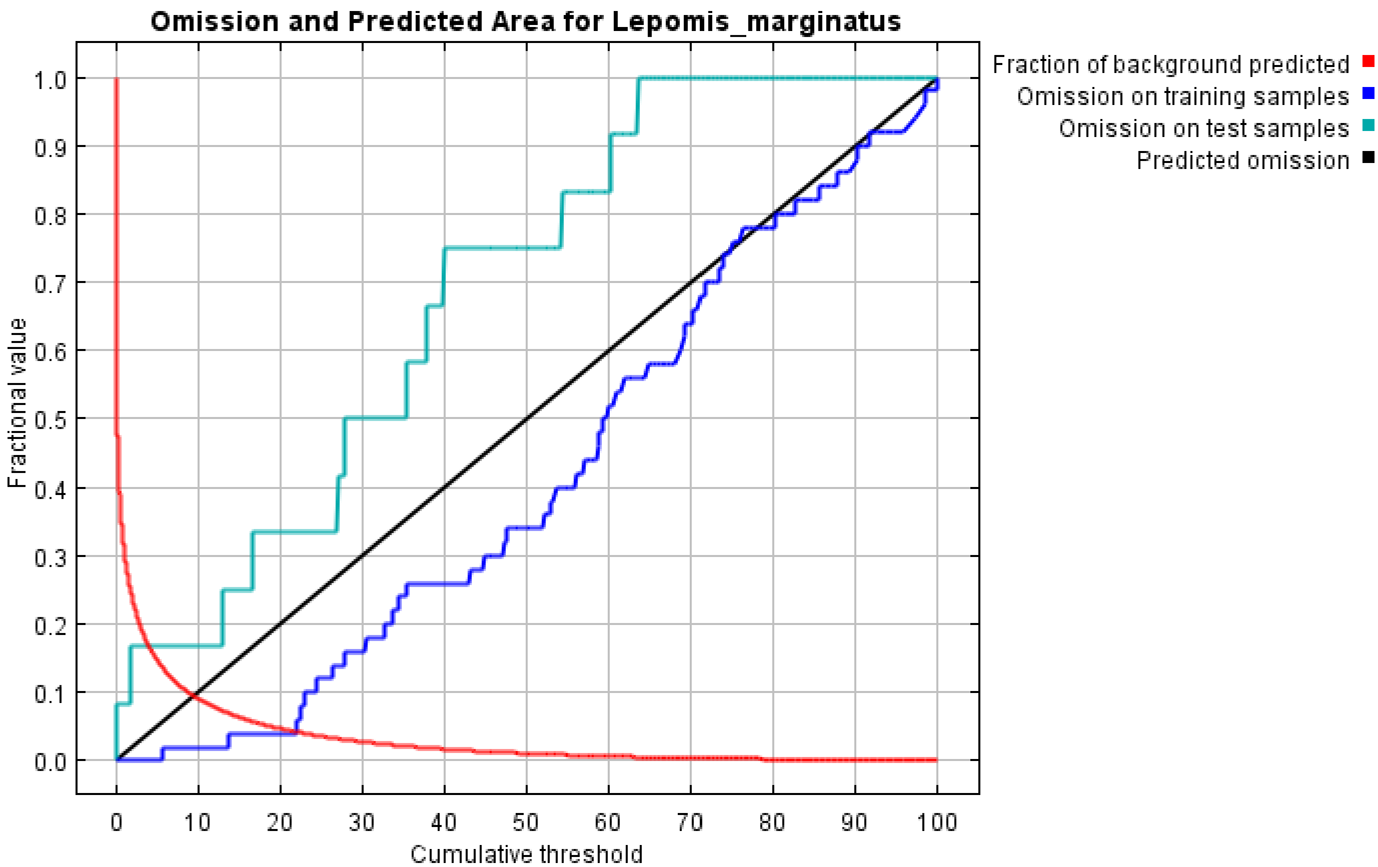
# Maxent model for Lepomis_marginatus

This page contains some analysis of the Maxent model for Lepomis_marginatus, created Sat Jan 25 09:37:23 EST 2014 using Maxent version 3.3.3k. If you would like to do further analyses, the raw data used here is linked to at the end of this page.

