

Package ‘KnowBR’

March 22, 2017

Version 1.2

Title Discriminating Well Surveyed Spatial Units from Exhaustive Biodiversity Databases

Author Cástor Guisande González and Jorge M. Lobo

Maintainer Cástor Guisande González <castor@uvigo.es>

Description It uses species accumulation curves and diverse estimators to assess, at the same time, the levels of survey coverage in multiple geographic cells of a size defined by the user. It also enables the geographical depiction of observed species richness, survey effort and completeness values including a background with administrative areas.

License GPL (≥ 2)

Encoding latin1

Depends R (≥ 3.0)

Suggests vegan, plotrix, fossil

NeedsCompilation no

Repository CRAN

Date/Publication 2017-03-22 19:17:16 UTC

R topics documented:

adworld	2
Beetles	2
KnowB	3

Index	16
--------------	-----------

adworld

GEOGRAPHICAL COORDINATES

Description

Latitude and longitude of all administrative areas

Usage

```
data(adworld)
```

Format

A matrix of many rows and 3 columns (Latitude, Longitude and name of the administrative area)

Source

Latitude and longitude coordinates of the administrative areas were obtained from the web page <https://www.openstreetmap.org>.

Beetles

Individual counts of species of beetles.

Description

This database includes 15,142 records belonging to 54 Iberian species of the Scarabaeidae (Coleoptera) previously compiled in the so called BANDASCA database (Lobo & Martín-Piera, 2002) also freely available in GBIF (<http://www.gbif.org/>). Individual counts, longitude and latitude of species occurrences of the family Scarabaeidae in the Iberian Peninsula are provided.

Usage

```
data(Beetles)
```

Format

A matrix with four columns: species, longitude, latitude and individual counts.

References

Lobo, J.M. & Martín-Piera, F. 2002. Searching for a predictive model for Iberian dung beetle species richness based on spatial and environmental variables. *Conservation Biology* 16: 158-173.

KnowB

Discriminating well surveyed spatial units from exhaustive biodiversity databases.

Description

Advances during the last decades in information technology allow us to store, retrieve, transmit and manipulate an unprecedented magnitude of massive information about species distributions (Guralnick *et al.*, 2007). Unfortunately, this compilation process suffers from three main shortcomings: i) almost all the compiled information represents to varying degrees of spatial precision solely those localities in which a species has been collected without any indication about the probability of its absence (the *unknown absences* drawback); ii) the general lack in the occurrence information of any measure of the survey effort carried out in each one of the considered spatial units (the *unknown efforts* drawback) and; iii) the lack of exhaustive occurrence information due to the fact that records belonging to the same species are generally considered redundant and not retained; thus, preventing an estimation of whether the presence of a species is occasional or frequent (the *unknown recurrence* drawback).

A major consequence of these shortcomings for biogeographical and conservation analyses is the uncertainty as to whether those areas without presence data reflect unoccupied or un-surveyed localities. Thus, geographic representations of the available records are suspiciously similar to the species richness maps (Hortal *et al.*, 2007). The so called Species Distribution data (SDMs), in which the available presence data is related with a variable number of environmental variables have been used in an attempt to overcome these deficiencies in the completeness of biodiversity databases and the inherent biases (Dennis & Hardy, 1999; Soberón *et al.*, 2000; Graham *et al.*, 2004; Lobo *et al.*, 2007); however, these have yielded inconsistent results. *Unknown absences* have promoted the almost universal use in SDMs of pseudo-absences or background absences selected at random from the considered territory. This has been the classic procedure followed for many years in Resource Selection Functions (Johnson, 1980) but is now recognized as unable to estimate the probability of occurrence (Hastie & Fithian, 2013) since it only reflects the intensity of the data used in the modeling procedure (Aarts *et al.*, 2012). The use of more appropriate evaluation methods to measure the discrimination performance of these models (Hijmans, 2012) clearly reveals the inconsistency of the results provided by the joint use of "unreal" absences and complex modeling algorithms. As the employment of statistical shortcuts based on data with an unknown level of error and bias can generate unreliable results, we advocate that the use and/or estimation of biodiversity data requires previous knowledge about the number, location and degree of completeness of relatively well-surveyed inventories. Such knowledge may allow us to discriminate those localities that can be considered to reflect relatively reliable absences, as a guide to the location of further surveys, and to ascertain the areas in which biodiversity data are insufficiently consistent.

