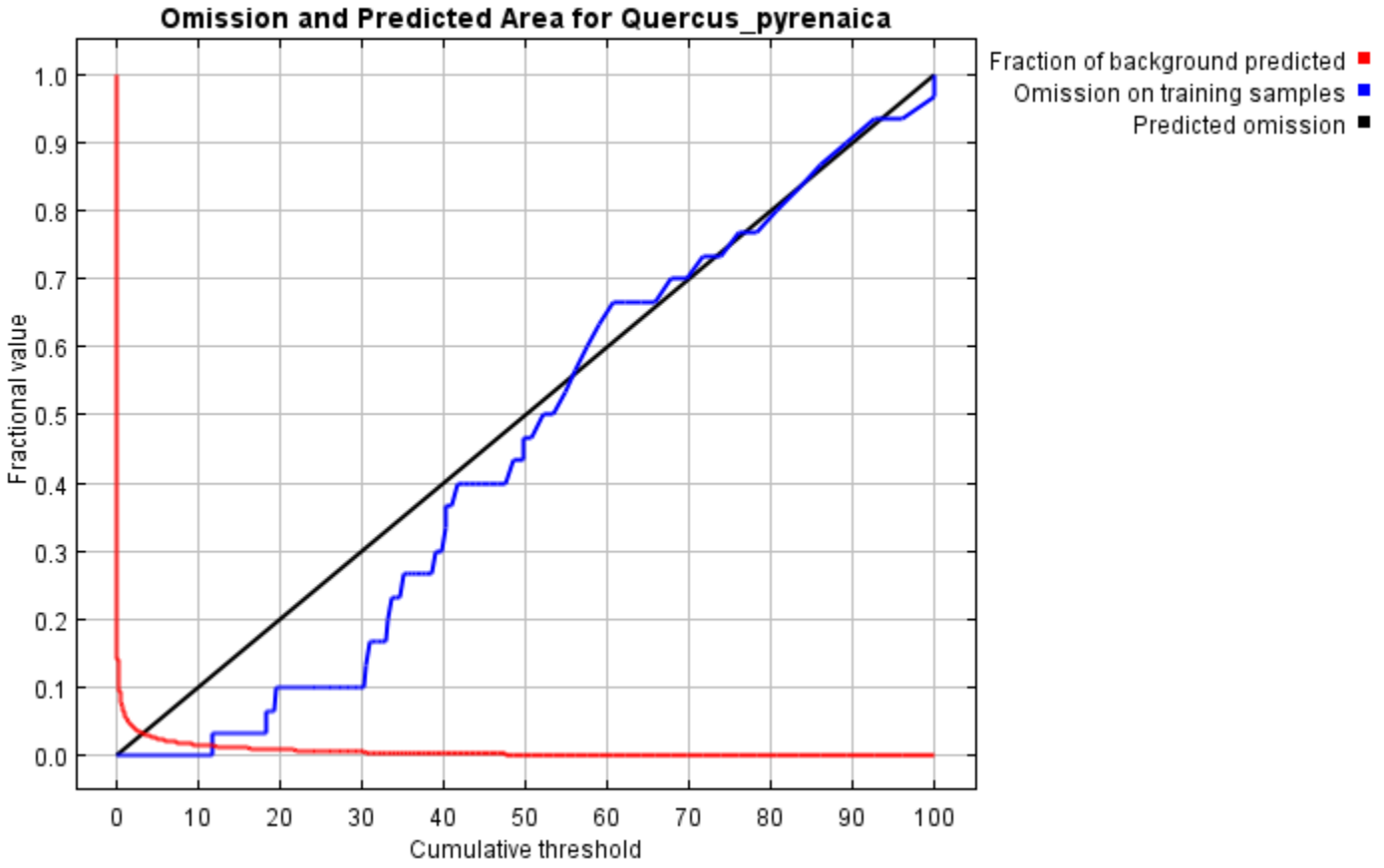
Maxent model for Quercus_pyrenaica

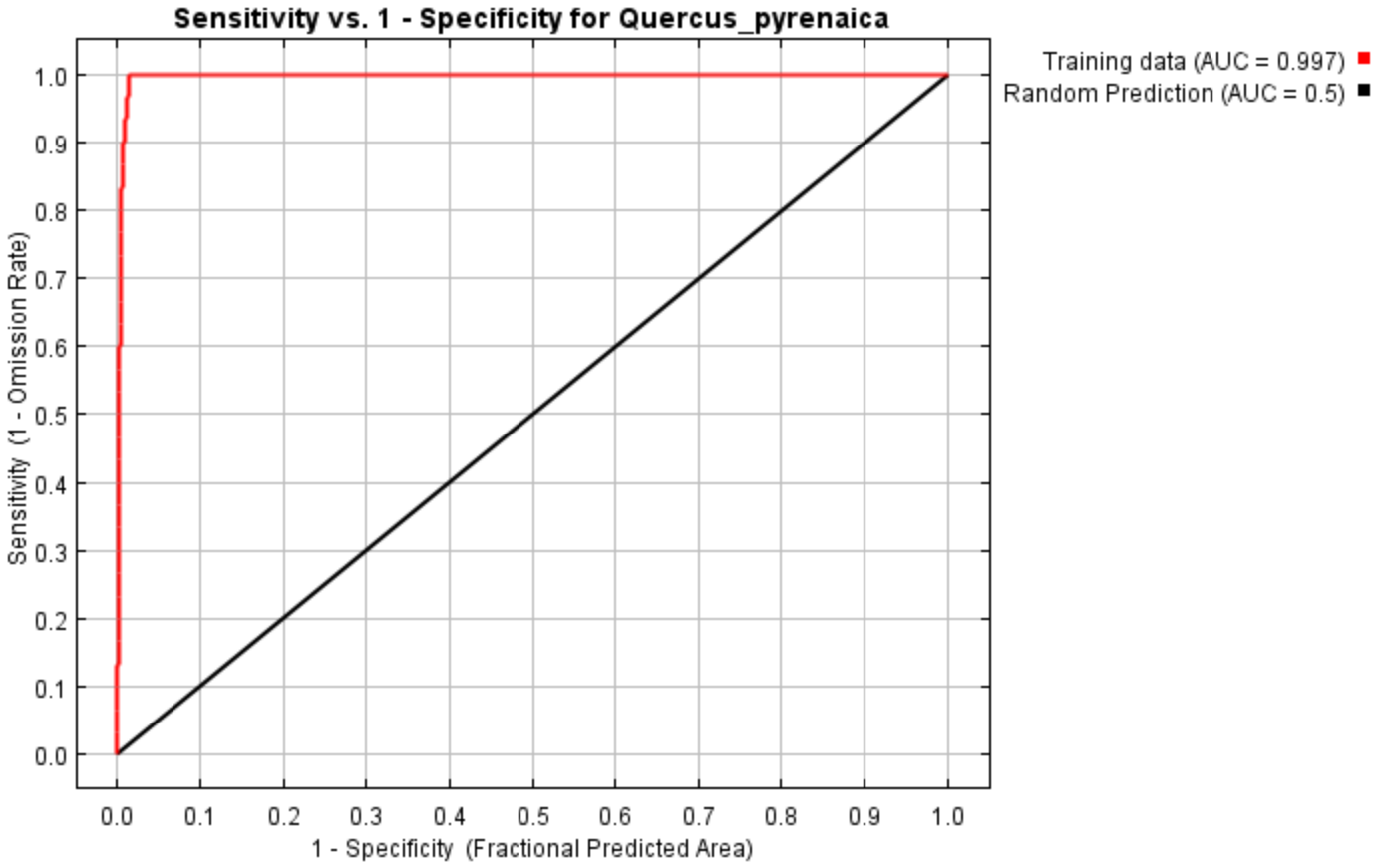
This page contains some analysis of the Maxent model for Quercus_pyrenaica, created Sat Feb 07 18:45:40 CET 2015 using Maxent version 3.3.3k. If you would like to do further analyses, the raw data used here is linked to at the end of this page.

