# Replicated maxent model for Eurasian_House_Sparrows

This page summarizes the results of 10 split-sample models for Eurasian_House_Sparrows, created Fri Oct 08 12:51:58 MDT 2010 using Maxent version 3.3.3a. The individual models are here: [0] [1] [2] [3] [4] [5] [6] [7] [8] [9]


Analysis of omission/commission

The following picture shows the test omission rate and predicted area as a function of the cumulative threshold, averaged over the replicate runs. The omission rate should be close to the predicted omission, because of the definition of the cumulative threshold.
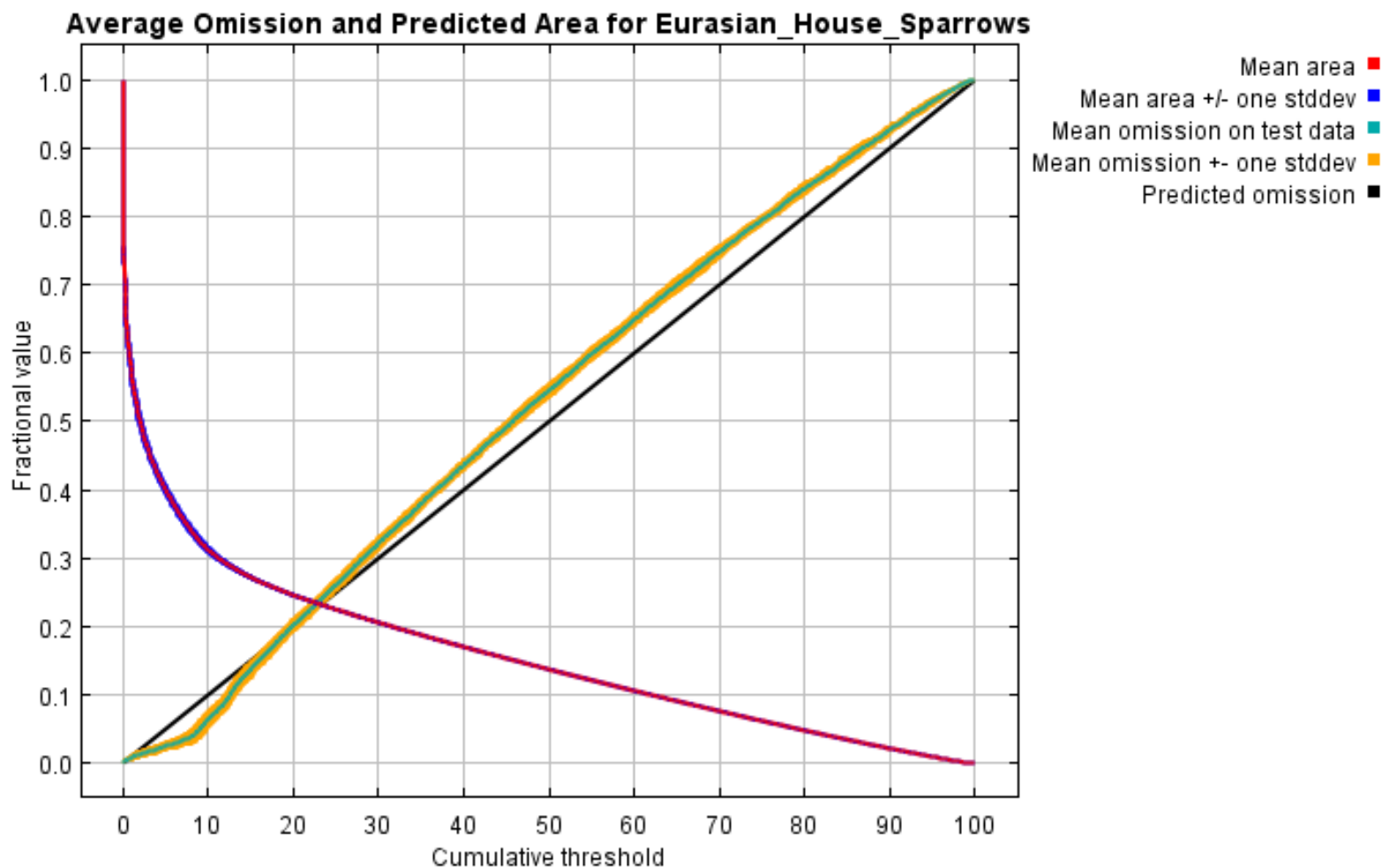

The next picture is the receiver operating characteristic (ROC) curve for the same data, again averaged over the replicate runs. Note that the specificity is defined using predicted area, rather than true commission (see the paper by Phillips, Anderson and Schapire cited on the help page for discussion of what this means). The average test AUC for the replicate runs is 0.838, and the standard deviation is 0.002.
