

Habitat suitability modeling test

2023-09-05

Mammalia

Functional group 11 ; Number of species: 2

Species name	Order
Rupicapra rupicapra	Cetartiodactyla
Talpa caeca	Eulipotyphla

In total, 4350 pixels of 1km² were occupied by the group within metropolitan France and 5000 pseudo-absences were randomly generated over the same territory (3 times, see figures below).

The final habitat suitability maps were generated using 4 families of machine learning algorithms (SDM models):

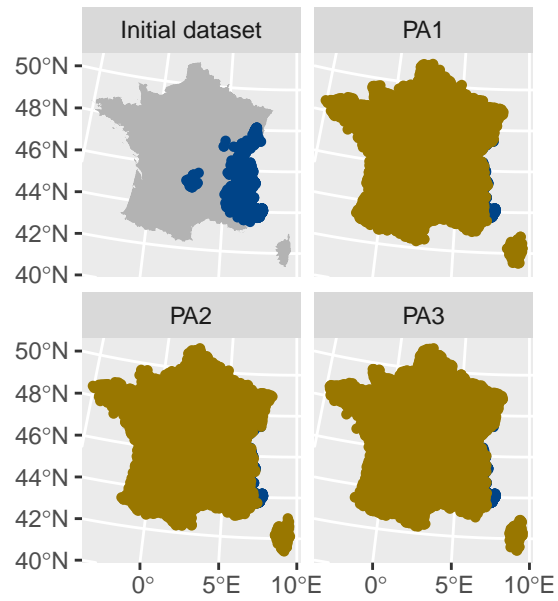
- neural network: artificial neural network (ANN)
- bagging model: random forest (RF)
- boosting model: Extreme gradient boosting (XGBOOST)
- maximum entropy model: MAXNET

For each family, 15 models were generated using different presence/pseudo-absence composition within the calibration/validation dataset. Each model was calibrated using 80% of the initial dataset and the remaining 20% was used to produce validation indicators (ROC and TSS). All models that passed the validation threshold (i.e., TSS > 0.4) were combined to produce ensemble models (either using probability weighted mean or committee averaging). Finally, we produced a map of presence probability for the group from each ensemble model.

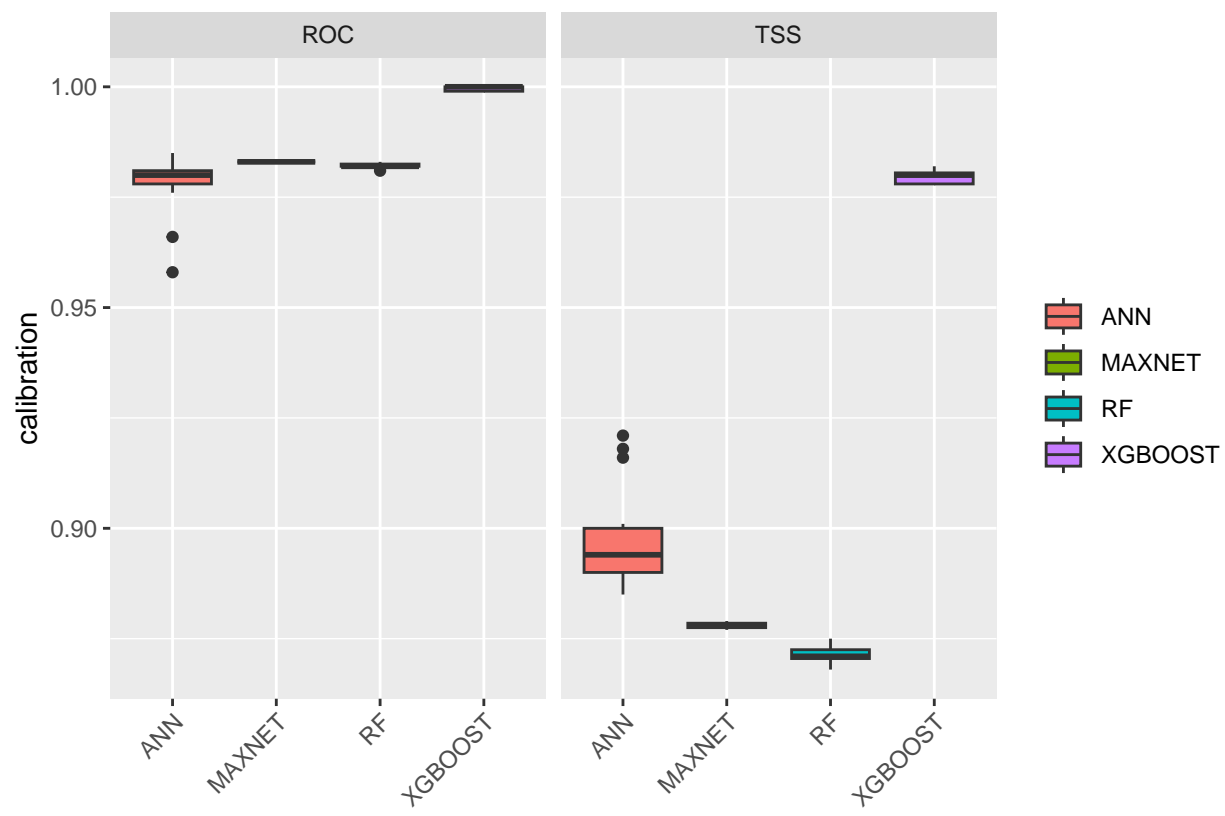
Explanatory variables that were used in these models included the same variables as the one used for environmental PCA (work task 1 for functional group definition).