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



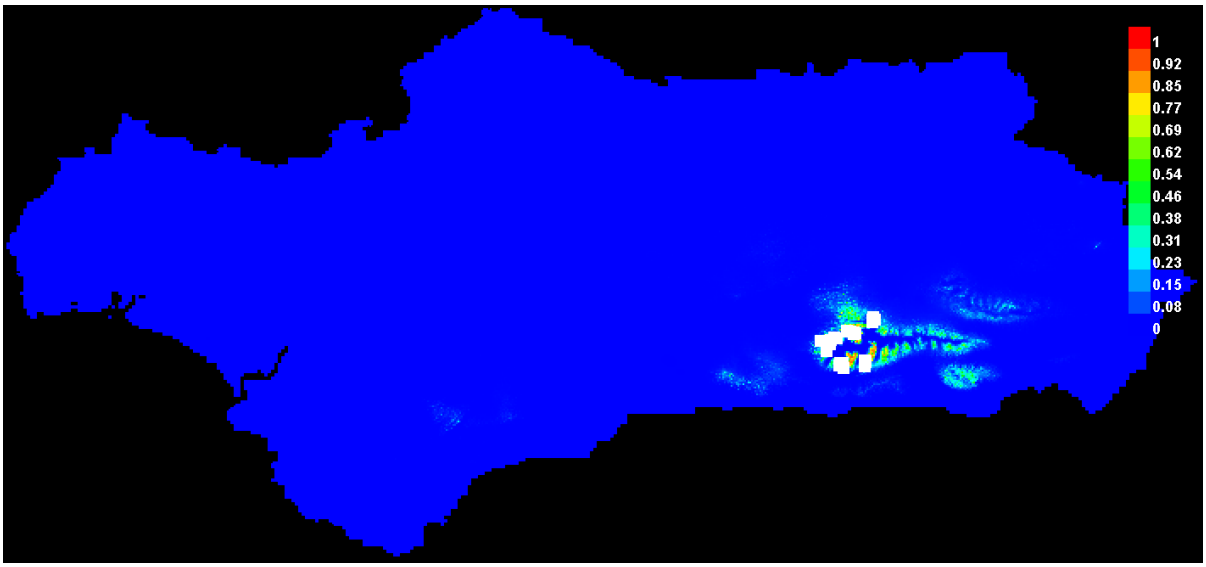
Some common thresholds and corresponding omission rates are as follows. If test data are available, binomial probabilities are calculated exactly if the number of test samples is at most 25, otherwise using a normal approximation to the binomial. These are 1-sided p-values for the null hypothesis that 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
1.000	0.005	Fixed cumulative value 1	0.060	0.000
5.000	0.045	Fixed cumulative value 5	0.026	0.000
10.000	0.117	Fixed cumulative value 10	0.016	0.000
11.688	0.137	Minimum training presence	0.015	0.000
30.163	0.342	10 percentile training presence	0.006	0.100
11.688	0.137	Equal training sensitivity and specificity	0.015	0.000
11.688	0.137	Maximum training sensitivity plus specificity	0.015	0.000

1.105	0.006	Balance training omission, predicted area and threshold value	0.058	0.000
11.688	0.137	Equate entropy of thresholded and original distributions	0.015	0.000