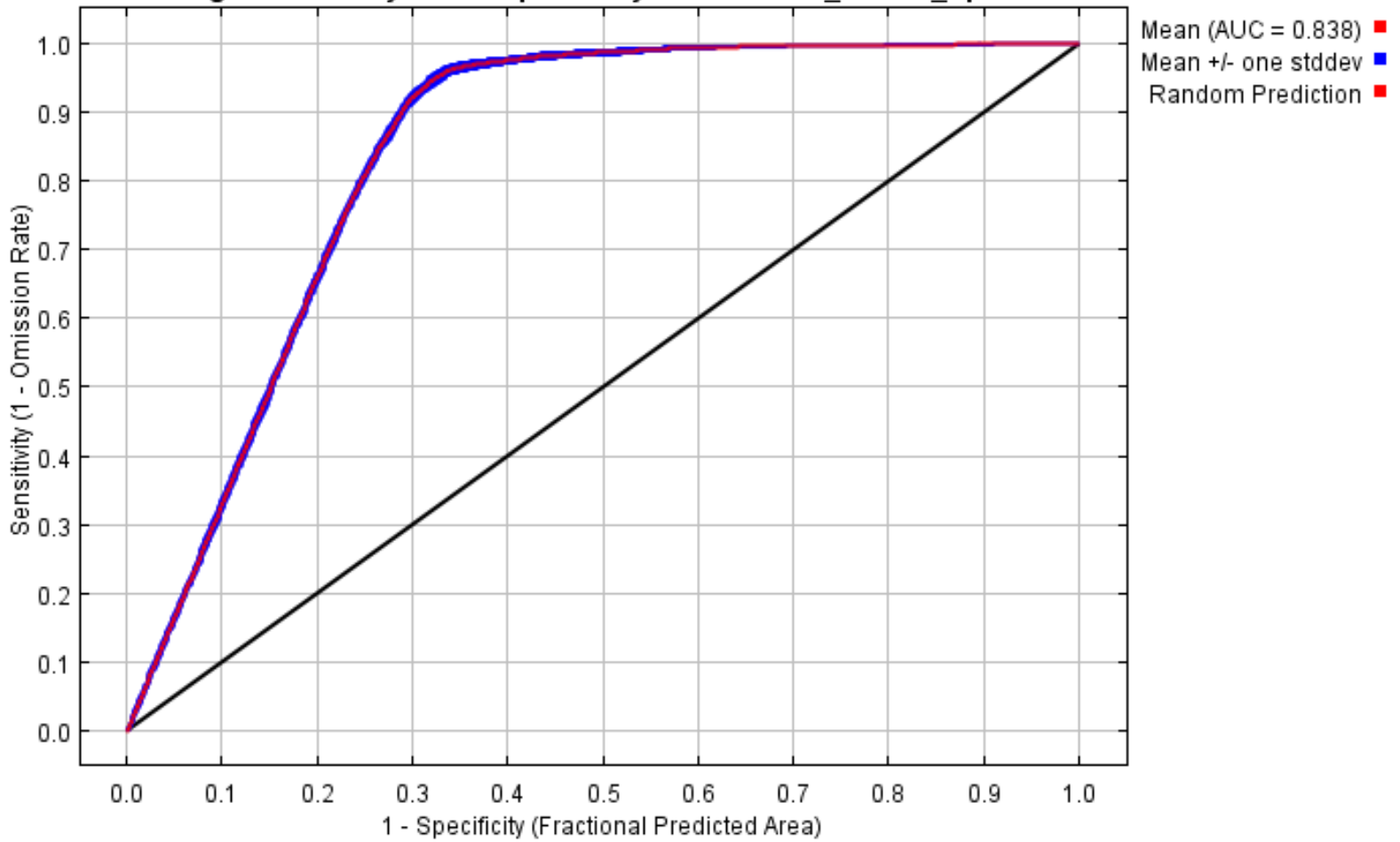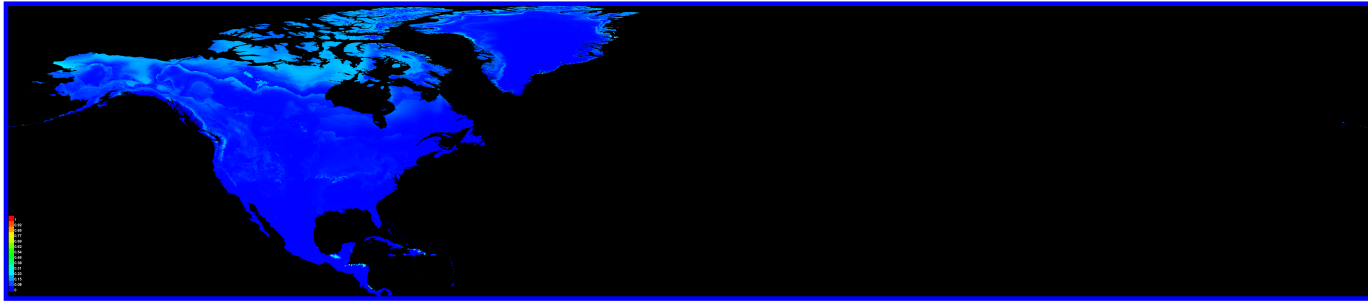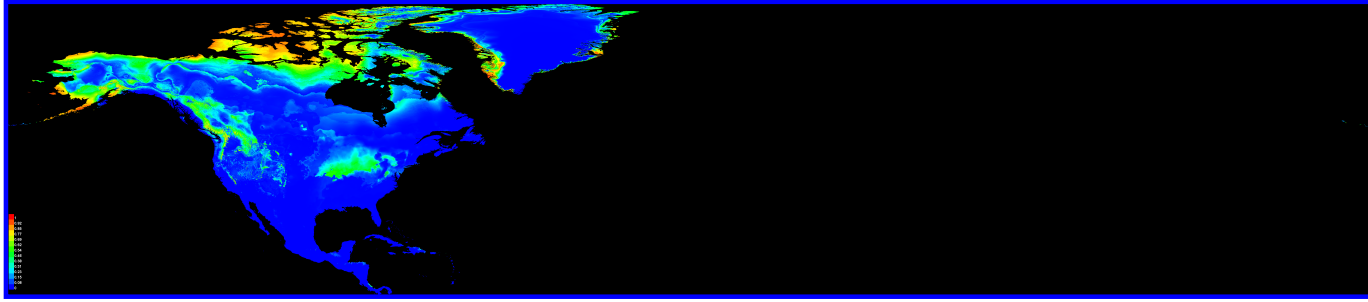
Pictures of the model

The following two pictures show the point-wise mean and standard deviation of the 10 models applied to the environmental layers in maxent.cache. Other available summary grids are min, max and median.
