Package ‘vandalico’

May 8, 2026

Type Package

Title Evaluation of Presence-Absence Models

Version 0.2.2

Date 2026-04-14

Maintainer Alberto Jiménez-Valverde <`alberto.jimenez.valverde@gmail.com`>

Description Collection of functions to evaluate presence-absence models. It comprises functions to adjust discrimination statistics for the representativeness effect through case-weighting, along with functions for visualizing the outcomes. Originally outlined in: Jiménez-Valverde (2022) The uniform AUC: dealing with the representativeness effect in presence-absence models. Methods Ecol. Evol, 13, 1224-1236.

Acknowledgments This work was supported by the Ministerio de Ciencia, Innovación y Universidades/Agencia Estatal de Investigación/FEDER (PID2023-147672NB-I00), and by CSIC (2024ICT253).

License GPL-3

Imports ROCR (>= 1.0-7), stats, graphics

Depends R (>= 4.3.0)

RoxygenNote 7.3.3

Encoding UTF-8

NeedsCompilation no

Author Alberto Jiménez-Valverde [aut, cre] (ORCID: <`https://orcid.org/0000-0001-9962-2106`>)

Repository CRAN

Date/Publication 2026-04-14 10:20:02 UTC

Contents

Index **9**

`AUCuniform` *Calculates the uniform* AUC *and uniform* Se*

Description

This function computes the uniform *AUC* (*uAUC*) and uniform *Se** (*uSe**) following Jiménez-Valverde (2022). A revised and improved formulation is available in `AUCuniform.2`, which users are encouraged to consider for most applications. `AUCuniform` is retained for completeness and reproducibility.

Usage

```
AUCuniform(
  mat,
  rep = 100,
  by = 0.1,
  deleteBins = NULL,
  plot = FALSE,
  plot.adds = FALSE
)
```

Arguments

`mat`	A matrix with two columns. The first column must contain the suitability values (i.e., the classification rule); the second column must contain the presences and absences.
`rep`	Number of sampling replications. By default, `rep` = 100.
`by`	Size of the suitability intervals (i.e., bins). By default, `by` = 0.1.
`deleteBins`	A vector (e.g., from 1 to 10 if `by` = 0.1) with the bins that have to be excluded (1 for [0,0.1), 10 for [0.9,1]) from the resampling procedure (trimming); `NULL` by default.
`plot`	Logical. Indicates whether or not the observed ROC curve is plotted.
`plot.adds`	Logical. Indicates whether or not the negative diagonal and the point of equivalence are added to the observed ROC plot.

Details

This function performs the stratified weighted bootstrap to calculate the uniform *AUC* (*uAUC*) and uniform *Se** (*uSe**) as suggested in Jiménez-Valverde (2022). A warning message will be shown if the sample size of any bin is zero. Another warning message will be shown if the sample size of any bin is lower than 15. In such case, trimming should be considered. The *AUC* (non-uniform) is estimated non-parametrically (Bamber 1975). *Se** is calculated by selecting the point that minimizes the absolute difference between sensitivity and specificity and by doing the mean of those values (Jiménez-Valverde 2020).

Value

A list with the following elements:

`AUC`: the *AUC* value (non-uniform), a numeric value between 0 and 1.

`Se`: the *Se** value (non-uniform), a numeric value between 0 and 1.

`bins`: a table with the sample size of each bin.

`suit.sim`: a matrix with the bootstrapped suitability values.

`sp.sim`: a matrix with the bootstrapped presence-absence data.

`uAUC`: a numeric vector with the (*uAUC*) values for each replication.

`uAUC.95CI`: a numeric vector with the sample (*uAUC*) quantiles corresponding to the probabilities 0.025, 0.5 and 0.975.

`uSe`: a numeric vector with the (*uSe**) values for each replication.

`uSe.95CI`: a numeric vector with the sample (*uSe**) quantiles corresponding to the probabilities 0.025, 0.5 and 0.975.

References

Bamber, D. (1975). The Area above the Ordinal Dominance Graph and the Area below the Receiver Operating Characteristic Graph. *J. Math. Psychol.*, 12, 387-415.

Jiménez-Valverde, A. (2020). Sample size for the evaluation of presence-absence models. *Ecol. Indic.*, 114, 106289.

Jiménez-Valverde, A. (2022). The uniform AUC: dealing with the representativeness effect in presence-absence models. *Methods Ecol. Evol.*, 13, 1224-1236.

Examples

```
suit<-rbeta(100, 2, 2) # Generate suitability values
random<-runif(100)
sp<-ifelse(random < suit, 1, 0) # Generate presence-absence data
result<-AUCuniform(cbind(suit, sp), plot = TRUE, plot.adds = TRUE)
result$uAUC.95CI[2] # Get the uAUC
```

`AUCuniform.2` *Calculates the uniform* AUC *and uniform* Se* *by the direct weighted trapezoidal estimation method.*

Description

This function computes the uniform *AUC* (*uAUC*) and uniform *Se** (*uSe**) using the direct weighted trapezoidal estimation method (Jiménez-Valverde 2025), instead of the stratified bootstrapping with inverse probability weighting method implemented in `AUCuniform` and originally proposed by Jiménez-Valverde (2022). Uniform statistics are design to account for the representativeness effect (Jiménez-Valverde 2022). This new method reduces bias and improves the coverage of confidence intervals relative to the original proposal. Additionally, the weight vector associated to each case can be customized.