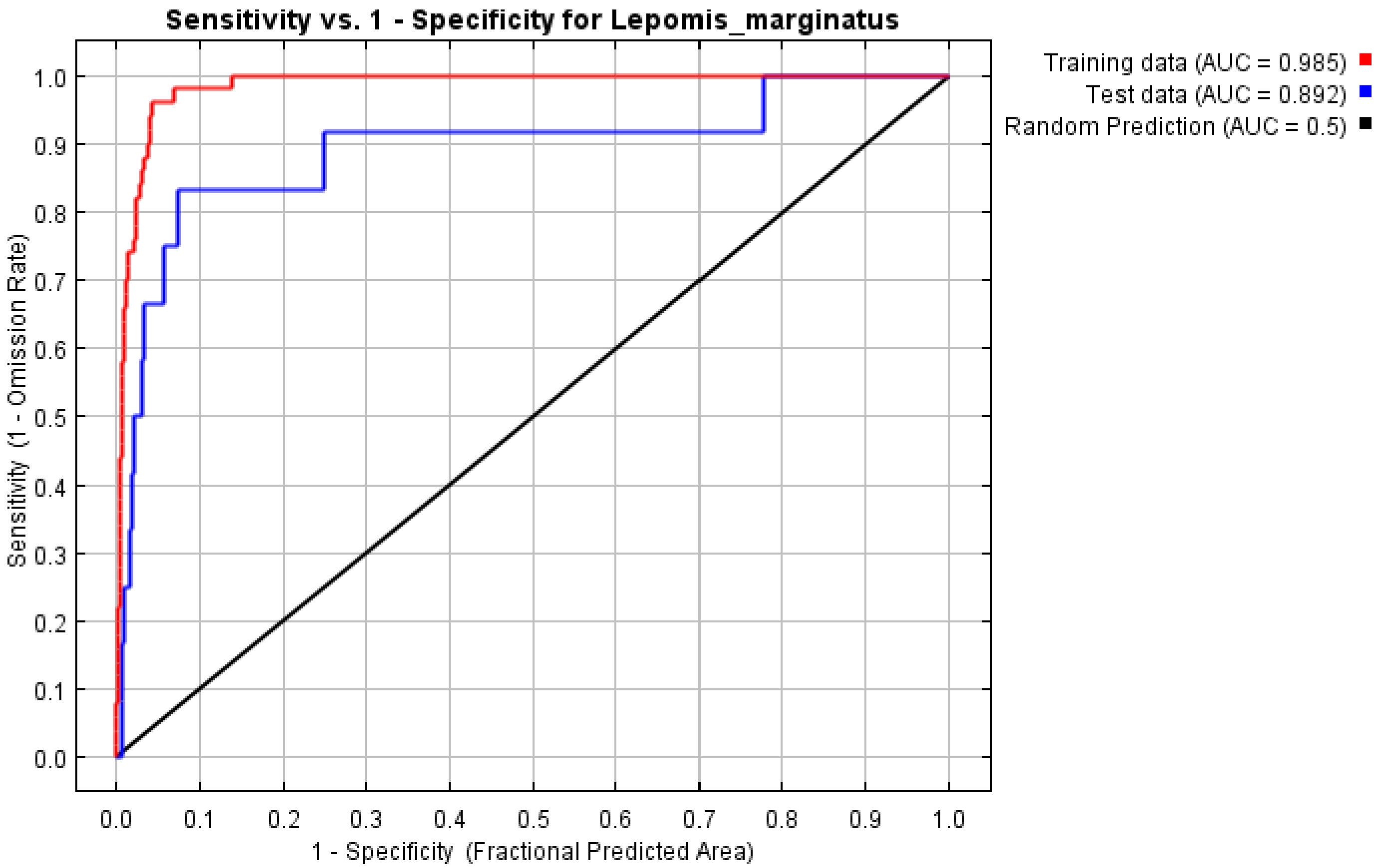
---

## 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 0.965 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.

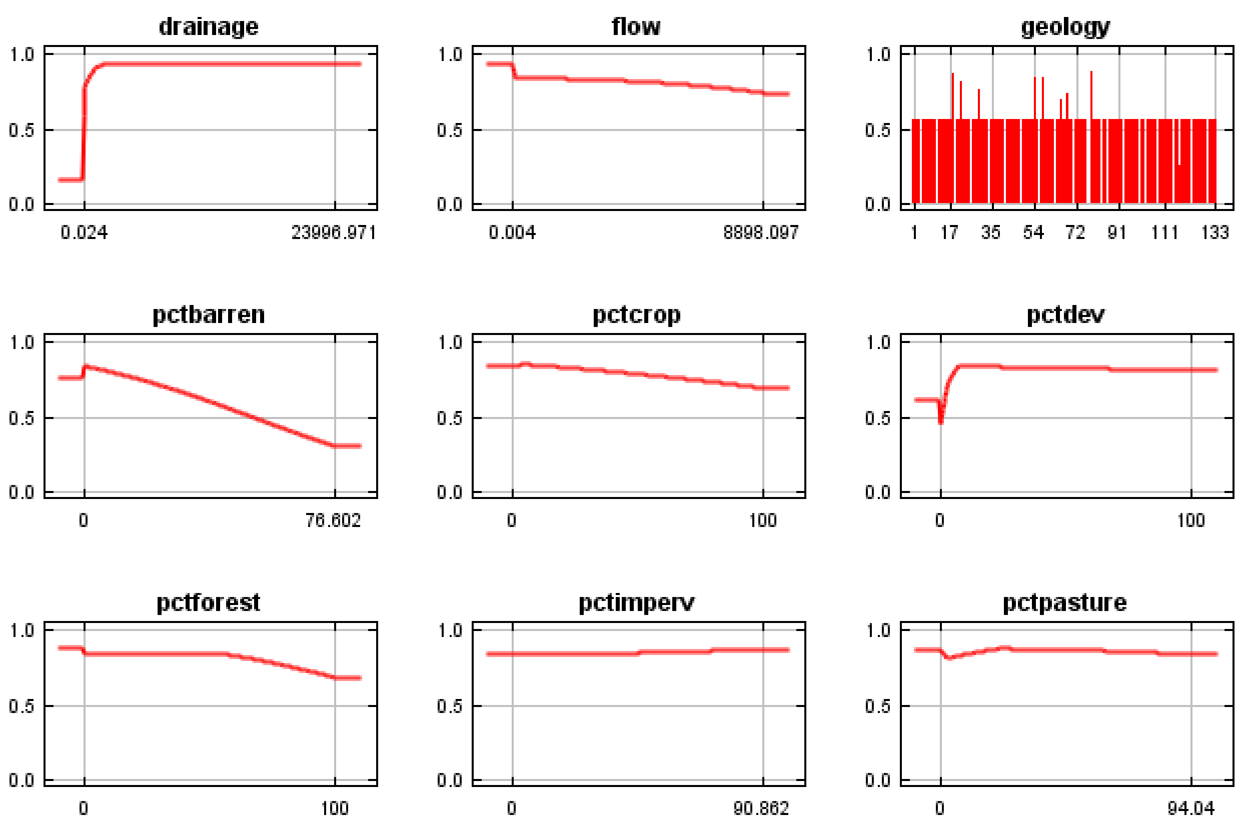
| **Cumulative threshold** | **Logistic threshold** | **Description** | **Fractional predicted area** | **Training omission rate** | **Test omission rate** | **P-value** |
|---|---|---|---|---|---|---|
| 1.000 | 0.007 | Fixed cumulative value 1 | 0.301 | 0.000 | 0.083 | 1.58E-5 |
| 5.000 | 0.038 | Fixed cumulative value 5 | 0.146 | 0.000 | 0.167 | 2.231E-7 |
| 10.000 | 0.090 | Fixed cumulative value 10 | 0.091 | 0.020 | 0.167 | 2.057E-9 |
| 5.498 | 0.042 | Minimum training presence | 0.138 | 0.000 | 0.167 | 1.269E-7 |
| 24.419 | 0.258 | 10 percentile training presence | 0.037 | 0.100 | 0.333 | 1.474E-9 |
| 21.909 | 0.231 | Equal training sensitivity and specificity | 0.042 | 0.040 | 0.333 | 4.298E-9 |
| 21.909 | 0.231 | Maximum training sensitivity plus specificity | 0.042 | 0.040 | 0.333 | 4.298E-9 |