KnowBR is utilized to estimate the completeness of species inventories in an unlimited number of spatial units or sites simultaneously without the necessity of using multiple datasets via a batch input mode. It allows species richness estimation using several estimators in multiple geographic cells (spatial units) and the high resolution depiction of the species richness and survey effort in the cells mapped onto a background with administrative areas. For this purpose, we use well-known and widely used estimators and species accumulation techniques that can be calculated at different resolutions and geographical extents. These techniques have been generally used to compare species richness and other diversity parameters between sites differing in their survey efforts. *KnowBR* is

designed for between site species richness comparisons using rarefaction and extrapolation analytic formulas such as those described in Chao & Jost (2012) and Colwell *et al.* (2012). It is intended to estimate how far these sites are from having reliable inventories, how many sites can be considered to be well surveyed, and where these sites are located. Thus, the rationale of this application consists in providing a tool capable of discriminating the location of well surveyed spatial units (WS_{sus}) over a wide geographic expanse using exhaustive databases in which all the available records of a group are gathered without discarding those seemingly redundant data belonging to the same species and locality.

Usage

```
KnowB(data, matrix=TRUE, cell=60, method="accumulation", curve="Clench",
estimator=1, cutoff=2, cutoffCompleteness=0, cutoffSlope=0.1, SpR=FALSE,
largematrix=FALSE, cormatrix=FALSE, Area="World", extent=TRUE, minLon,
maxLon, minLat, maxLat, colbg="#FFFFFF", colcon="#C8C8C8", colf="black",
pro=TRUE, inc=0.005, exclude=NULL, colexc=NULL, colfexc="black",
colscale=c("#FFFFFF", "#C8FFFFFF", "#64FFFFFF", "#00FFFFFF", "#64FF64FF",
"#C8FF00FF", "#FFFF00FF", "#FFC800FF", "#FF6400FF", "#FF0000FF"), legend.pos="y",
breaks=10, xl=0, xr=0, yb=0, yt=0, asp, lab=NULL, xlab="Longitude",
ylab="Latitude", main1="Actual richness", main2="Predicted richness",
main3="Residuals", main4="Records", main5="Completeness", main6="Slope",
cex.main=1.6, cex.lab=1.4, cex.axis=1.2, family="sans", font.main=2, font.lab=1,
font.axis=1, lwdP=0.6, lwdC=0.1, trans=c(1,1), log=c(0,0), ndigits=0, ini=NULL,
end=NULL, ComCut=TRUE, SlopeCut=TRUE, file1="Actual richness.csv",
file2="Predicted richness.csv", file3="Residuals.csv",
file4="List of species.csv", file5="Species per site.csv",
file6="Estimators.csv", file7="Species per record.csv",
file8="Records.csv", file9="Completeness.csv", file10="Slope.csv",
file11="Standard error of the estimators.csv", na="NA", dec=".",
row.names=FALSE, fileEncoding="", jpg=TRUE, jpg1="Actual richness.jpg",
jpg2="Predicted richness.jpg", jpg3="Residuals.jpg",
jpg4="Correlation matrix.jpg", jpg5="Records.jpg", jpg6="Completeness.jpg",
jpg7="Slope.jpg", cex=1.5, pch=15, cex.labels=1.5, pchcol="red", ask=FALSE)
```

Arguments

data	The data is introduced as a CSV file following two simple formats: one including the longitude and latitude of each spatial unit and as many columns as species (see format A in the following table), and other in which only four columns are included (see format B; species name, longitude, latitude and a number reflecting the incidence or abundance of the species). Both incidence and abundance data can be used, although we advocate the use of incidence information since it very difficult to find adequate abundance data reflecting survey effort differences at larger scales. The CSV file may be obtained using ModestR (see details).
------	---

A					B			
Longitude	Latitude	spp1	spp2	spp3	Species	Longitude	Latitude	Counts
-4.694	34.800	10			spp1	-4.694	34.800	10
-4.641	34.741	1			spp1	-4.641	34.741	1
-2.244	34.893	8			spp1	-2.244	34.893	8
-1.443	35.098		3		spp2	-1.443	35.098	3
-4.713	36.140			2	spp3	-4.713	36.140	2
-3.224	38.675			12	spp3	-3.224	38.675	12
-3.213	38.787			23	spp3	-3.213	38.787	23

The primary matrix used in *KnowBR* has a special characteristic - it must be derived from an exhaustive database including all the available georeferenced information including even those apparently redundant records of a species from the same locality provided that is a difference in some of the collection conditions for a species at a locality (i.e. date of capture, food source, collector, type of microhabitat, etc.). Thus, any difference in any database field value yields a new database record regardless of the number of individuals (see for example Lobo & Martín-Piera, 2002). As biodiversity data can derive from heterogeneous sources with different collector methodologies, no universal sampling effort measure capable of offering reliable comparisons exists and the number of database records is used as a surrogate (see Soberón *et al.*, 2007; Lobo, 2008). This approach is particularly appropriate for poorly surveyed groups and/or regions lacking sufficient information to correct unequal sampling efforts arising from standardized survey protocols.

