

Package ‘assignR’

December 4, 2024

Type Package

Title Infer Geographic Origin from Isotopic Data

Version 2.4.3

Description Routines for re-scaling isotope maps using known-origin tissue isotope data, assigning origin of unknown samples, and summarizing and assessing assignment results. Methods are adapted from Wunder (2010, in ISBN:9789048133536) and Vander Zanden, H. B. et al. (2014) <doi:10.1111/2041-210X.12229> as described in Ma, C. et al. (2020) <doi:10.1111/2041-210X.13426>.

Imports mvnfast, rlang, geosphere, terra (>= 1.7-23)

Depends R (>= 4.0)

Suggests knitr, rmarkdown, testthat, covr

VignetteBuilder knitr

License GPL-3

Language en-US

Encoding UTF-8

LazyData true

NeedsCompilation no

Author Chao Ma [aut],
Gabe Bowen [aut, cre]

Maintainer Gabe Bowen <gabe.bowen@utah.edu>

Repository CRAN

Date/Publication 2024-12-04 15:20:09 UTC

Contents

assignR	2
c.wDist	3
calRaster	4
d2h_lrNA	5
getIsoscapes	6

isoStack	7
jointP	8
knownOrig	8
naMap	10
oddsRatio	11
pdRaster	12
plot.isoStack	13
plot.QA	14
plot.wDist	15
QA	16
qtlRaster	19
refTrans	20
sr_MI	21
states	22
stds	22
subOrigData	24
unionP	26
wDist	26
wrld_simpl	28
Index	29

 assignR

Tools for inferring geographic origin from isotopic data

Description

Routines for rescaling isoscapes using known-origin tissue isotope data, assigning origin of unknown samples, and summarizing and assessing assignment results.

Author(s)

Maintainer: Gabe Bowen <gabe.bowen@utah.edu> **Authors:** Chao Ma, Gabe Bowen

See Also

<https://spatial-lab.github.io/assignR/>

c.wDist	<i>Combine method for wDist objects</i>
---------	---

Description

Combine statistics from one or more [wDist](#) objects in a single data frame.

Usage

```
## S3 method for class 'wDist'  
c(...)
```

Arguments

... One or more [wDist](#) objects

Value

data.frame containing sample IDs, distance, and bearing statistics for each sample in ...

See Also

[wDist](#)

Examples

```
# load hydrogen isotope data for human hair in North America  
d = subOrigData(group = "Modern human", mask = naMap, niter = 100)  
  
# rescale from environmental isoscape to tissue isoscape  
r = calRaster(known = d, isoscape = d2h_lrNA, mask = naMap)  
  
# four unknown-origin examples  
id = c("A", "B", "C", "D")  
d2H = c(-110, -90, -105, -102)  
un = data.frame(id,d2H)  
  
# posterior probabilities  
pp = pdRaster(r, unknown = un, mask = naMap)  
  
# random collection locations  
sites = d$data[sample(seq(length(d$data)), 4),]  
  
# generate a wDist object  
wd = wDist(pp, sites)  
  
# combine stats and print  
c(wd)
```

calRaster

*Rescale isoscape using linear regression***Description**

This function uses known-origin tissue data to rescale a map of environmental isotope values to a map of tissue value (and associated uncertainty) using a linear regression model.

Usage

```
calRaster(known, isoscape, mask = NULL, interpMethod = 2, NA.value = NA,
  ignore.NA = TRUE, genplot = TRUE, outDir = NULL, verboseLM = TRUE)
```