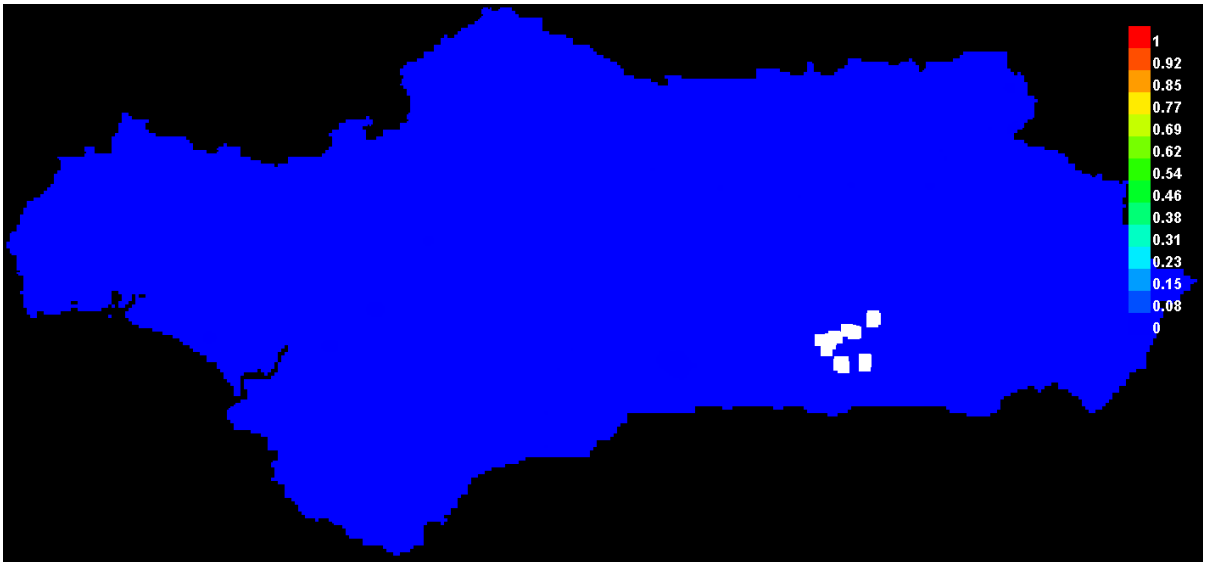
Pictures of the model

This is a representation of the Maxent model for Quercus_pyrenaica. 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.



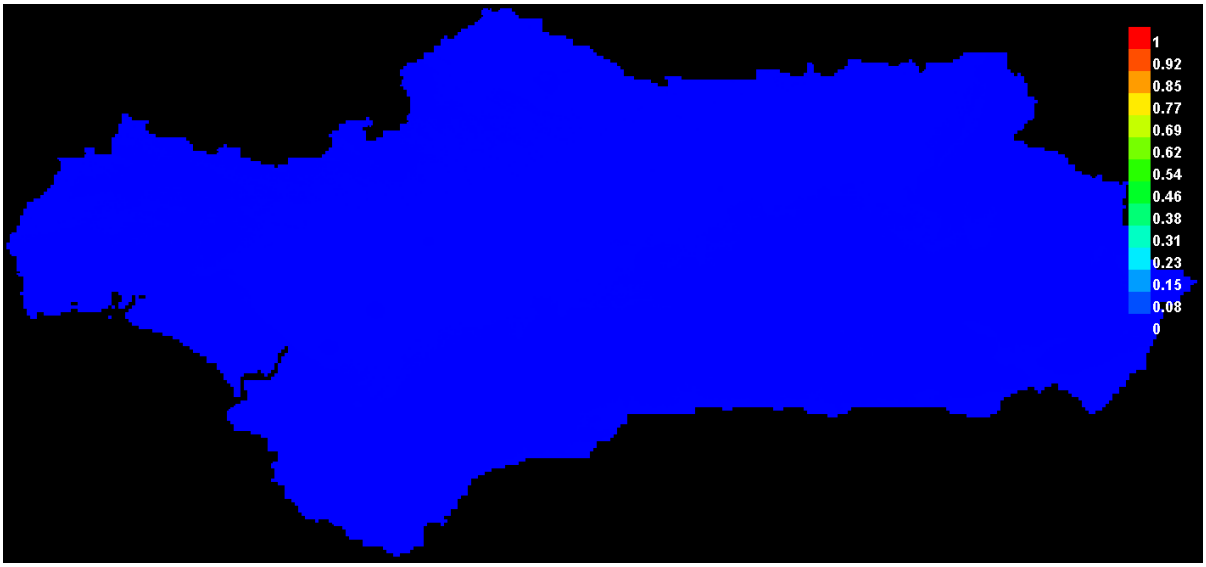
Click here to interactively explore this prediction using the Explain tool. If clicking from your browser does not succeed in starting the tool, try running the script in C:\Users\Helena\Documents\Granada\máster \asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\resultados \Quercus_pyrenaica_explain.bat directly. This tool requires the environmental grids to be small enough that they all fit in memory.