matrix	If TRUE the format of the data frame is species, longitude, latitude and the incidence or abundance of the species (default, format A of the table showed above), if FALSE the format of the data frame is longitude, latitude and the rest of columns are the species with the incidence or abundance values for each site (format B of the table showed above).
cell	Resolution of the cells (spatial units) in minutes on which calculations were carried out. In the present version the user can select any resolution between 1 and 60 minutes but other spatial units may be selected in the near future by including a shape file with the desired regular or irregular polygons.
method	<p>This defines the type of data used and the general method followed in the calculations.</p> <p>Option "incidence" should be used when incidence-based data is used (presence-absence).</p> <p>Option "abundance" uses counts or abundances of species in each spatial unit. When incidence data type is selected all counts must be one.</p> <p>Option "accumulation" (default) estimates accumulation curves and extrapolated predicted richness according to the analytical formulas of Ugland <i>et al.</i> (2003) for the <i>exact</i> estimator (default option) and Coleman <i>et al.</i> (1981; <i>coleman</i> estimator). The accumulated number of species with the addition of database records is also calculated by performing 200 random permutations of the original data with replacement. This procedure smooths the species accumulation curve and avoids potential spurious effects in the curve resulting from the order of record addition. In this case, the smoothed accumulation curve is adjusted to either a Clench or an Exponential function (Soberón and Llorente 1993) to</p>

	<p>calculate the asymptotic extrapolated values (<i>random</i> estimator). The user must take into account the <i>random</i> procedure may generate slightly different final slopes and completeness values depending on the selected random samples. In all these three cases <i>KnowBR</i> calculate: i) completeness values as the percentage of observed species against the predicted number of species, and ii) the slope between the two last steps of the so generated species accumulation curves. This final slope can be considered a measure of the degree of exhaustiveness of the inventory observed in each spatial unit (the attained final rate of increase in the number of species with the survey effort; see Hortal and Lobo (2005)).</p>
curve	<p>The smoothed accumulation curve generated by the <i>random</i> accumulation curve can be adjusted to a "Clench" (default) or to "Exponential" function (Soberón & Llorente, 1993) calculating the asymptotic extrapolated values to further derive a completeness percentage (the percentage representing the observed number of species against the predicted one).</p>
estimator	<p>Vector that defines the used estimator (see details section).</p> <p>In the case of "incidence" data several non-parametric estimators can be calculated:</p> <p>0 the mean predicted species richness obtained with all these five methods is used to estimate the completeness.</p> <p>1 is Chao.</p> <p>2 is ICE.</p> <p>3 is first order jackknife (default estimator for option incidence).</p> <p>4 is the second order jackknife.</p> <p>5 is bootstrap.</p> <p>For "abundance" data the options are:</p> <p>0 the mean predicted species richness obtained with all these two methods is used to estimate the completeness.</p> <p>1 for Chao (unbiased variant).</p> <p>2 for ACE (default estimator for option abundance).</p> <p>In "accumulation" default option:</p> <p>0 the mean predicted species richness obtained with all these four methods is used to estimate the completeness.</p> <p>1 is "exact" that finds the expected (mean) species richness according to Ugland <i>et al.</i> (2003) (default estimator for the option accumulation).</p> <p>2 is "random" that adds sites at random performing 200 random permutations.</p> <p>3 is "coleman" that finds the expected richness following Coleman <i>et al.</i> (1982).</p> <p>4 is "rarefaction" finds the mean when accumulating individuals instead of sites.</p> <p>In each case a completeness percentage is calculated and mapped (the percentage representing the observed number of species against the predicted one) for the selected estimator. Predicted species richness and completeness values were calculated following the specaccum function of the vegan R package. The final slope of the observed data in each spatial unit is also calculated for each one of the formerly mentioned estimators.</p>
cutoff	<p>This number reflects the ratio between the number of database records and the number of species. If this ratio is lower than the selected threshold value in each</p>

	considered spatial unit, any one of the estimators will be calculated and these spatial units are considered as lacking information.
cutoffCompleteness	If the value of completeness is lower than this threshold, any one of the estimators will be calculated and the spatial unit considered as lacking information.
cutoffSlope	If method = "accumulation" and the slope is higher than this threshold, any one of the estimators will be calculated and the spatial unit considered as lacking information.
SpR	If the format of the data frame is species, longitude, latitude and the record, incidence or abundance of the species (matrix=TRUE), then generating the species by record file is computationally intensive. If this argument is FALSE (default) the file is generated but without the longitude and latitude of the sites.
largematrix	When there many species and/or many records resulting in a species per record matrix with more than 2^{31} cells, it is impossible to create the CSV file with the species per record due to memory limits in R. If this argument is TRUE, the function creates a TXT file with the species per record, but the process is computationally intensive, so it may takes several hours and it may create a large TXT file. The default value is FALSE.
cormatrix	If TRUE a correlation matrix among estimators is plotted.
Area	A character with the name of the administrative area or a vector with several administrative areas (see details).
extent	If TRUE the minimum and maximum longitudes and latitudes are delimited by the minimum and maximum of the data (default). If FALSE the minimum and maximum longitudes and latitudes are delimited by the arguments Area and, minLat, maxLat, minLon and maxLon.
minLon, maxLon	Optionally it is possible to define the minimum and maximum longitude (see details).
minLat, maxLat	Optionally it is possible to define the minimum and maximum latitude (see details).
colbg	Background color of the map (in some cases this is the sea).
colcon	Background color of the administrative areas.
colf	Color of administrative areas border.
pro	If it is TRUE an automatic calculation is made in order to correct the aspect ratio y/x along latitude.
inc	Adds some room along the map margins with the limits x and y thus not exactly the limits of the selected areas.
exclude	A character with the name of the administrative area or a vector with several administrative areas that can be plotted with a different color on the map.
colexc	Background color of areas selected in the argument exclude.
colfexc	Color of borders of the areas selected in the argument exclude.
colscale	Palette color.
legend.pos	Whether to have a horizontal (x) or vertical (y) color gradient.
breaks	Number of breakpoints of the color legend.

