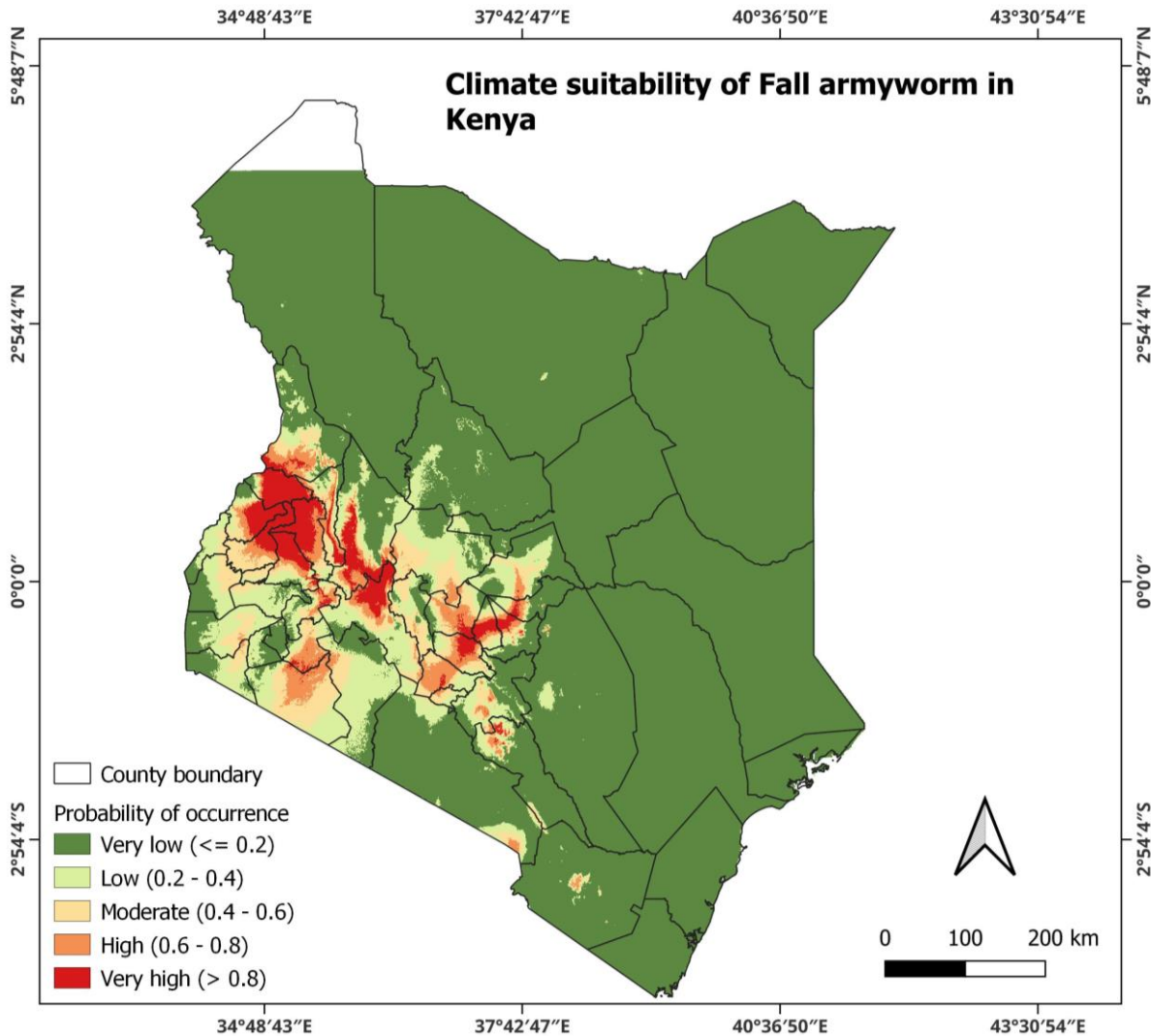
Blending mode: Normal

Buttons: OK, Cancel, Apply, Help



Map development

Map development



- Add all the map features to produce a map.



Food and Agriculture
Organization of the
United Nations





Thank you



Contact: eabdel-rahman@icipe.org,
ekimathi@icipe.org

Co-organized by



Food and Agriculture
Organization of the
United Nations



Supported by



EUROPEAN UNION



USAID
FROM THE AMERICAN PEOPLE



Implemented by

giz Deutsche Gesellschaft
für Internationale
Zusammenarbeit (GIZ) GmbH