Assessment of habitat suitability for the globally threatened Cock-tailed Tyrant,

***Alectrurus tricolor* (Passeriformes: Tyrannidae)**

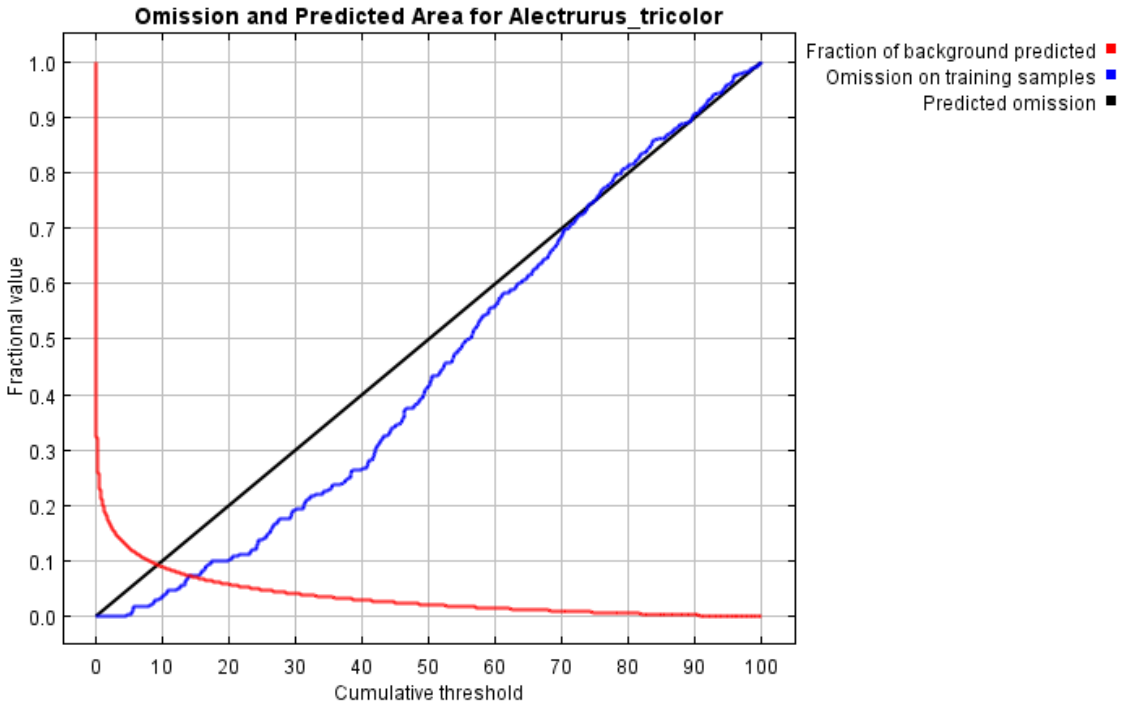
Pier Cacciali, Lorena Sforza

Appendix S1

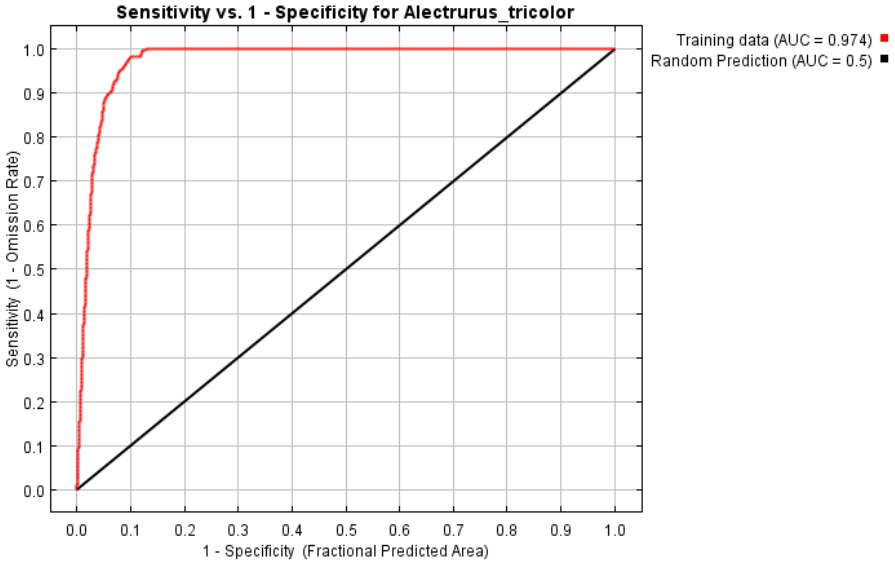
Maxent model for *Alectrurus tricolor*

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 0.963 rather than 1; in practice the test AUC may exceed this bound.