<code>xl,xr,yb,yt</code>	The lower left and upper right coordinates of the color legend in user coordinates.
<code>asp</code>	The y/x aspect ratio.
<code>lab</code>	A numerical vector of the form <code>c(x, Y)</code> which modified the default method by which axes are annotated. The values of <code>x</code> and <code>y</code> give the (approximate) number of tick marks on the <code>x</code> and <code>y</code> axes.
<code>xlab</code>	A title for the <code>x</code> axis.
<code>ylab</code>	A title for the <code>y</code> axis.
<code>main1</code>	An overall title for the plot of the actual species richness.
<code>main2</code>	An overall title for the plot of the predicted species richness.
<code>main3</code>	An overall title for the plot of the residuals.
<code>main4</code>	An overall title for the plot of the records
<code>main5</code>	An overall title for the plot of the completeness
<code>main6</code>	Only for the accumulation method. An overall title for the plot of the slope between the last species richness value and the previous value for each one of the accumulation methods.
<code>cex.main</code>	The magnification to be used for main titles relative to the current setting of <code>cex</code> .
<code>cex.lab</code>	The magnification to be used for <code>x</code> and <code>y</code> labels relative to the current setting of <code>cex</code> .
<code>cex.axis</code>	The magnification to be used for axis annotation relative to the current setting of <code>cex</code> .
<code>family</code>	The name of a font family for drawing text.
<code>font.main</code>	The font to be used for plot main titles.
<code>font.lab</code>	The font to be used for <code>x</code> and <code>y</code> labels.
<code>font.axis</code>	The font to be used for axis annotation.
<code>lwdP</code>	Line width of the plot.
<code>lwdC</code>	Line width of the borders.
<code>trans</code>	It is possible to multiply or divide the dataset by a value. For a vector with two values, the first may be 0 (divide) or 1 (multiply), and the second number is the value of the division or multiplication.
<code>log</code>	It is possible to apply a logarithmic transformation to the dataset. For a vector with two values, the first may be 0 (do not log transform) or 1 (log transformation), and the second number is the value to be added in case of log transformation.
<code>ndigits</code>	Number of decimals in legend of the color scale.
<code>ini</code>	Maps. Minimum to be considered in the maps of the actual and predicted richness.
<code>end</code>	Maps. Maximum to be considered in the maps of the actual and predicted richness.
<code>ComCut</code>	Maps. If TRUE the minimum to be considered in the map of completeness is the one defined in the argument <i>cutoffCompleteness</i> .

SlopeCut	Maps. If TRUE the maximum to be considered in the map of the slope is the one defined in the argument <i>cutoffSlope</i> .
file1	CSV file. A character string naming the file with the actual richness.
file2	CSV file. A character string naming the file with the predicted richness.
file3	CSV file. A character string naming the file with the residuals.
file4	CSV file. A character string naming the file with the list of species.
file5	CSV file. A character string naming the file with the species incidences per site.
file6	CSV file. A character string naming the file with the estimators per site.
file7	CSV file. A character string naming the file with the species per records.
file8	CSV file. A character string naming the file with the records.
file9	CSV file. A character string naming the file with the completeness.
file10	CSV file. A character string naming the file with the slopes of the accumulation analyses.
file11	CSV file. A character string naming the file with the standard error of the estimators.
na	CSV files. The string to use for missing values in the data.
dec	CSV files. The string to use for decimal points in numeric or complex columns: must be a single character.
row.names	CSV files. Either a logical value indicating whether the row names of x are to be written along with x, or a character vector of row names to be written.
fileEncoding	CSV files. Character string: if non-empty declares the encoding to be used on a file (not a connection) so the character data can be re-encoded as they are written.
jpg	If TRUE the plots are exported to jpg files instead of using the windows device.
jpg1	Name of the jpg file with the values of the actual richness.
jpg2	Name of the jpg file with the values of the predicted richness.
jpg3	Name of the jpg file with the residuals.
jpg4	Name of the jpg file with the correlation matrix among estimators.
jpg5	Name of the jpg file with the records.
jpg6	Name of the jpg file with the completeness.
jpg7	Name of the jpg file with the slopes of the accumulation analyses.
cex	A numerical value giving the amount by which plotting symbols should be magnified relative to the default in the correlation matrix plot.
pch	Either an integer specifying a symbol or a single character to be used as the default in plotting points in the correlation matrix plot.
cex.labels	Size of labels in the correlation matrix plot.
pchcol	Color of the symbols in the correlation matrix plot.
ask	If TRUE (and the R session is interactive) the user is asked for input before a new figure is drawn.