Arguments

known	subOrigData or SpatVector. Known-origin tissue isotope data from the subOrigData function or provided by user. User-provided data must be formatted as a subOrigData object (see subOrigData) or a SpatVector with point geometry in which the first data field contains the measured tissue isotope value and the second the 1 standard deviation uncertainty on that value.
isoscape	SpatRaster. Isoscape raster with two layers. The first one is the mean and the second is one standard deviation.
mask	SpatVector. Polygon layer that constrains the area of the output rasters. If this is not provided, the entire area of isoscape is returned.
interpMethod	numeric. 1 or 2. Designate one of two methods for extracting values from isoscape. If 1, values for the cell in which a point falls are returned. If 2 the returned values are interpolated (bilinear interpolation) from the values of the four adjacent grid cells.
NA.value	NA or numeric. Value representing the absence of data in isoscape. Commonly used values include NA, -9999, and 9999.
ignore.NA	logical. If NA values are extracted from isoscape at the location of samples in known, ignore these values and proceed with calRaster.
genplot	logical. Plot the results.
outDir	character string. Directory path to which output will be saved. If NULL no files are written.
verboseLM	logical. Print out the linear regression results.

Value

Returns an object of class “rescale”.

isoscape.rescale

SpatRaster. isoscape rescaled using a linear regression model between values of known and isoscape. The first layer is the mean prediction and the second the standard deviation.

lm.data	data.frame. Known origin data and extracted isoscape values used for linear regression modeling.
lm.model	list. Linear regression model.

See Also

[pdRaster](#)

Examples

```
# load hydrogen isotope data for human hair in North America
d = subOrigData(group = "Modern human", mask = naMap, niter = 100, genplot = FALSE)

# rescale from environmental isoscape to tissue isoscape
r = calRaster(d, d2h_lrNA, naMap)
```

d2h_lrNA	<i>data: low resolution North American growing season H isoscape</i>
----------	--

Description

Interpolated growing season precipitation H isoscape from waterisotopes.org.

Usage

```
d2h_lrNA
```

Format

SpatRaster with two layers. The first layer is the mean prediction and the second is 1 standard deviation

References

- Bowen, G. J. (2018) Gridded maps of the isotopic composition of meteoric waters. <http://www.waterisotopes.org>.
- Bowen, G. J., Wassenaar, L. I. and Hobson, K. A. (2005) Global application of stable hydrogen and oxygen isotopes to wildlife forensics. *Oecologia*, **143**, 337–348.
- IAEA/WMO (2018) Global Network of Isotopes in Precipitation. The GNIP Database. <https://nucleus.iaea.org/wiser>.

Examples

```
library(terra)

plot(d2h_lrNA)
```

getIsoscapes

Download and unpack isoscapes from the web

Description

This function retrieves gridded isotope maps from waterisotopes.org, unpacks the zip archives, and bundles the map layers as a RasterStack.

Usage

```
getIsoscapes(isoType = "GlobalPrecipGS", timeout = 1200)
```

Arguments

isoType	character string indicating which isoscapes are requested: see 'Details'.
timeout	integer. Maximum allowed file download time, in seconds. Some isoscape archives exceed 2 GB in size and may require long download times on slow connections. This option may not work on all system configurations.

Details

Accepted isoType values are:

"GlobalPrecipGS" Global growing-season precipitation H and O isotope values

"GlobalPrecipMA" Global mean-annual precipitation H and O isotope values

"GlobalPrecipMO" Global monthly precipitation H and O isotope values

"GlobalPrecipALL" Global mean-annual and monthly precipitation H and O isotope values

"USPrecipMA" High-resolution contiguous USA mean-annual precipitation H and O isotope values

"USPrecipMO" High-resolution contiguous USA monthly precipitation H and O isotope values

"USPrecipALL" High-resolution contiguous USA mean-annual and monthly precipitation H and O isotope values

"USSurf" High-resolution contiguous USA surface water H and O isotope values

"USTap" High-resolution contiguous USA surface water H and O isotope values

"USGround" Contiguous USA groundwater H and O isotope values in 7 depth intervals

"GlobalSr" High-resolution bioavailable Sr isotope ratios for the global land surface

"USSr" High-resolution contiguous USA Sr isotope ratios

"CaribSr" High-resolution Sr isotope ratios for the circum-Caribbean region

Value

RasterStack containing the requested isoscape layers.