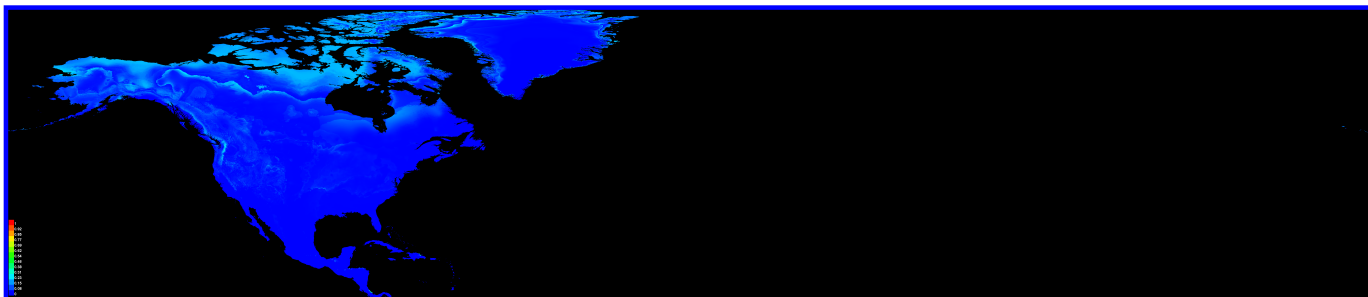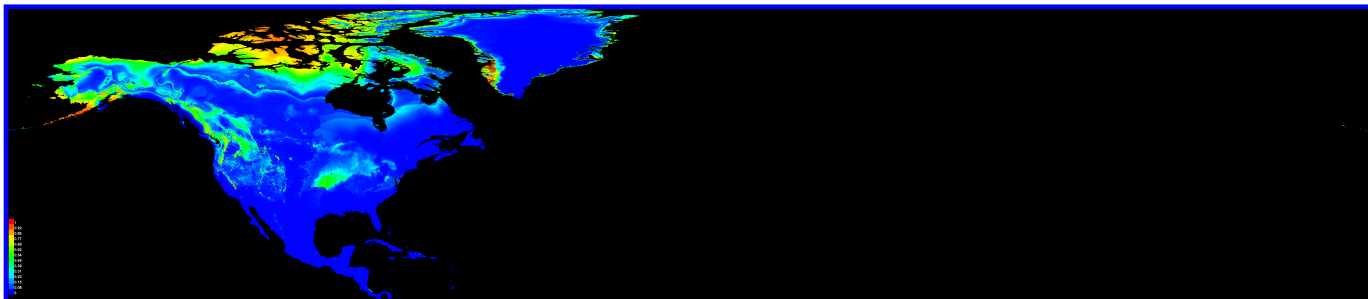
The following two pictures show the point-wise mean and standard deviation of the 10 models applied to the environmental layers in b2a_2020_avg. Other available summary grids are min, max and median.



The following two pictures show the point-wise mean and standard deviation of the 10 models applied to the environmental layers in b2a_2050_avg. Other available summary grids are min, max and median.
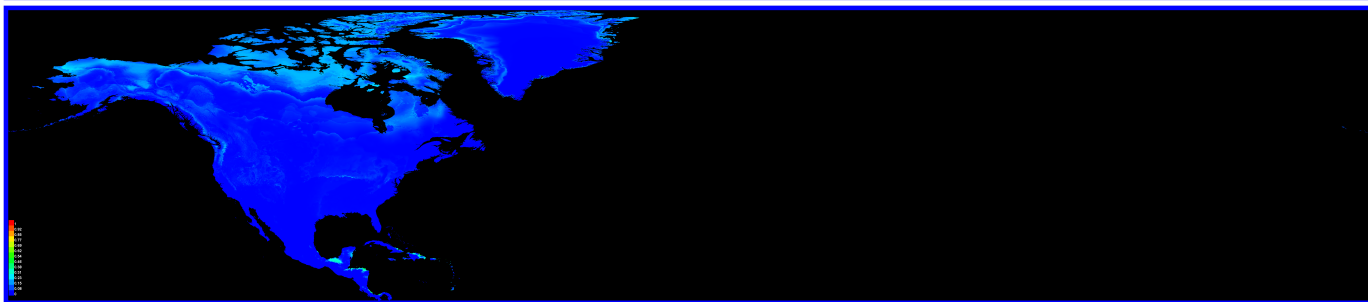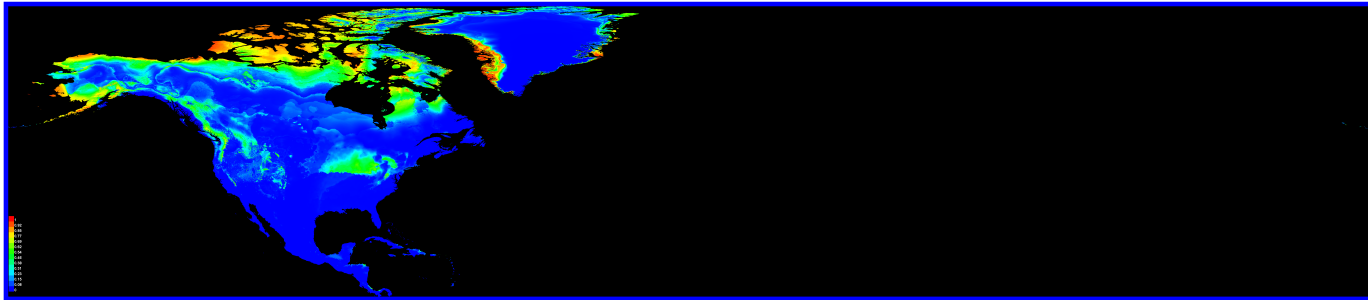

Response curves