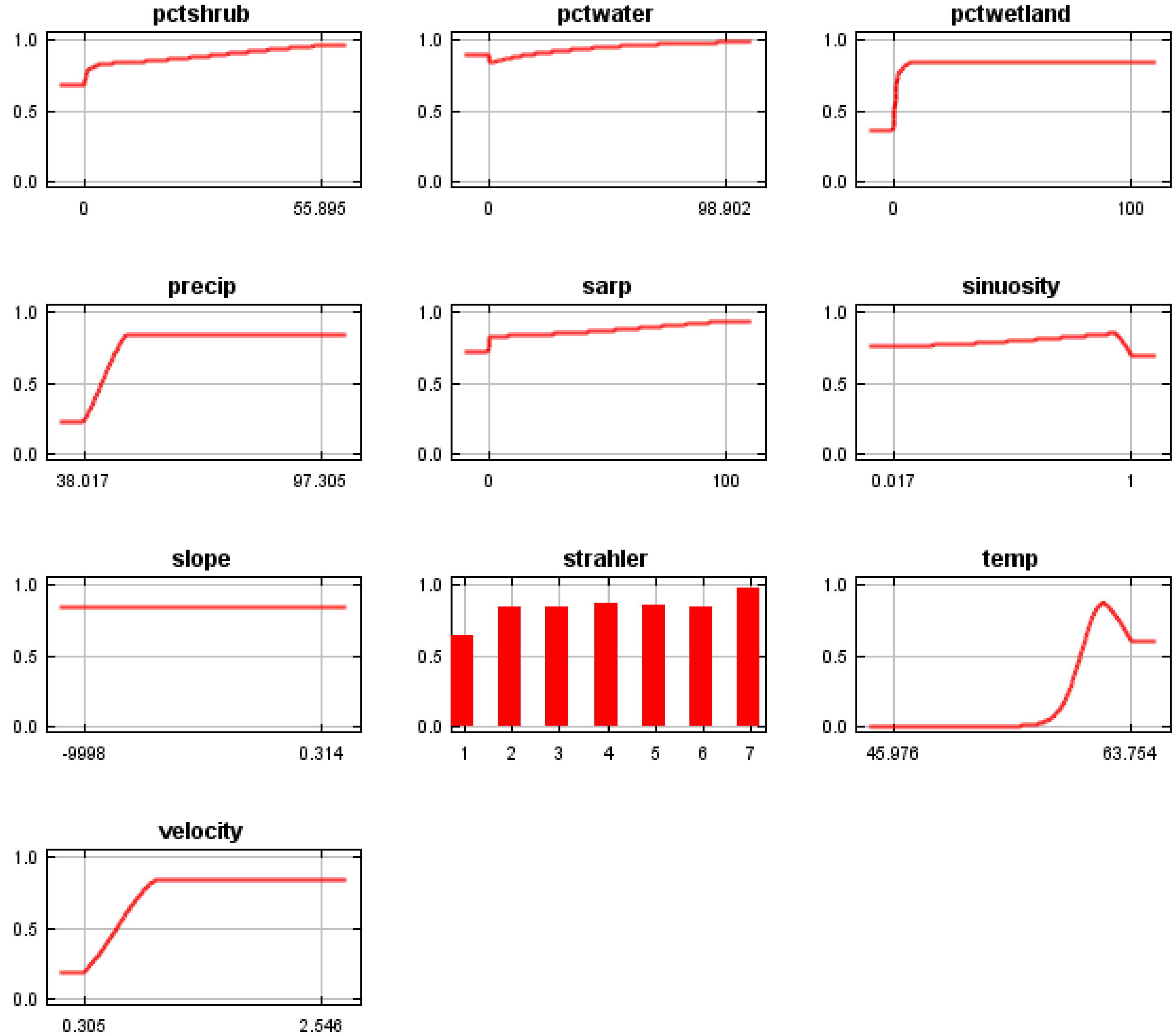
| | | | | | | |
|---|---|---|---|---|---|---|
| 3.997 | 0.030 | Equal test sensitivity and specificity | 0.167 | 0.000 | 0.167 | 7.845E-7 |
| 12.816 | 0.120 | Maximum test sensitivity plus specificity | 0.074 | 0.020 | 0.167 | 2.647E-10 |
| 3.743 | 0.027 | Balance training omission, predicted area and threshold value | 0.173 | 0.000 | 0.167 | 1.108E-6 |
| 13.698 | 0.131 | Equate entropy of thresholded and original distributions | 0.069 | 0.040 | 0.250 | 6.705E-9 |

Click here to interactively explore this prediction using the Explain tool. If clicking from your browser does not succeed in starting the tool, try running the script in D:\V2Maxent\Output\Lepomis_marginatus_explain.bat directly. This tool requires the environmental grids to be small enough that they all fit in memory.

