BIOMOD_JointEvaliation

# Determines the most suitable algorithm from biomod2 package for your species distribution model according to the joint consideration of metrics of predictive performance and overfitting.

## Description

This function chooses the most suitable algorithm available in the biomod2 package (Thuiller et al., 2009) for your species distribution model according to the joint consideration of metrics of predictive performance and overfitting. The election is based on a weighted sum of scores obtained by the algorithm. A score corresponds to the ranking obtained by the algorithm over-under the others according to a specific predictive performance or overfitting metric.

## Usage

BIOMOD_AlgoEvaluation(

  models = x,

  ensMods = NULL,

  metrics = NULL,

  weights = NULL

)

## Arguments

`models`
Object of class biomod2 of individual models (different algorithms) fitted for a species trough BIOMOD_Modeling function. It can be a list containing fitted models for various species, where each item in the list is the modelling result for a particular species. Those models should have at least ten "RUNs" for each species, that is, ten calibration of models using randomly a certain proportion of the data (i.e., fitting) and another for evaluation (i.e., testing)

`ensMods`
Optional. Object of class biomod2 of ensemble models constructed for a species trough BIOMOD_EnsembleModeling function. It can be a list containing ensemble models for various species. Ensemble models listed must have the same species order as in the "models" argument.

`metrics`
A character vector containing metrics of predictive performance and overfitting to be used to evaluate and rank the different species distribution model algorithms. By default the vector is c("AUC", "TSS", "KAPPA", "OF.AUC", "OF.TSS", "OF.KAPPA"), being the first three elements predictive

performance metrics and the other three overfitting metrics measured through the same predictive performance, respectively. Sensitivity and specificity using the codes "SEN" and "SPE" can also be used as predictive performance metrics and the overfitting measured through those metrics using "OF.SEN" and "OF.SPE" codes.

weigthts    A numeric vector containing the weights for the predictive performance and overfitting metrics. By default, each metric has the same weight if you don't specify the values. The order of the weights must be consistent with the order of the "metrics" argument vector. Make sure that the sum of the weights of the vector will be equal to 1.

## Details

The function chooses the most suitable algorithm available in the biomod2 package (Thuiller et al., 2009) for your species distribution model according to the joint consideration of metrics of predictive performance and overfitting. The election is made to the algorithm that reaches the highest value of a weighted sum of scores. A score corresponds to ranking obtained by an algorithm over-under the others according to a specific predictive performance or overfitting metric. To obtain the sum of scores, the mean value of each metric is first obtained by:

$$\bar{M}_{ij} = \frac{1}{n}\sum_{k=1}^{n} M_{ijk},$$

where $i$ corresponds to the $i$-th metric of predictive performance or overfitting for $j$-th algorithm of $k$-th-fold cross-validation iteration. Then, the score for each metric of predictive performance is obtained normalizing metrics result trough:

$$M.Score_{ij} = \frac{\bar{M}_{ij} - min(\bar{M}_{ij})}{max(\bar{M}_{ij}) - min(\bar{M}_{ij})},$$

while for overfitting metrics was by:

$$M.Score_{ij} = -\left(\frac{\bar{M}_{ij} - min(\bar{M}_{ij})}{max(\bar{M}_{ij}) - min(\bar{M}_{ij})}\right) + 1.$$

In this way, metrics scores can vary in a range that is between 0 and 1. Therefore, if an algorithm obtains a value equal to 1 for a predictive performance or overfitting, it means that the algorithm showed the highest predictive performance or the lowest overfitting, respectively.

Lastly, the final score (sums of scores) of an algorithm is obtained by:

$$F_j = \sum_{i=1}^{n} M.Score_{ij} \times W_i,$$

where $W$ corresponds to the weight assigned to the $i$-th metric. By default, the function uses the same weight for each metric.

The predictive performance can be measured by three metrics used to measure SDM's discrimination capacity: 1) area under the curve of a receiver operating characteristic (AUC; Fielding & Bell, 1997); b) the true skill statistic (TSS; Allouche et al., 2006) and; c) Cohen's Kappa (KAPPA; Cohen, 1960).

Overfitting are measured through same predictive performance metrics indicated above. They are calculated as the difference between predictive performance presented by an algorithm on data used for calibration and its predictive performance presented on data used for testing. For example, the overfitting measured trough AUC is calculated by:

$$OF.AUC_j = \frac{1}{n} \sum_{k=1}^{n} AUC.trainData_{jk} - AUC.testData_{jk},$$

where $j$ is the $j$-th algorithm and $k$ is the $k$-th-fold cross-validation iteration.

## Value

A data.frame with the best ranked algorithm according to the sum of weighted scores, segregated by modeled species. In addition, output files are generated with detailed information about results by species in the "algoRankResults" folder

## References

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species

distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, *43*(6), 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*(1), 38–49.

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, *32*(3), 369–373. https://doi.org/10.1111/j.1600-0587.2008.05742.x