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. The curves show the mean response of the 10 replicate Maxent runs (red) and and the mean +/- one standard deviation (blue, two

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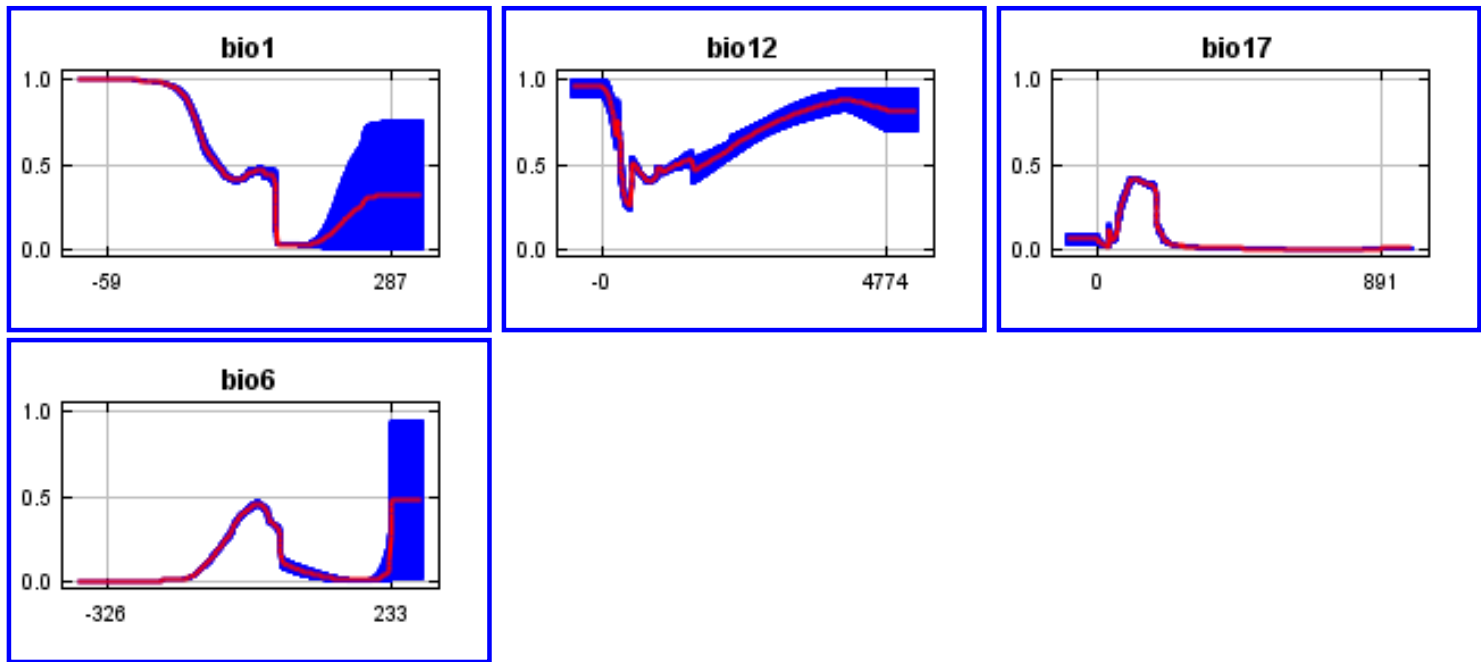
In contrast to the above marginal response curves, each of the following curves represents a different model, namely, a Maxent model created using only the corresponding variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by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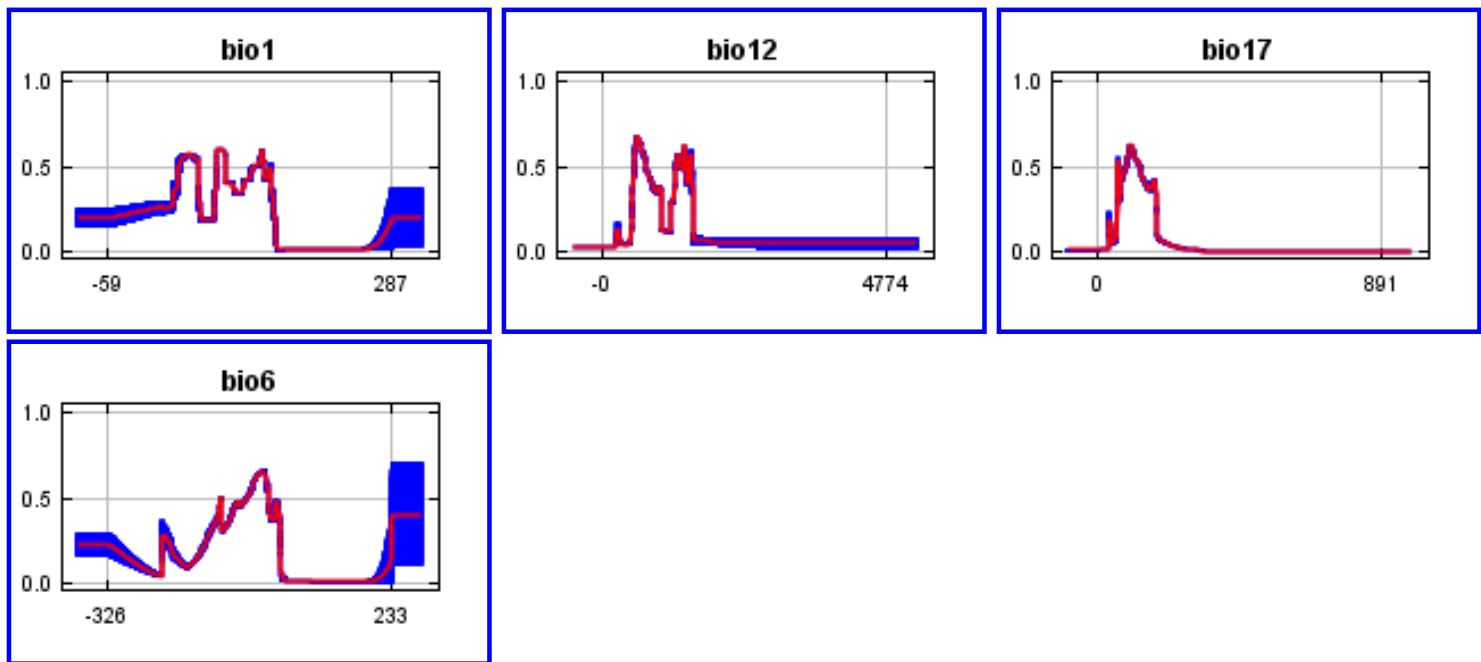

Analysis of variable contributions

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. Values shown are averages over replicate runs.

| Variable	Percent contribution	Permutation importance
bio17	57.7	42.6
bio6	18.5	29.7
bio1	17.7	19.5
bio12	6.1	8.1

The following picture shows the results of the jackknife test of variable importance. The environmental variable with highest gain when used in isolation is bio17, which therefore appears to have the most useful information by itself. The environmental variable that decreases the gain the most when it is omitted is bio17, which therefore appears to have the most information that isn't present in the other variables. Values shown are averages over replicate runs.