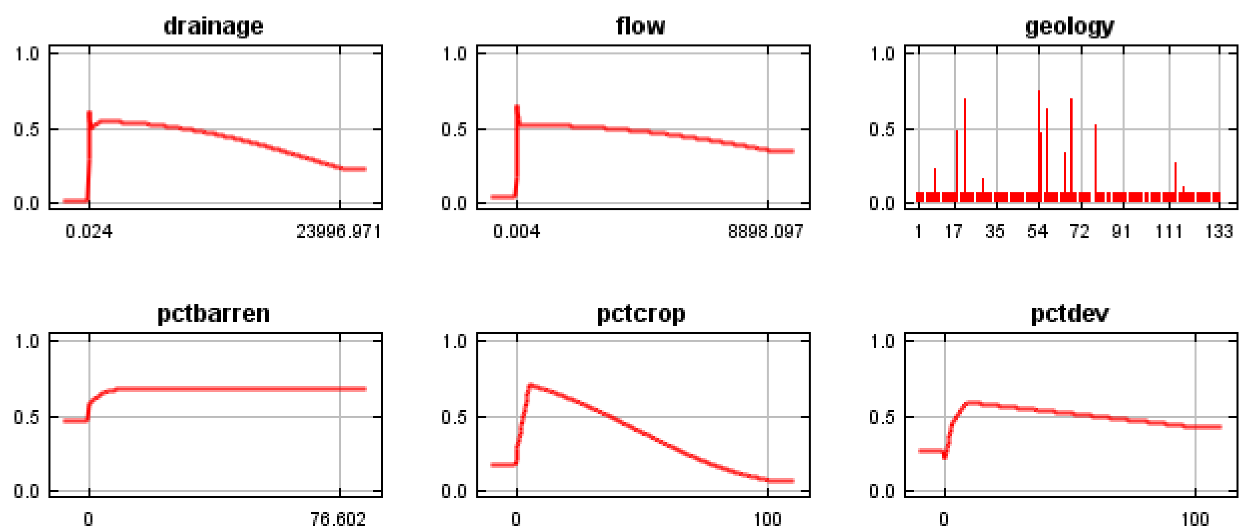
## Response curves

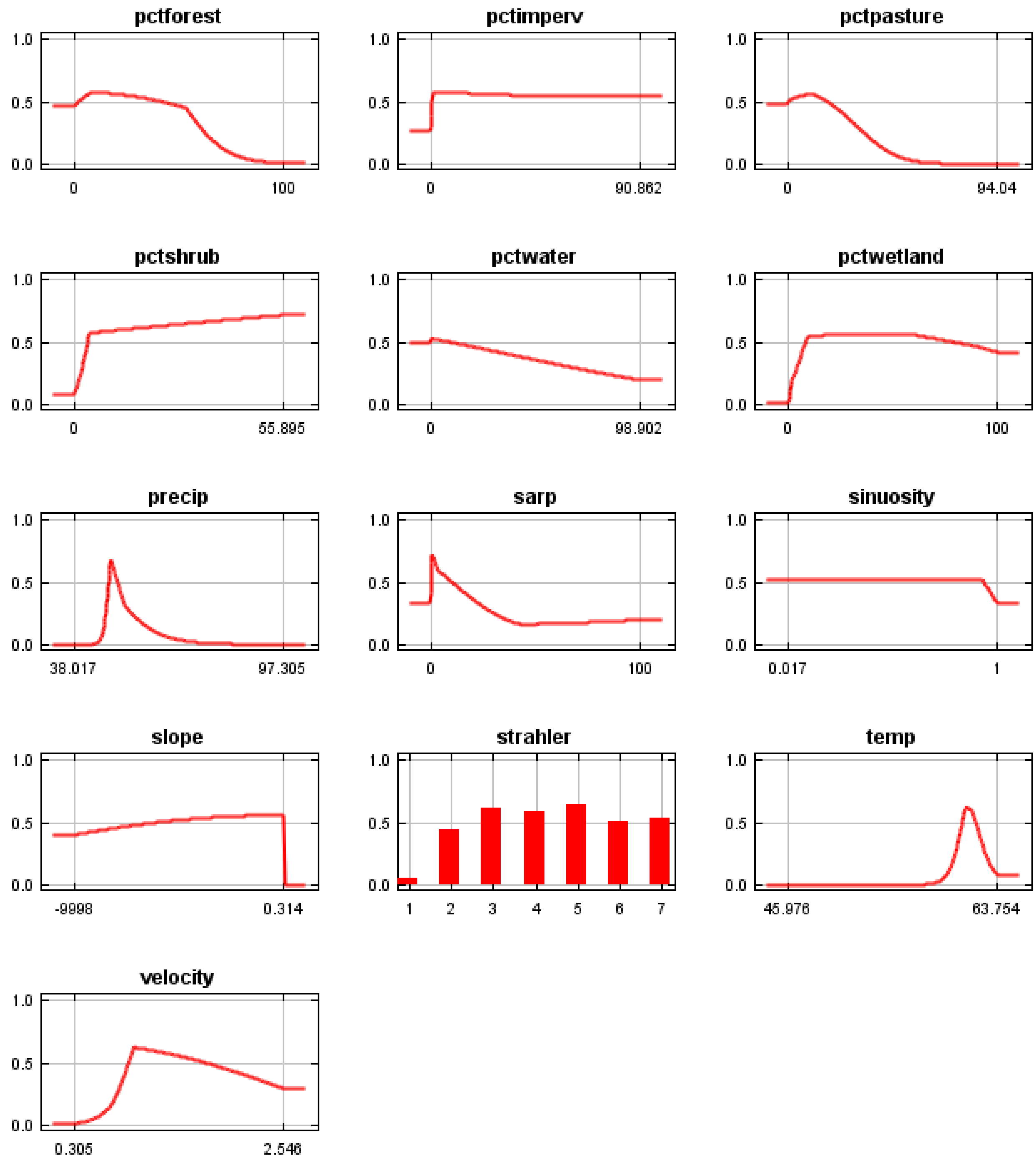
These curves show how each environmental variable affects the Maxent prediction. The curves show how the logistic prediction changes as each environmental variable is varied, keeping all other environmental variables at their average sample value. Click on a response curve to see a larger version. Note that the curves can be hard to interpret if you have strongly correlated variables, as the model may depend on the correlations in ways that are not evident in the curves. In other words, the curves show the marginal effect of changing exactly one variable, whereas the model may take advantage of sets of variables changing together.