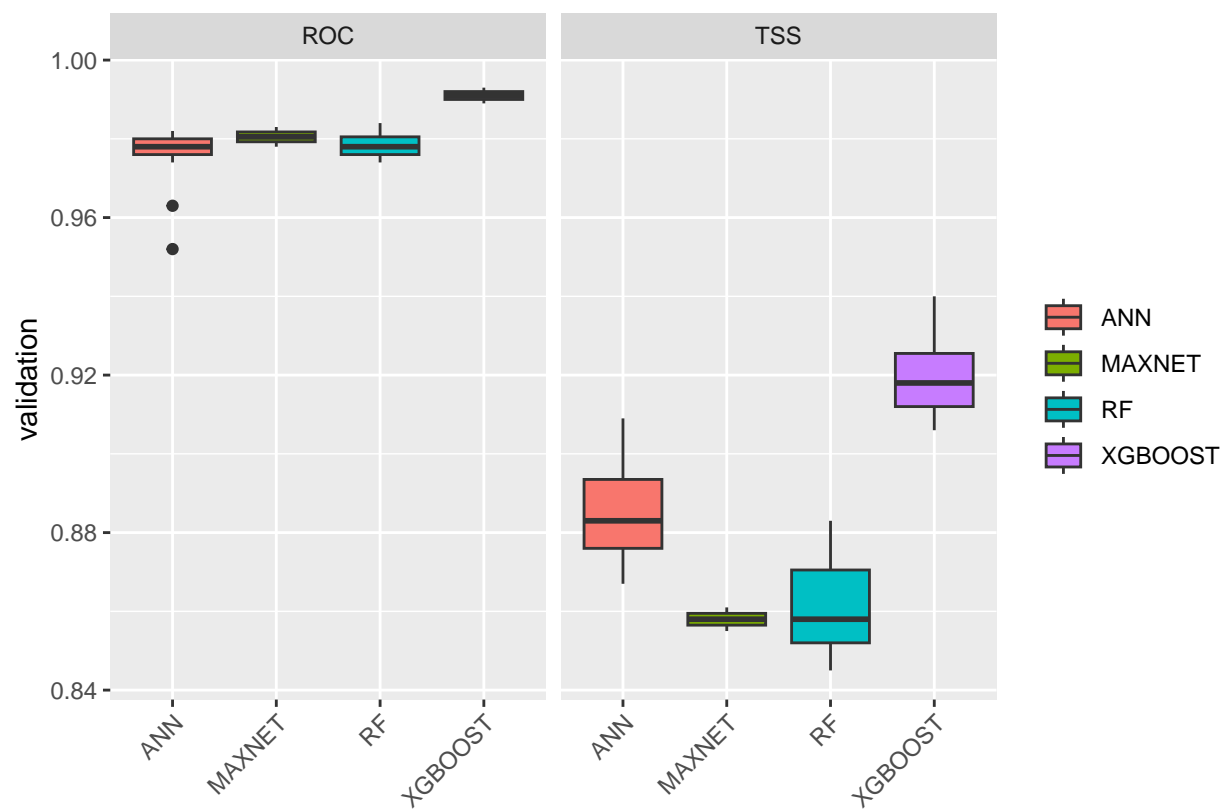
Distribution of presences and generated pseudo-absences:

Presences		True Absences		Pseudo-Absences	
●	calibration	●	calibration	●	calibration
●	validation	●	validation	●	validation
▲	evaluation	▲	evaluation		

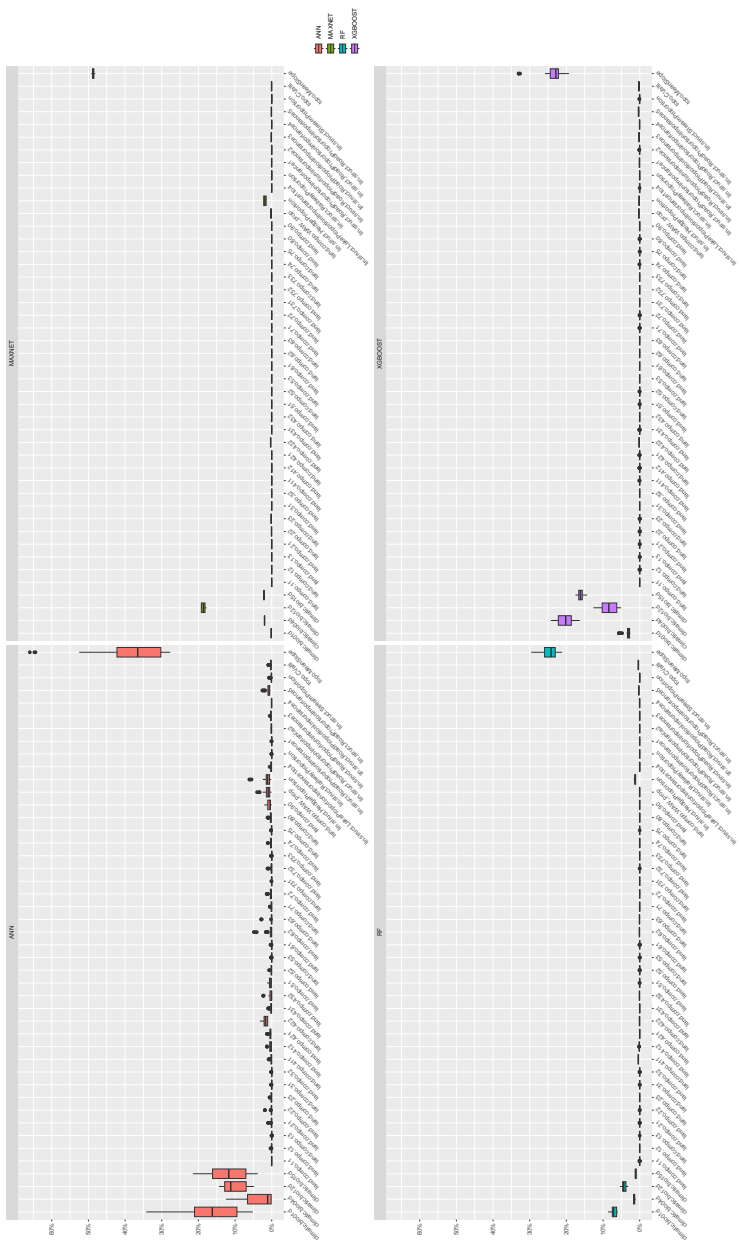


Distribution of evaluation metrics for the calibration and validation datasets:





Importance of the environmental variables for the calculation of presence probabilities:



Projection of the ensemble models over the whole territory to predict group presence probabilities (multiplied by 1000 here):