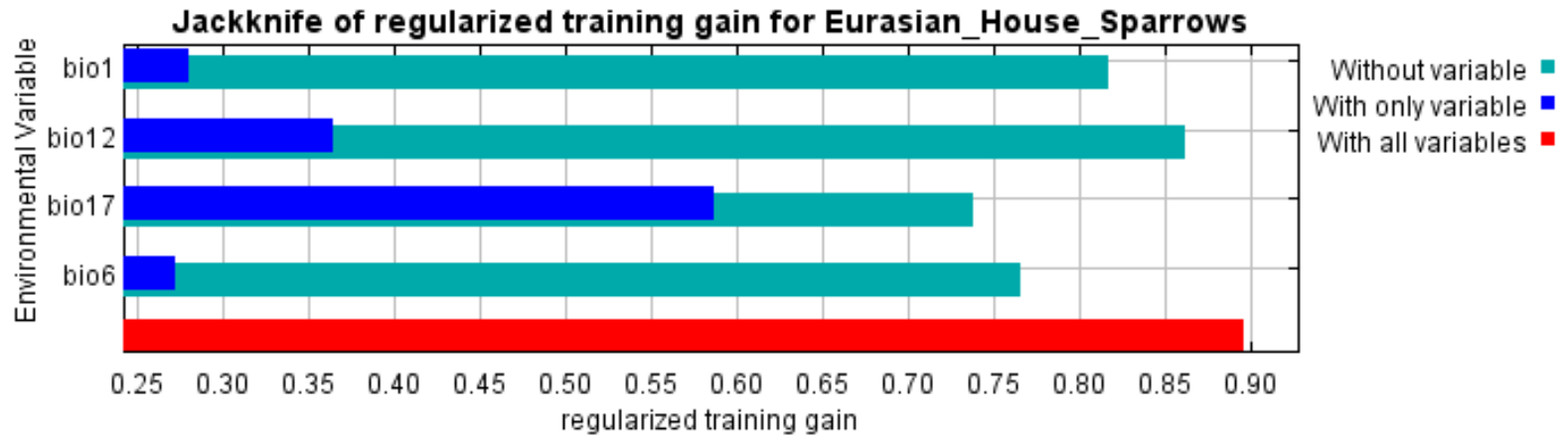
The next picture shows the same jackknife test, using test gain instead of training gain. Note that conclusions about which variables are most important can change, now that we're looking at test data.
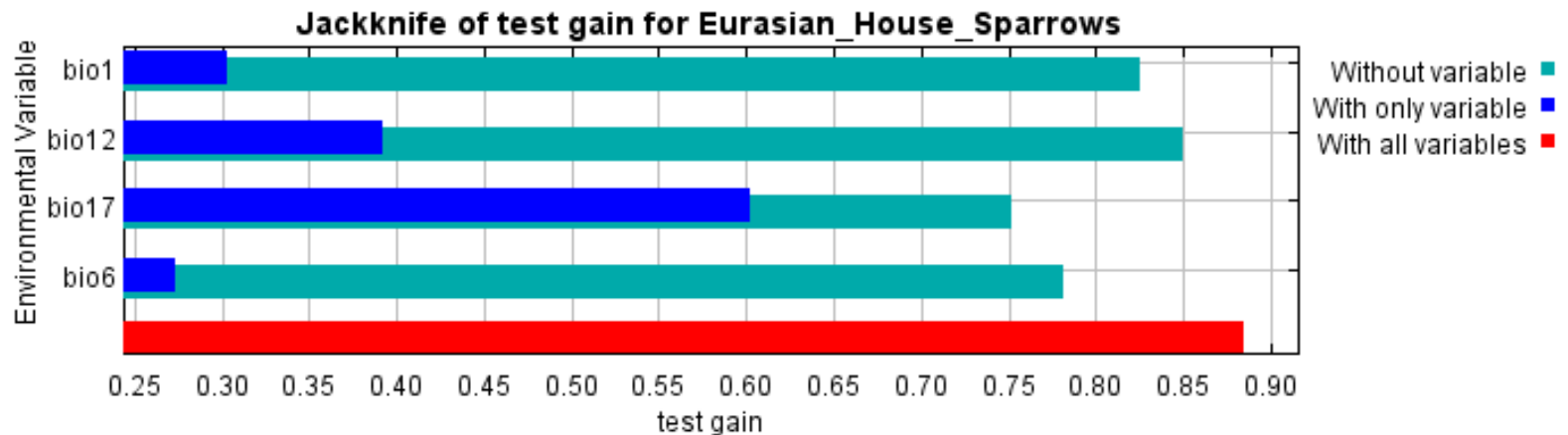
Lastly, we have the same jackknife test, using AUC on test data.