In contrast to the above marginal response curves, each of the following curves represents a different model, namely, a Maxent model created using only the corresponding variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other variables. They may be easier to interpret if there are strong correlations between variables.






# 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 |
|---|---|---|
| pctwetland | 23 | 7.5 |
| drainage | 21.2 | 15.1 |
| geology | 20.8 | 4.8 |
| strahler | 12.3 | 2.8 |
| temp | 6.4 | 57.4 |
| pctshrub | 5.4 | 1.2 |
| pctdev | 2.1 | 3.3 |
| pctwater | 1.6 | 0.4 |
| pctimperv | 1.5 | 0 |
| velocity | 1.4 | 3.2 |
| sarp | 1.3 | 0.2 |
| precip | 0.9 | 1.5 |
| pctbarren | 0.7 | 0.4 |
| pctpasture | 0.4 | 0.3 |
| flow | 0.3 | 1.6 |
| sinuosity | 0.2 | 0.3 |
| pctcrop | 0.2 | 0 |
| pctforest | 0.1 | 0 |
| slope | 0 | 0 |

# 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
The coefficients of the model
The omission and predicted area for varying cumulative and raw thresholds
The prediction strength at the training and (optionally) test presence sites
Results for all species modeled in the same Maxent run, with summary statistics and (optionally) jackknife results

Regularized training gain is 2.670, training AUC is 0.985, unregularized training gain is 3.220.
Unregularized test gain is 0.986.
Test AUC is 0.892, standard deviation is 0.061 (calculated as in DeLong, DeLong & Clarke-Pearson 1988, equation 2).
Algorithm terminated after 500 iterations (41 seconds).

The follow settings were used during the run:
50 presence records used for training, 12 for testing.
10040 points used to determine the Maxent distribution (background points and presence points).

Environmental layers used: drainage flow geology(categorical) pctbarren pctcrop pctdev pctforest pctimperv pctpasture pctshrub pctwater pctwetland precip sarp sinuosity slope strahler(categorical) temp velocity
Regularization values: linear/quadratic/product: 0.193, categorical: 0.250, threshold: 1.500, hinge: 0.500
Feature types used: linear quadratic hinge
responsecurves: true
pictures: false
outputdirectory: D:\V2Maxent\Output
samplesfile: D:\V2Maxent\OccurrenceData\SpeciesOver24.csv
environmentallayers: D:\V2Maxent\EnvVariables
randomtestpoints: 20
Command line used:

Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E Lepomis_marginatus responsecurves nopictures outputdirectory=D:\V2Maxent\Output samplesfile=D:\V2Maxent\OccurrenceData\SpeciesOver24.csv environmentallayers=D:\V2Maxent\EnvVariables randomtestpoints=20 -t geology -t strahler