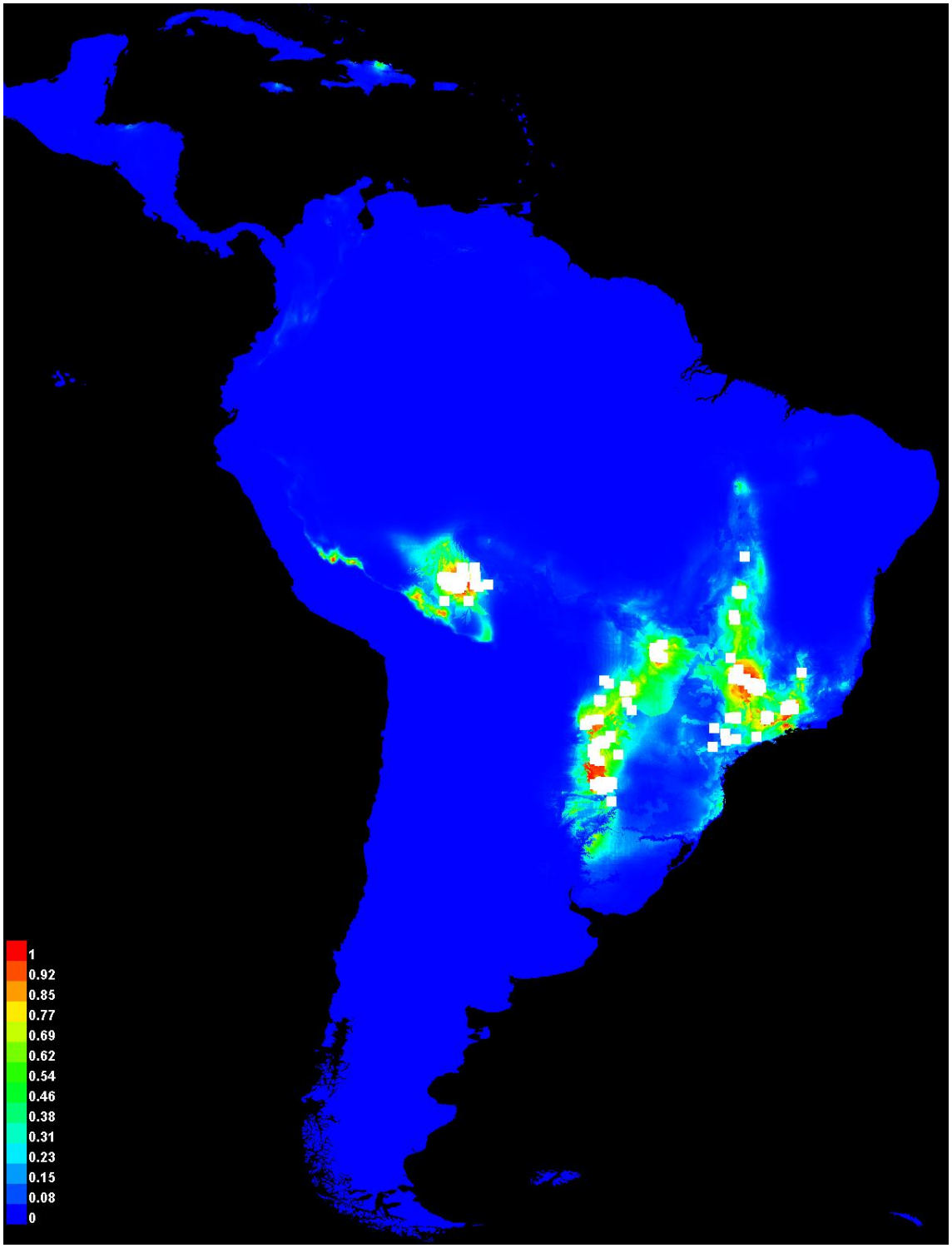


Some common thresholds and corresponding omission rates are as follows. If test data are available, binomial probabilities are calculated exactly if the number of test samples is at most 25, otherwise using a normal approximation to the binomial. These are 1-sided p-values for the null hypothesis that 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	0.020	Fixed cumulative value 1	0.204	0.000
5.000	0.094	Fixed cumulative value 5	0.124	0.004
10.000	0.184	Fixed cumulative value 10	0.090	0.039
4.535	0.087	Minimum training presence	0.129	0.000
19.957	0.366	10 percentile training presence	0.059	0.100
14.143	0.266	Equal training sensitivity and specificity	0.074	0.074
8.025	0.147	Maximum training sensitivity plus specificity	0.101	0.017
2.011	0.039	Balance training omission, predicted area and threshold value	0.170	0.000
8.447	0.153	Equate entropy of thresholded and original distributions	0.098	0.022

Pictures of the model

This is a representation of the Maxent model for Alectrurus_tricolor. Warmer colors show areas with better predicted conditions. White dots show the presence locations used for training, while violet dots show test locations. Click on the image for a full-size version.