Usage

```
AUCuniform.2(
  mat,
  by = 0.1,
  deleteBins = NULL,
  w = NULL,
  plot = FALSE,
  plot.compare = FALSE,
  plot.adds = FALSE
)
```

Arguments

`mat`	A matrix with two columns. The first column must contain the classification rule (e.g., the suitability values); the second column must contain the presences and absences.
`by`	The size of the intervals used to divide the classification rule (i.e., bins width). By default, `by` = 0.1. This argument is only used when `w` = `NULL`.
`deleteBins`	A vector (e.g., from 1 to 10 if `by` = 0.1) with the bins that have to be excluded (1 for [0,0.1), 10 for [0.9,1]) from the calculation of the uniform statistics. The default is `NULL`. This argument is only used when `w` = `NULL`.
`w`	A vector with the weights associated with each case. If `NULL` (default), each case is weighted by the inverse of the sample size of its corresponding bin, and the uniform *AUC* (*uAUC*) and uniform *Se** (*uSe**) are calculated (see Details).
`plot`	Logical. If `TRUE`, the standard (unweighted) ROC curve is plotted (gray dots).
`plot.compare`	Logical. If `TRUE`, the weighed ROC curve is plotted (black line).
`plot.adds`	Logical. If `TRUE`, adds the negative diagonal and the points of equivalence (weighted and unweighted) to the ROC plot.

Details

This function calculates the uniform *AUC* (*uAUC*) and uniform *Se** (*uSe**) using the direct weighted trapezoidal estimation method proposed by Jiménez-Valverde (2025). To compute the uniform statistics, the *w* parameter must be set to *NULL* (default). The data set is divided into bins (defined by the parameter `by`) based on the values of the first vector in the input matrix `mat` (the classification rule). Each observation is assigned a weight equal to one divided by the number of observations in the corresponding bin. Then, the uniform discrimination statistics are calculated via the direct weighted trapezoidal estimation method such that, for each threshold, the weighted true positive and false positive rates are cumulatively updated by summing the weights of the presences and absences, respectively, with that score (Jiménez-Valverde 2025). The calculation of the uniform statistics requires the classification rule (`mat[,1]`) to range between 0 and 1, and the value of `by` to divide 1 exactly. If any of this conditions are not match, the function stops. A warning message is displayed if (1) the sample size is lower than 30, (2) any bin has a sample size of zero, or (3) any bin has a sample size between 1 and 15. In the latter case, trimming should be considered using `deleteBins`, in which case the uniform statistics are computed excluding the selected bins. See Jiménez-Valverde (2022) for further details.

Alternatively, users may wish to downweight the importance of certain observations relative to others for reasons unrelated to the representativeness effect (Jiménez-Valverde 2025). For this purpose, the weights associated to each case can be fully customized with the `w` parameter (see Examples). The length of the weight vector has to be equal to `dim(mat)[1]`.

The standard *AUC* (non-uniform, unweighted) is estimated non-parametrically by the trapezoidal rule, which is equivalent to the Wilcoxon-based estimation (Hanley & McNeil 1982) used in `AUCuniform`. *Se** is calculated as in `AUCuniform`.

Value

A list with the following elements:

`AUC`: the standard *AUC* value (unweighted), a numeric value between 0 and 1.

`Se`: the standard *Se** value (unweighted), a numeric value between 0 and 1.

`bins`: a table with the sample size of each bin (returned only if `w = NULL`).

`uAUC`: the uniform *AUC* value (returned only if `w = NULL`).

`uSe`: the uniform *Se** value (returned only if `w = NULL`).

`wAUC`: the weighted *AUC* estimated with the vector `w` (returned only if `w` is not `NULL`).

`wSe`: the weighted *Se** estimated with the vector `w` (returned only if `w` is not `NULL`).

`TP`: a vector with the true positive rate for every threshold.

`FP`: a vector with the false positive rate for every threshold.

`TP.W`: a vector with the weighted true positive rate for every threshold.

`FP.W`: a vector with the weighted false positive rate for every threshold.

References

Hanley, J. A. & McNeil, B. J. (1982). The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology.*, 143, 29-36.

Jiménez-Valverde, A. (2022). The uniform AUC: dealing with the representativeness effect in presence-absence models. *Methods Ecol. Evol.*, 13, 1224-1236.

Jiménez-Valverde, A. (2025). Refining uniform discrimination metrics: towards a case-by-case weighting evaluation in species distribution models with presence-absence data. *Ecol. Evol.*, 15, e72573.