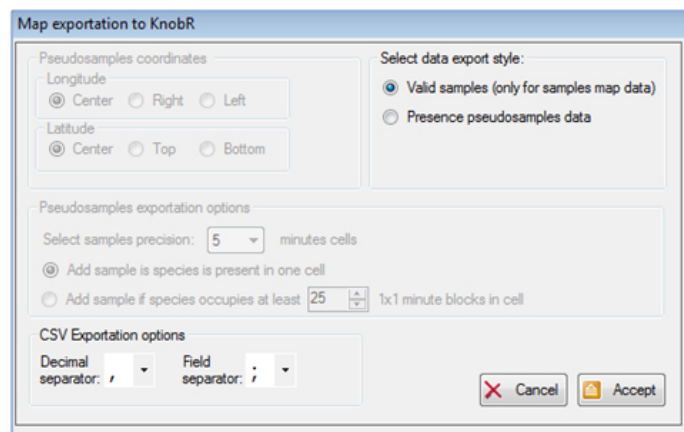
Details

The CSV file required in the argument *data* with the format B (species, longitude, latitude and the incidence) may be obtained using ModestR (available at the web site www.ipez.es/ModestR) as shown in the following screenshot (Export/Export maps of the select branch/To RWizard Applications/To KnowBR).

The menu shown in the following screenshot is obtained when it is possible to export the valid samples or pseudosamples. The pseudosamples are grid cells for instance of 5' x 5', 30' x 30', 1° x 1°, etc. Therefore, the output of ModestR is a list of species within each of the grid cells with the cell size defined by the user. It is therefore possible to obtain the number of records for each species within the grid cell or just the records available for all the species, with the format described above.

It is possible to use all, just one or several estimators, i.e., in the "incidence" method c(1,2), indicates to use Chao and ICE estimators. In the "accumulation" method, c(2,3) indicates to use the "random" and "coleman" methods. When several estimators are selected, the output value is the mean of all of them.

Area = "World" to plot the entire world. If the coordinates minLon, maxLon, minLat and maxLat are not specified, they are calculated automatically based on the selected administrative areas.