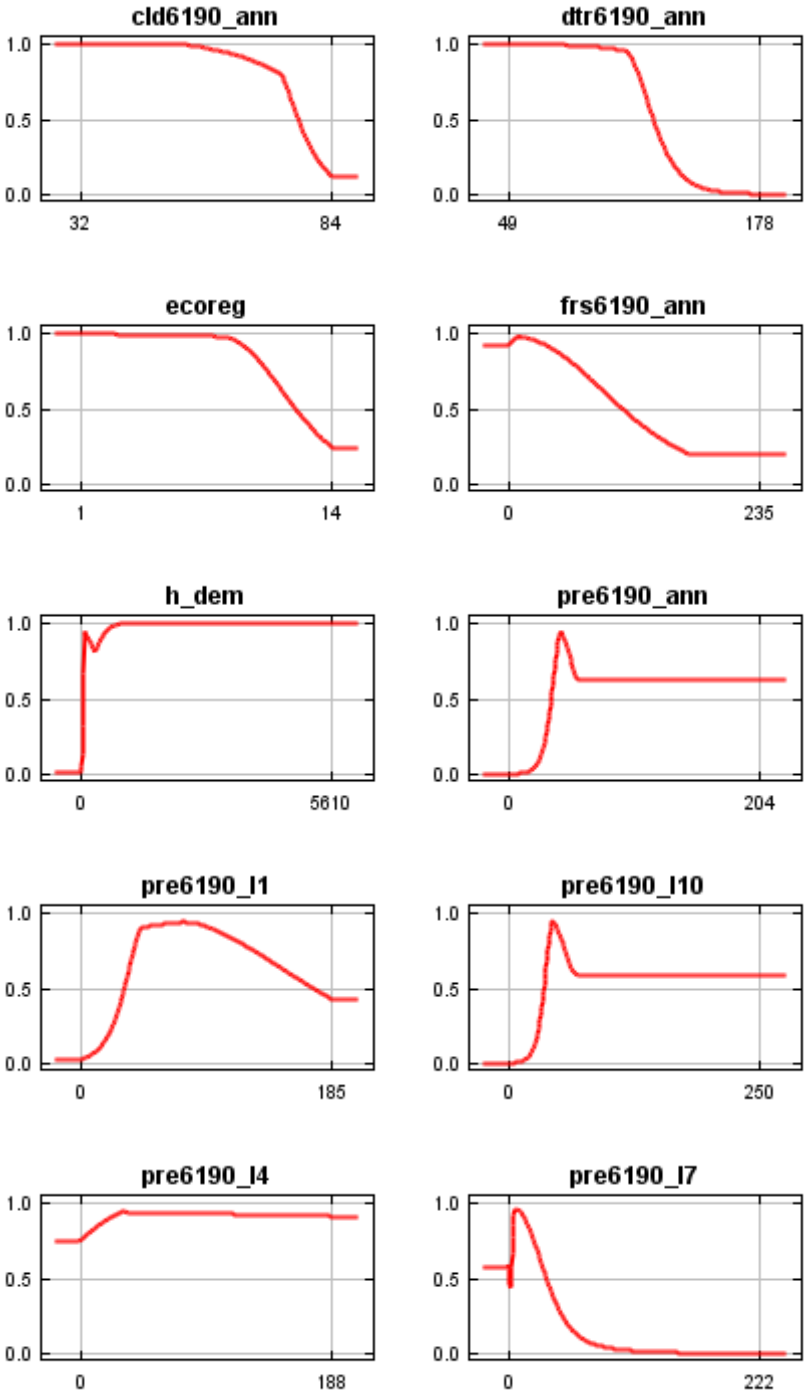
Analysis of variable contributions

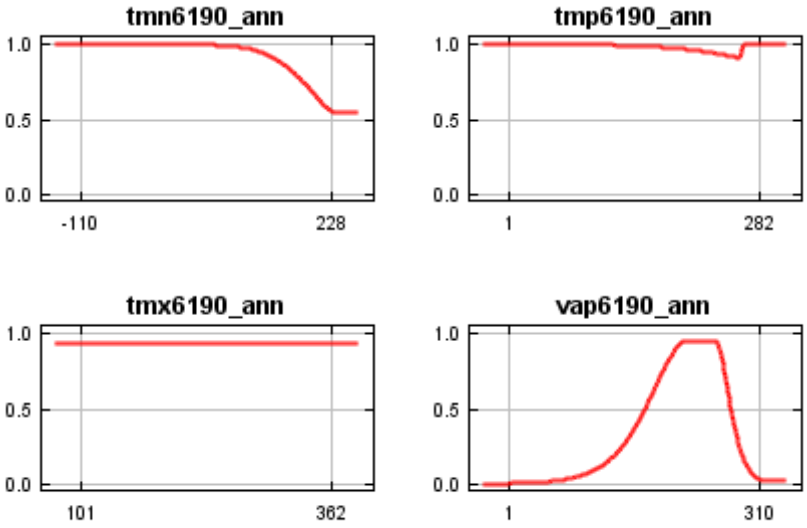
The following table gives estimates of relative contributions of the environmental variables to the Maxent model. To determine the first estimate, in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the corresponding variable, or subtracted from 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
ecoreg	33.5	0.7
pre6190_l7	15.4	11.6
pre6190_l10	15.1	18.8
pre6190_l1	14.1	5.1
vap6190_ann	7.8	9.3
h_dem	3.7	9.3
dtr6190_ann	3.4	11.5
cld6190_ann	3.2	6.2
pre6190_ann	2.1	19.8
tmp6190_ann	1	2.5
frs6190_ann	0.7	0.9
tmn6190_ann	0.1	4.1
pre6190_l4	0	0
tmx6190_ann	0	0

Response curves

These curves show how each environmental variable affects the Maxent prediction. The 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. 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.


cld6190_ann
1.0
0.5
0.0
32
84
dtr6190_ann
1.0
0.5
0.0
49
178
ecoreg
1.0
0.5
0.0
1
14
frs6190_ann
1.0
0.5
0.0
0
235
h_dem
1.0
0.5
0.0
0
5610
pre6190_ann
1.0
0.5
0.0
0
204
pre6190_l1
1.0
0.5
0.0
0
185
pre6190_l10
1.0
0.5
0.0
0
250
pre6190_l4
1.0
0.5
0.0
0
188
pre6190_l7
1.0
0.5
0.0
0
222


tmn6190_ann
1.0
0.5
0.0
-110
228
tmp6190_ann
1.0
0.5
0.0
1
282
tmx6190_ann
1.0
0.5
0.0
101
362
vap6190_ann
1.0
0.5
0.0
1
310