This is the projection of the Maxent model for Quercus_pyrenaica onto the environmental variables in C:\Users\Helena\Documents\Granada\máster\asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones \sesión 6\ficheros\variables\futuro_B2\2050. 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.



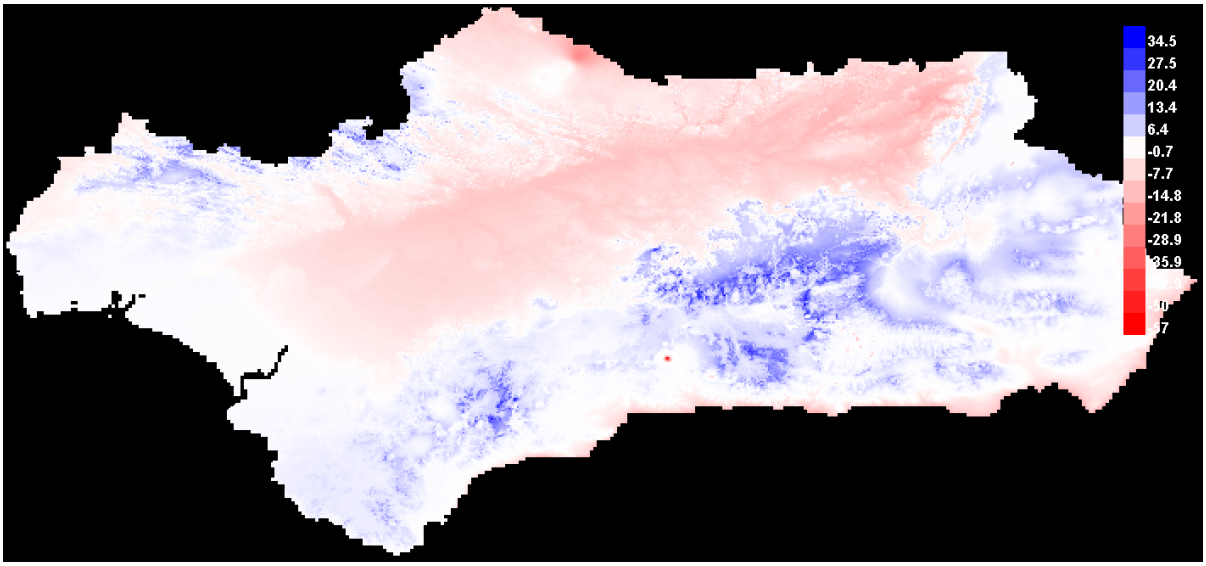
Click here to interactively explore this prediction using the Explain tool. If clicking from your browser does not succeed in starting the tool, try running the script in C:\Users\Helena\Documents\Granada\máster \asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\resultados \Quercus_pyrenaica_2050_explain.bat directly. This tool requires the environmental grids to be small enough that they all fit in memory.