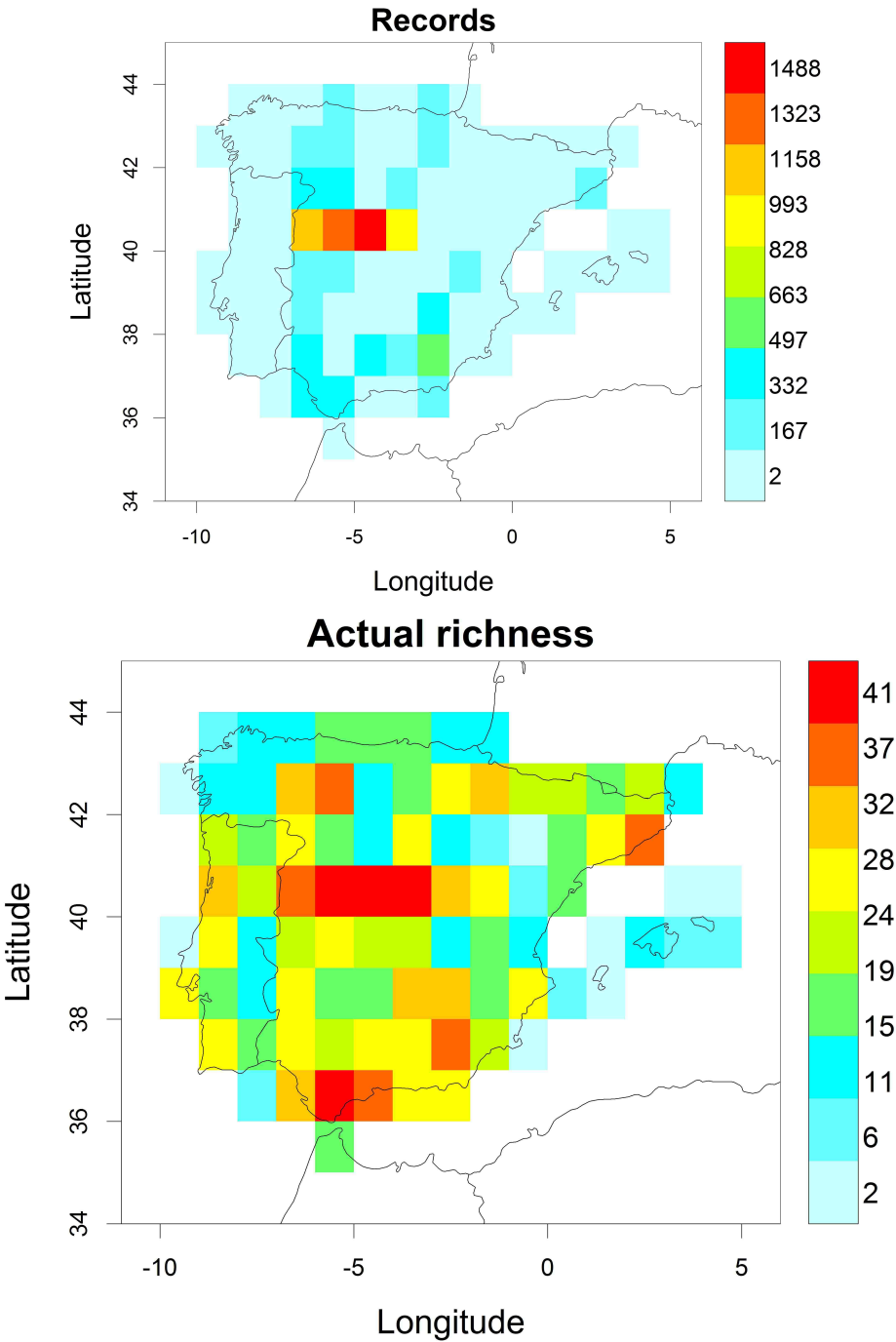


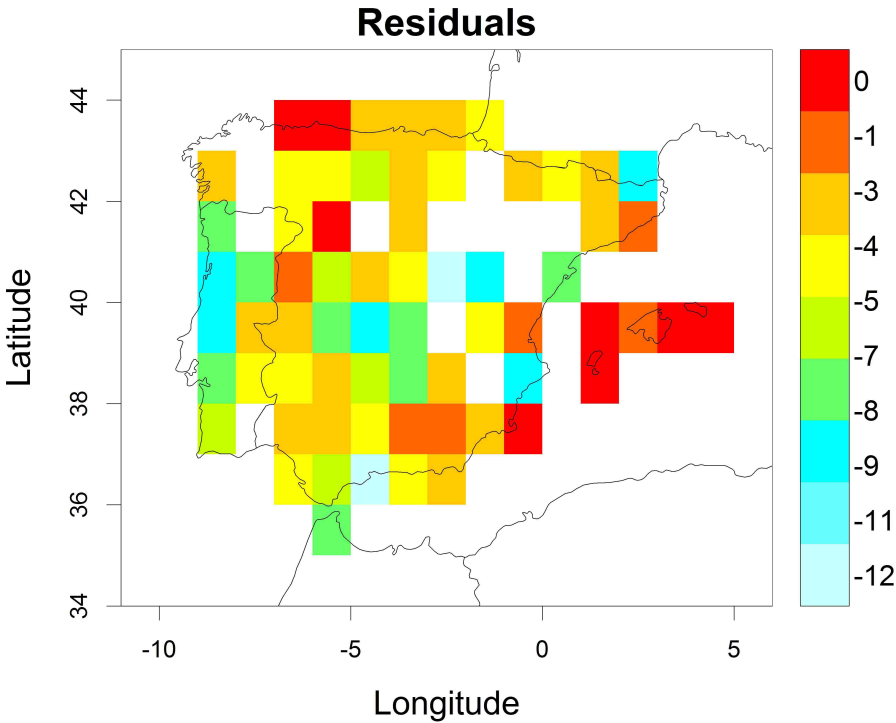
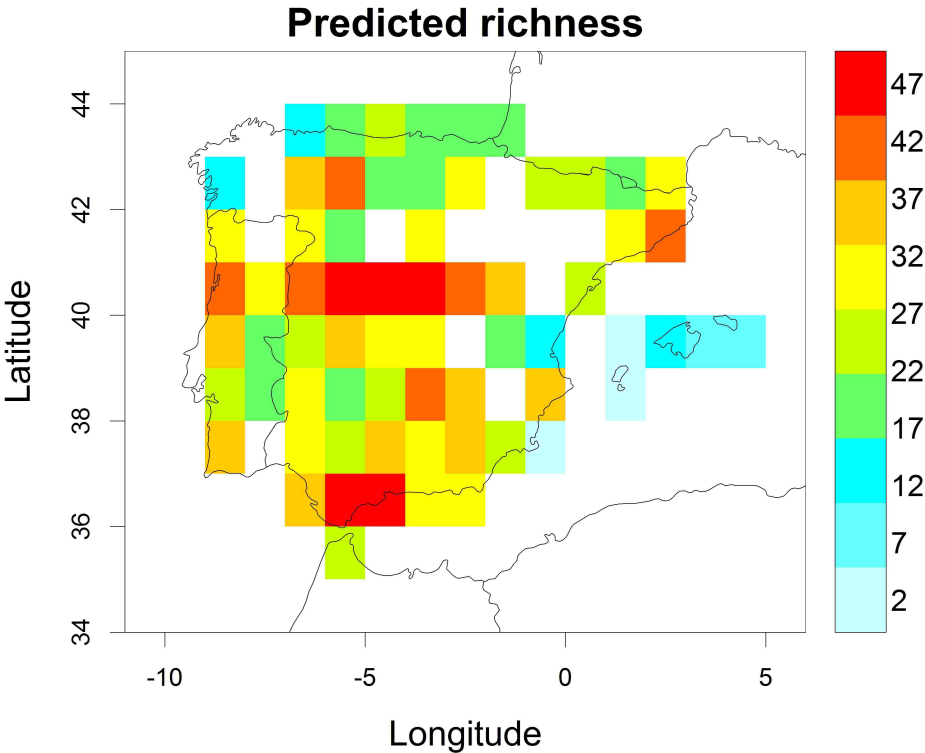
FUNCTIONS