Jackknife of AUC for Eurasian_House_Sparrows
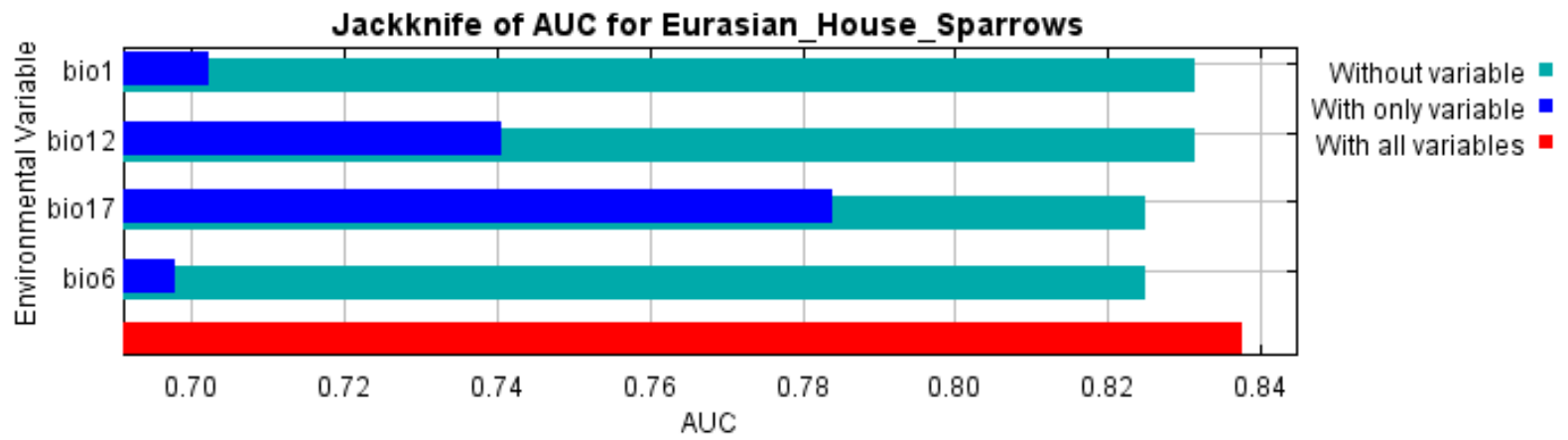
Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E Eurasian_House_Sparrows responsecurves jackknife outputdirectory=I:\SpeciesData\HouseSparrows\Maxent \ALL_ETS2 projectionlayers=I:\SpeciesData\HouseSparrows\1950to2000_NorthAmerica\ASCII\maxent.cache,I: \SpeciesData\HouseSparrows\Future_mean\NorthAmerica\B2a\2020\ASCII\b2a_2020_avg,I:\SpeciesData \HouseSparrows\Future_mean\NorthAmerica\B2a\2050\ASCII\b2a_2050_avg samplesfile=I:\SpeciesData \HouseSparrows\ETSdata\EHS_0_2010_30sec_nozero_HS30minbckgrnd.csv environmentallayers=I:\SpeciesData \HouseSparrows\ETSdata\EHS_0_2010_30sec_nozero_HS30minbckgrnd.csv randomseed skipifexists nowritemess randomtestpoints=30 replicates=10 replicatetype=subsample appendtoresultsfile maximumiterations=5000 -N bio10 -N bio11 -N bio13 -N bio14 -N bio15 -N bio16 -N bio18 -N bio19 -N bio2 -N bio3 -N bio4 -N bio5 -N bio7 -N bio8 -N bio9