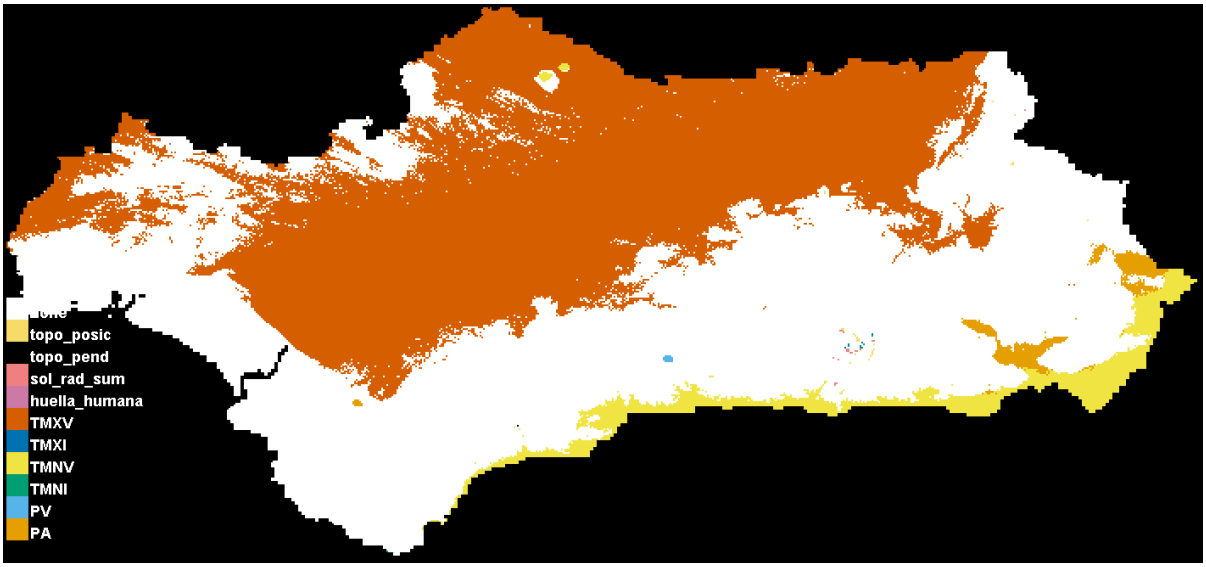
The following picture shows where the prediction is most affected by variables being outside their training range, while projecting the Maxent model onto the environmental variables in C:\Users\Helena\Documents \Granada\máster\asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\variables \futuro_B2\2050. The values shown in the picture give the absolute difference in predictions when using vs not using clamping. (Clamping means that environmental variables and features are restricted to the range of values encountered during training.) Warmer colors show areas where the treatment of variable values outside their training ranges is likely to have a large effect on predicted suitability.