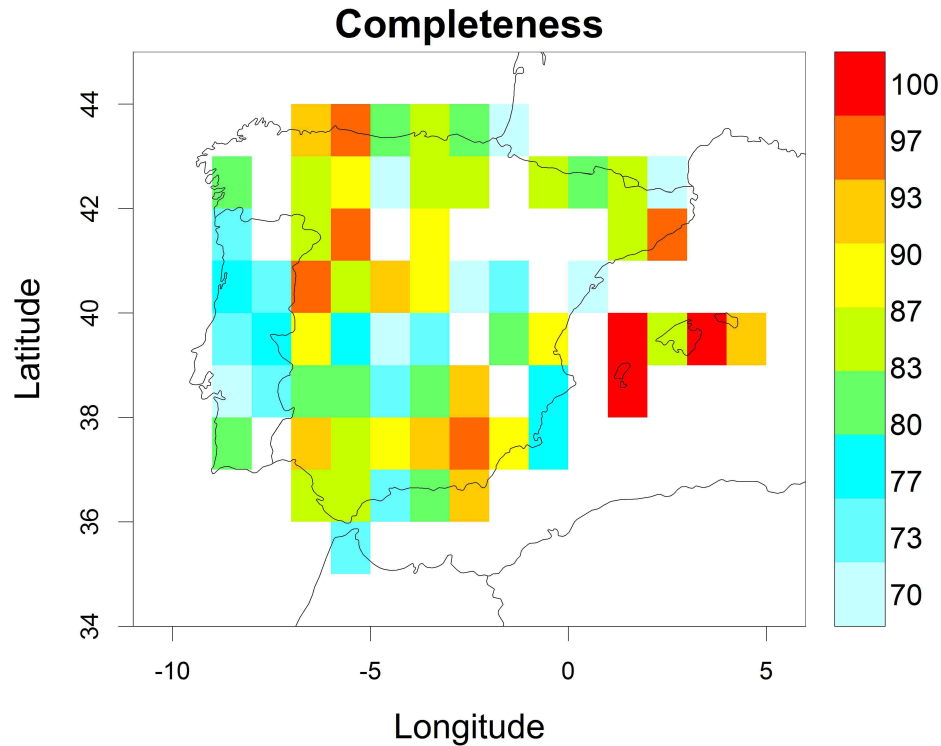
The estimators Chao, first order jackknife, second order jackknife and bootstrap of the option incidence were estimated with the function `specpool` of the package `vegan` (Oksanen *et al.*, 2014). The estimator ICE of the option incidence was estimated with the function `ICE` of the package `fossil` (Vavrek, 2014). The estimators Chao1 and ACE of the option abundance were estimated with the function `estimateR` of the package `vegan` (Oksanen *et al.*, 2014). The estimators exact, random and coleman of the option accumulation were estimated with the function `specaccum` of the package `vegan` (Oksanen *et al.*, 2014).

EXAMPLE

The database of the example includes 15,142 records for the 54 Iberian species of the Scarabaeidae (Coleoptera) previously compiled in the so called BANDASCA database (Lobo & Martín-Piera, 2002). The following maps show the records, actual and predicted species richness, the residuals and completeness obtained in cells of 60' x 60' using the first order jackknife estimator.







Any of these maps may be easily modified using the RWizard application *MapsR*, with the exported CSV files detailing the actual species richness, the predicted species richness, the residuals and/or the completeness.

Value

CSV files: 1) Actual richness, 2) Predicted richness, 3) Residuals, 4) List of species, 5) Species per site, 6) Estimators, 7) Species per record, 8) Records, 9) Completeness, 10) Slope (only if the method is "accumulation") and 11) Standard error of the estimators.