Examples

```
# In this first example, a data set is simulated in such a way that the
# classification rule is well-calibrated, i.e., the observed proportion of
# positive cases equates to the simulated probabilities of presence. Since
# the objective is to calculate the uAUC to account for the environmental
# representativeness effect (see Jiménez-Valverde 2022), weights are
# automatically calculated and no w vector is needed.

n <- 1000 # Set the sample size
hs <- rbeta(n, 2, 2) # Simulated probabilities (the classification rule)
random <- runif(n)
```

```
sp <- ifelse(random < hs, 1, 0) # Observed presence-absence data

result <- AUCuniform.2(cbind(hs, sp), plot = TRUE, plot.compare = TRUE)

result$AUC  # Get the standard AUC
result$uAUC # Get the uniform AUC. Note how it is close to the reference value
            # of 0.83 since the probability values  (the classification rule)
            # are simulated to be well-calibrated (see Jiménez-Valverde 2022)

# In this second set of examples, the objective is not to calculate the
# uniform AUC, but to assign specific weights to certain observations. These
# examples corresponds to some of those provided in Table 1 of
# Jiménez-Valverde (2025).

hs <- seq(1, 0.05, by = -0.05) # Generate the classification rule
sp <- c(0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0) # Observed presence-absence data

wa <- ifelse(sp == 0, 0.2, 1) # The vector of weights for each case
result.a <- AUCuniform.2(cbind(hs, sp), w = wa, plot = TRUE, plot.compare = TRUE)

result.a$AUC  # Get the standard AUC
result.a$wAUC # Get the weighted AUC. Since every case within each category of
              # sp received the same weight, the weighted AUC value equals the
              # standard AUC value

wb <- c(rep(1, 19), 0.2) # The vector of weights for each case
result.b <- AUCuniform.2(cbind(hs, sp), w = wb, plot = TRUE, plot.compare = TRUE)

result.b$wAUC # Get the weighted AUC. Since a low weight is assigned to an
              # instance of absence associated with a low probability value,
              # the weighted AUC is lower than the standard AUC value.

wc <- c(0.2, rep(1, 19)) # The vector of weights for each case
result.c <- AUCuniform.2(cbind(hs, sp), w = wc, plot = TRUE, plot.compare = TRUE)

result.c$wAUC # Get the weighted AUC. Since a low weight is assigned to an
              # instance of absence associated with a high probability value,
              # the weighted AUC is higher than the standard AUC value
```

`CALplot` *Calibration graph*

Description

A function to plot a calibration graph.

Usage

```
CALplot(mat, by = 0.1)
```

Arguments

`mat`	A matrix with two columns. The first column must contain the suitability values (i.e., the classification rule); the second column must contain the presences and absences.
`by`	Size of the suitability intervals (bins). By default, `by` = 0.1.

Details

Dots for bins with 15 or more cases are shown in solid black; dots for bins with less than 15 cases are shown empty (see Jiménez-Valverde et al. 2013). This way, by plotting the calibration graph before running `AUCuniform`, one can get a glimpse of how reliable *uAUC* or *uSe** can be expected to be.

Value

This function returns a calibration plot

References

Jiménez-Valverde, A., Acevedo, P., Barbosa, A. M., Lobo, J. M. & Real, R. (2013). Discrimination capacity in species distribution models depends on the representativeness of the environmental domain. *Global Ecol. Biogeogr.*, 22, 508-516.

Examples

```
suit<-rbeta(100, 2, 2) # Generate suitability values
random<-runif(100)
sp<-ifelse(random < suit,1 , 0) # Generate presence-absence data
CALplot(cbind(suit, sp))
```

`HSgraph` *Suitability values distribution graph*

Description

A function to visualize the distribution of the suitability values associated to presences, absences, and all cases together.

Usage

```
HSgraph(mat, breaks = 10, hist.total = TRUE)
```

Arguments

`mat`	A matrix with two columns. The first column must contain the suitability values (i.e., the classification rule); the second column must contain the presences and absences.
`breaks`	Number of cells for the total histogram. By default, `breaks` = 10.
`hist.total`	Logical. Indicates whether or not the distribution of suitability values for all the cases together is graphed.

Details

In blue, the distribution of the suitability values associated to presences. In red, the distribution of the suitability values associated to absences. This graph helps to understand why the *AUC* (or *Se**) is greater, equal to, or less than the *uAUC* (or *uSe**) (see Jiménez-Valverde 2022).

Value

This function returns a multiple histogram.

References

Jiménez-Valverde, A. (2022). The uniform AUC: dealing with the representativeness effect in presence-absence models. *Methods Ecol. Evol.*, 13, 1224-1236.

Examples

```
suit<-rbeta(100, 2, 2) # Generate suitability values
random<-runif(100)
sp<-ifelse(random < suit, 1 , 0) # Generate presence-absence data
HSgraph(cbind(suit, sp), breaks = 20, hist.total = TRUE)
```

Index