The following two pictures compare the environmental similarity of variables in 2050 to the environmental data used for training the model. In the first picture (MESS), areas in red have one or more environmental variables outside the range present in the training data, so predictions in those areas should be treated with

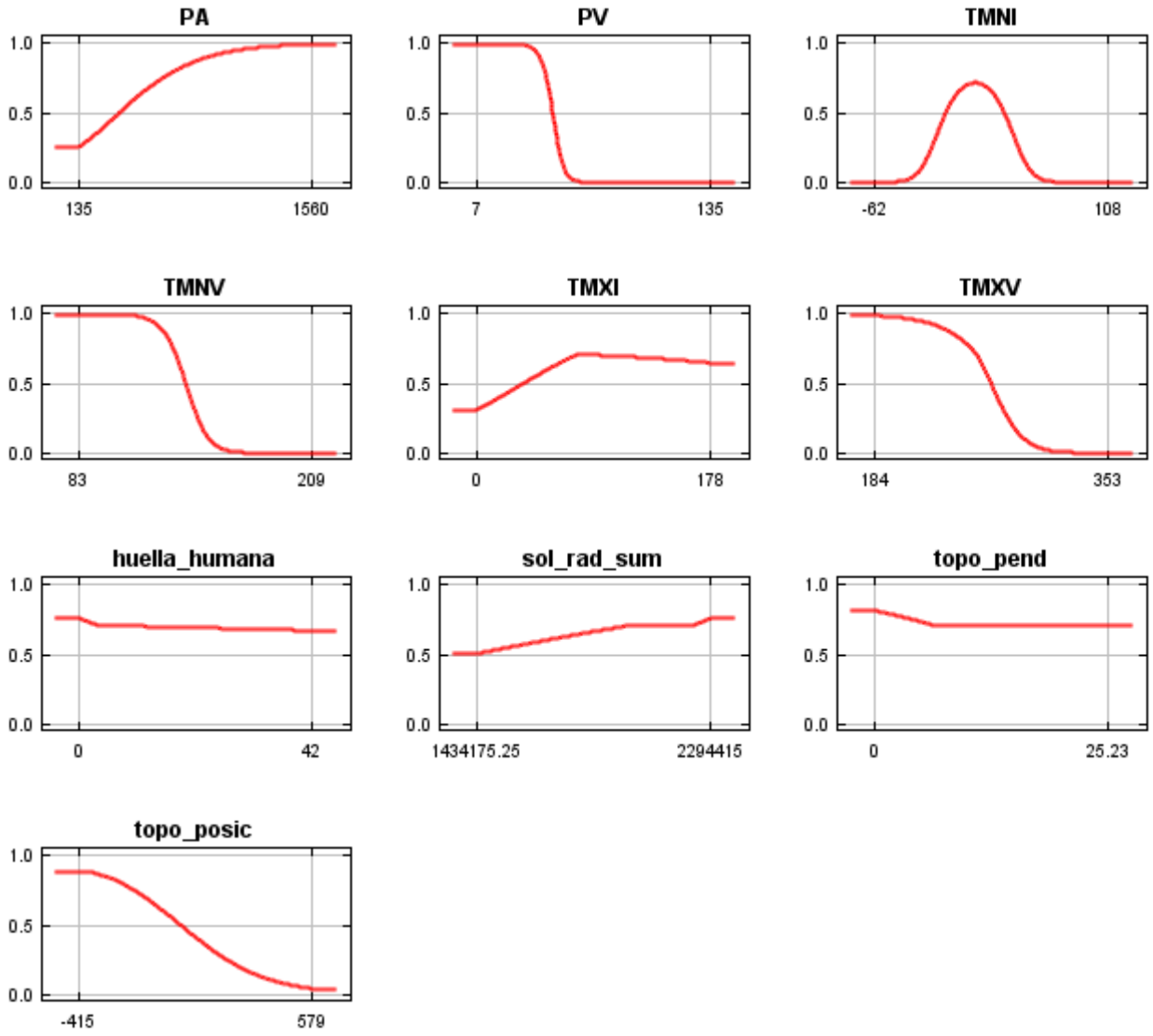
strong caution. The second picture (MoD) shows the most dissimilar variable, i.e., the one that is furthest outside its training range. For details, see Elith et al., Methods in Ecology and Evolution, 2010