JPG files with maps: Actual richness, 2) Predicted richness, 3) Residuals, 4) Correlation matrix (only if the argument *cormatrix* is TRUE), 5) Records, 6) Completeness and 7) Slope (only if the method is "accumulation").

Source

Spatial database of the location of the world's administrative areas (or administrative boundaries) was obtained from the Web Site <http://www.openstreet.org/>.

References

Aarts, G., Fieberg, J. & Matthiopoulos, J. 2012. Comparative interpretation of count, presence-absence and point methods for species distribution models. *Methods in Ecology and Evolution* 3: 177-187.

- Chao, A. & Jost, L. 2012. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology* 93: 2533-2547.
- Coleman, B.D, Mares, M.A., Willis, M.R. & Hsieh, Y. 1982. Randomness, area and species richness. *Ecology* 63: 1121-1133.
- Colwell, R.K., Chao, A., Gotelli, N.J., Lin, S-Y., Mao, C. X., Chazdon, R.L. & Longino, J.T. 2012. Models and estimations linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology* 5: 3-21.
- Dennis R.L.H. & Hardy, P.B. 1999. Targeting squares for survey: predicting species richness and incidence for a butterfly atlas. *Global Ecology and Biogeography* 8: 443-454.
- Graham, C.H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A.T. 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution* 19: 497-503.
- Guralnick, R.P., Hill, A.W. & Lane, M. 2007. Towards a collaborative, global infrastructure for biodiversity assessment. *Ecology Letters* 10: 663-672.
- Hastie, T. & Fithian, W. 2013. Inference from presence-only data; the ongoing controversy. *Ecography* 36:864-867
- Hijmans, R.J. 2012. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* 93:679-688.
- Hortal, J. & Lobo, J.M. 2005. An ED-based protocol for the optimal sampling of biodiversity. *Biodiversity and Conservation*, 14: 2913-2947.
- Hortal, J., Lobo, J.M. & Jiménez-Valverde, A., 2007. Limitations of biodiversity databases: case study on seed-plant diversity in Tenerife (Canary Islands). *Conservation Biology* 21, 853-863.
- Johnson, D.H. 1980. The comparison of usage and availability measurements for evaluating resource preference. *Ecology* 61: 65-71.
- Lobo, J.M. & Martín-Piera, F. 2002. Searching for a predictive model for Iberian dung beetle species richness based on spatial and environmental variables. *Conservation Biology* 16: 158-173.
- Lobo, J.M., Baselga, A., Hortal, J., Jiménez-Valverde, A. & Gómez, J.F. 2007. How does the knowledge about the spatial distribution of Iberian dung beetle species accumulate over time? *Diversity and Distributions* 13:772-780.
- Lobo, J.M. 2008. Database records as a surrogate for sampling effort provide higher species richness estimations. *Biodiversity and Conservation* 17: 873-881.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H. & Wagner, H. 2014. Community Ecology Package. R package version 2.0-10. Available at: <https://CRAN.R-project.org/package=vegan>.
- Soberón, J. & Llorente, B.J. 1993. The use of species accumulation functions for the prediction of species richness. *Conservation Biology* 7: 480-488.
- Soberón, J., Jiménez, R., Golubov, J. & Koleff, P., 2007. Assessing completeness of biodiversity databases at different spatial scales. *Ecography* 30, 152-160.
- Soberón, J.M., Llorente, J. & Oñate, L. 2000. The use of specimen-label databases for conservation purposes: an example using Mexican Papilionid and Pierid butterflies. *Biodiversity and Conservation* 9 :1441-1466.
- Ugland, K.I., Gray J. S. & Ellingsen, K.E. 2003. The species-accumulation curve and estimation of species richness. *Journal of Animal Ecology* 72: 888-897.
- Vavrek, M.J. 2014. Palaeoecological and Palaeogeographical Analysis Tools. R package version 0.3.7. Available at: <https://CRAN.R-project.org/package=fossil>.

Examples

```
#Example 1 using the method incidence and the estimator first
#order jackknife. If using RWizard, for a better quality of the
#geographic coordinates, replace data(adworld) by @_Build_AdWorld_
```

```
data(adworld)
data(Beetles)
KnowB(data=Beetles, method="incidence", estimator=3, jpg=FALSE)
```

```
## Not run:
```

```
#Only to be used with RWizard.
#Example 2 using the method accumulation (default).
```

```
@_Build_AdWorld_
data(Beetles)
KnowB(Beetles, cell=15)
```

```
## End(Not run)
```

Index

*Topic **Beetles**

Beetles, [2](#)

*Topic **KnowB**

KnowB, [3](#)

*Topic **adworld**

adworld, [2](#)

adworld, [2](#)

Beetles, [2](#)

estimateR, [10](#)

ICE, [10](#)

KnowB, [3](#)

specaccum, [6](#), [10](#)

specpool, [10](#)