Maxent model for species

This page contains some analysis of the Maxent model for species, created Mon Oct 07 22:37:06 EDT 2024 using Maxent version 3.4.3. If you would like to do further analyses, the raw data used here is linked to at the end of this page.

## Analysis of omission/commission

The following picture shows the omission rate and predicted area as a function of the cumulative threshold. The omission rate is is calculated both on the training presence records, and (if test data are used) on the test records. The omission rate should be close to the predicted omission, because of the definition of the cumulative threshold.

The next picture is the receiver operating characteristic (ROC) curve for the same data. Note that the specificity is defined using predicted area, rather than true commission (see the paper by Phillips, Anderson and Schapire cited on the help page for discussion of what this means). This implies that the maximum achievable AUC is less than 1. If test data is drawn from the Maxent distribution itself, then the maximum possible test AUC would be 0.958 rather than 1; in practice the test AUC may exceed this bound.

Some common thresholds and corresponding omission rates are as follows. If test data are available, binomial probabilities are calculated exactly if the number of test samples is at most 25, otherwise using a normal approximation to the binomial. These are 1-sided p-values for the null hypothesis that test points are predicted no better than by a random prediction with the same fractional predicted area. The "Balance" threshold minimizes 6 \* training omission rate + .04 \* cumulative threshold + 1.6 \* fractional predicted area.

| Cumulative threshold | Cloglog threshold | Description | Fractional predicted area | Training omission rate |
| --- | --- | --- | --- | --- |
| 1.000 | 0.046 | Fixed cumulative value 1 | 0.137 | 0.002 |
| 5.000 | 0.236 | Fixed cumulative value 5 | 0.103 | 0.022 |
| 10.000 | 0.359 | Fixed cumulative value 10 | 0.088 | 0.076 |
| 0.031 | 0.001 | Minimum training presence | 0.273 | 0.000 |
| 12.887 | 0.403 | 10 percentile training presence | 0.081 | 0.098 |
| 11.207 | 0.376 | Equal training sensitivity and specificity | 0.085 | 0.085 |
| 4.322 | 0.219 | Maximum training sensitivity plus specificity | 0.106 | 0.016 |
| 0.934 | 0.043 | Balance training omission, predicted area and threshold value | 0.138 | 0.002 |
| 3.718 | 0.191 | Equate entropy of thresholded and original distributions | 0.109 | 0.013 |

(A link to the Explain tool was not made for this model. The model uses product features, while the Explain tool can only be used for additive models.)

## Analysis of variable contributions

The following table gives estimates of relative contributions of the environmental variables to the Maxent model. To determine the first estimate, in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the corresponding variable, or subtracted from it if the change to the absolute value of lambda is negative. For the second estimate, for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is reevaluated on the permuted data, and the resulting drop in training AUC is shown in the table, normalized to percentages. As with the variable jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated.

| Variable | Percent contribution | Permutation importance |
| --- | --- | --- |
| bio19 | 81.9 | 79.7 |
| bio18 | 8.1 | 5.8 |
| bio9 | 3.8 | 4.4 |
| bio2 | 2.5 | 1.8 |
| bio8 | 2.4 | 2.6 |
| bio3 | 1.2 | 5.6 |

## Raw data outputs and control parameters

The data used in the above analysis is contained in the next links. Please see the Help button for more information on these.

[The model applied to the training environmental layers](http://docs.google.com/species.csv)

[The coefficients of the model](http://docs.google.com/species.lambdas)

[The omission and predicted area for varying cumulative and raw thresholds](http://docs.google.com/species_omission.csv)

[The prediction strength at the training and (optionally) test presence sites](http://docs.google.com/species_samplePredictions.csv)

[Results for all species modeled in the same Maxent run, with summary statistics and (optionally) jackknife results](http://docs.google.com/maxentResults.csv)

Regularized training gain is 2.215, training AUC is 0.961, unregularized training gain is 2.287.

Algorithm terminated after 500 iterations (11 seconds).

The follow settings were used during the run:

447 presence records used for training.

10418 points used to determine the Maxent distribution (background points and presence points).

Environmental layers used (all continuous): bio18 bio19 bio2 bio3 bio8 bio9

Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500

Feature types used: hinge product linear quadratic

outputdirectory: C:/Users/Mrochs/Documents/GitHub/Mroch-Homework/maxent

samplesfile: C:/Users/Mrochs/Documents/GitHub/Mroch-Homework/maxent/presence

environmentallayers: C:/Users/Mrochs/Documents/GitHub/Mroch-Homework/maxent/absence

autorun: true

visible: false

prefixes: false

Command line used: autorun -e C:/Users/Mrochs/Documents/GitHub/Mroch-Homework/maxent/absence -o C:/Users/Mrochs/Documents/GitHub/Mroch-Homework/maxent -s C:/Users/Mrochs/Documents/GitHub/Mroch-Homework/maxent/presence -z

Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E species outputdirectory=C:/Users/Mrochs/Documents/GitHub/Mroch-Homework/maxent samplesfile=C:/Users/Mrochs/Documents/GitHub/Mroch-Homework/maxent/presence environmentallayers=C:/Users/Mrochs/Documents/GitHub/Mroch-Homework/maxent/absence autorun novisible noprefixes