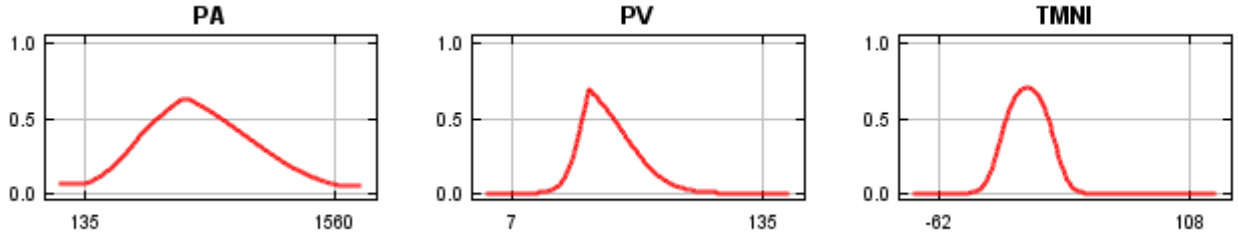


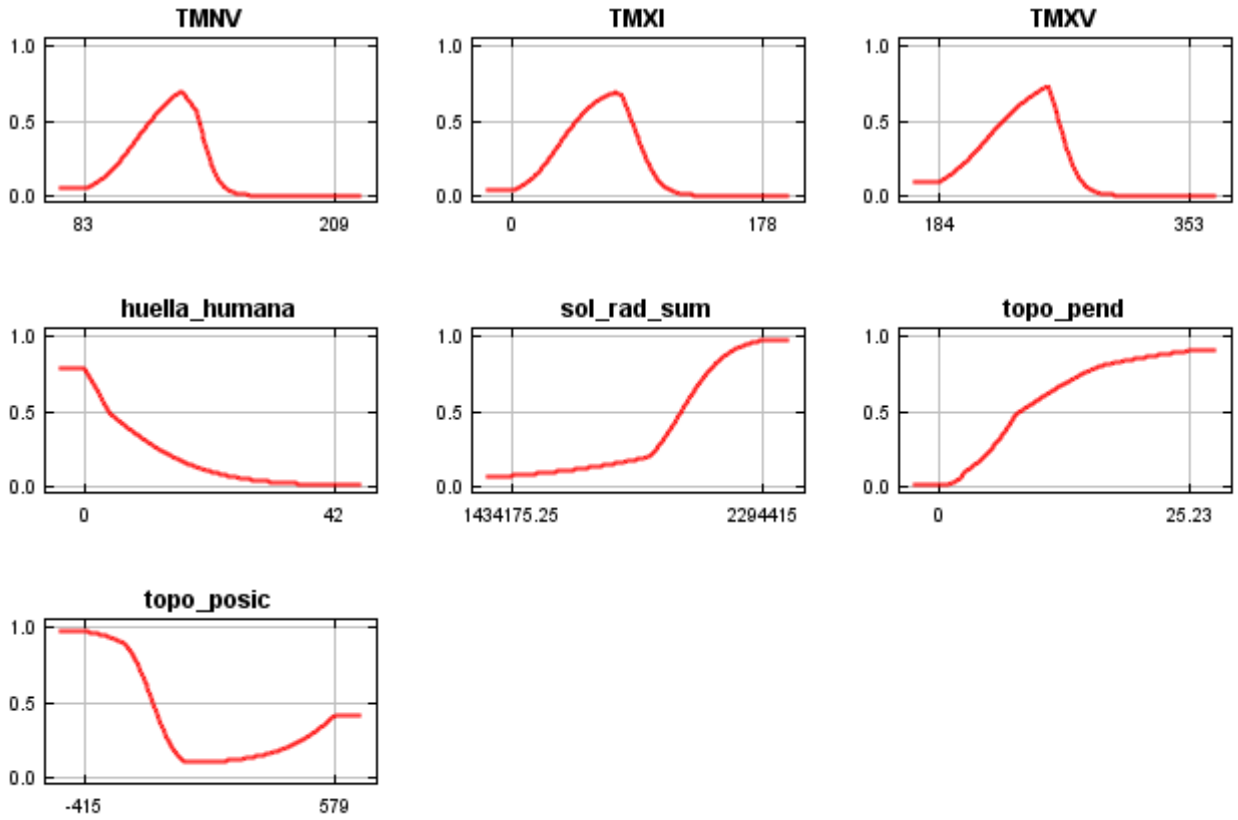
Response curves

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
TMXV	46.5	28.5
topo_posic	25.3	1.6
TMXI	15.3	0.1
TMNV	5.3	55.8
PV	3.6	7.5
TMNI	2.8	5.6

sol_rad_sum	0.5	0.1
PA	0.4	0.6
topo_pend	0.2	0.1
huella_humana	0.2	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 model applied to the environmental layers in C:\Users\Helena\Documents\Granada\máster\asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\variables\futuro_B2\2050
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 4.204, training AUC is 0.997, unregularized training gain is 4.589.
Algorithm terminated after 500 iterations (9 seconds).

The follow settings were used during the run:
30 presence records used for training.
10027 points used to determine the Maxent distribution (background points and presence points).
Environmental layers used (all continuous): PA PV TMNI TMNV TMXI TMXV huella_humana sol_rad_sum topo_pend topo_posic
Regularization values: linear/quadratic/product: 0.250, categorical: 0.250, threshold: 1.700, hinge: 0.500
Feature types used: hinge linear quadratic
responsecurves: true
outputdirectory: C:\Users\Helena\Documents\Granada\máster\asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\resultados
projectionlayers: C:\Users\Helena\Documents\Granada\máster\asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\variables\futuro_B2\2050
samplesfile: C:\Users\Helena\Documents\Granada\máster\asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\presencia\presencia_q_pyrenaica.csv
environmentallayers: C:\Users\Helena\Documents\Granada\máster\asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\variables\presente

Command line used:

Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E Quercus_pyrenaica responsecurves "outputdirectory=C:\Users\Helena\Documents\Granada\máster \asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\resultados" "projectionlayers=C:\Users\Helena\Documents\Granada\máster\asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\variables\futuro_B2\2050" "samplesfile=C:\Users \Helena\Documents\Granada\máster\asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\presencia\presencia_q_pyrenaica.csv" "environmentallayers=C:\Users\Helena\Documents \Granada\máster\asignaturas master\En curso\Ecoinformatica\CLASE\Sesiones\sesión 6\ficheros\variables \presente"