References

https://wateriso.utah.edu/waterisotopes/pages/data_access/ArcGrids.html

Examples

```
## Not run:
iso = getIsoscapes("CaribSr")

## End(Not run)
```

isoStack	<i>Stack isoscapes</i>
----------	------------------------

Description

Combine multiple isoscapes into a single data object, including optional reconciliation of raster properties.

Usage

```
isoStack(..., clean = TRUE)
```

Arguments

...	Two or more SpatRaster isoscapes, each with two layers, or rescale objects as produced by calRaster . For objects other than rescale, the first layer in each should contain the tissue-specific isoscape and the second the isoscape prediction uncertainty (1 standard deviation).
clean	logical. Reconcile differences in raster properties within ...?

Details

If `clean = TRUE` all raster layers are projected to the projection of the first object in ... and then resampled to the highest spatial resolution and smallest common spatial extent within ... Finally, cells containing NA in any layer within ... are masked across all layers.

If `clean = FALSE` any differences in raster properties between isoscapes will produce an error.

Value

Returns an object of class “isoStack”, a list containing the isoscapes objects in ... after any cleaning.

Examples

```
#stack H and Sr isoscapes
h_s = isoStack(d2h_1rNA, sr_MI)
```

jointP *Joint probability of origin*

Description

Joint probability for individuals of common origin (product of probabilities)

Usage

```
jointP(pdR)
```

Arguments

pdR SpatRaster of probability density maps, e.g., as produced by pdRaster. All layers must have common spatial extent and projection.

Value

SpatRaster.

Examples

```
# load hydrogen isotope data for human hair in North America
d = subOrigData(group = "Modern human", mask = naMap, genplot = FALSE)

# rescale from environmental isoscape to tissue isoscape
r = calRaster(d, d2h_lrNA, naMap, genplot = FALSE)

# four unknown-origin examples
id = c("A", "B", "C", "D")
d2H = c(-110, -90, -105, -102)
un = data.frame(id, d2H)

# posterior probabilities
pp = pdRaster(r, un, mask = naMap, genplot = FALSE)

# joint probability for individuals of common origin
jointP(pp)
```

knownOrig *data: database of H and O isotope data for tissues of known origin*

Description

This dataset consists of hydrogen and oxygen isotope values and metadata for human hair, insect wings, and bird feathers of known geographic origin.

Usage

knownOrig

Format

list.

sites SpatVector with 5 fields. WGS84 unprojected geometry.

- [, 1] Site_ID: Unique ID
- [, 2] Site_name: Site name or descriptor
- [, 3] State: State or province of collection site, where recorded
- [, 4] Country: Country of collection site, where recorded
- [, 5] Site_comments: Site comments

samples data.frame with 15 fields.

- [, 1] Sample_ID: Unique ID
- [, 2] Sample_ID_orig: ID used in original data report
- [, 3] Site_ID: ID for sample collection site
- [, 4] Dataset_ID: ID for dataset from which sample is derived
- [, 5] Taxon: Genus and species name
- [, 6] Group: Biological group (informal)
- [, 7] Source_quality: Code indicating level of certainty in geographic origin
- [, 8] Age_class: Code for age of individual
- [, 9] Material_type: Tissue sampled, e.g., “Hair”
- [, 10] Matrix: Compound measured, e.g., “Keratin”
- [, 11] d2H: Hydrogen isotope value (permil)
- [, 12] d2H.sd: Reported analytical uncertainty for hydrogen isotope value (permil)
- [, 13] d18O: Oxygen isotope value (permil)
- [, 14] d18O.sd: Reported analytical uncertainty for oxygen isotope value (permil)
- [, 15] Sample_comments: Sample comments

sources data.frame with 17 fields.

- [, 1] Dataset_ID: Unique ID
- [, 2] Dataset_name: Short name or descriptor
- [, 3] Citation: Bibliographic citation for study
- [, 4] Sampling_method: How material was subsampled for analysis, if reported
- [, 5] Sample_powdered: Was sample powdered prior to analysis (Y/N/NA)?
- [, 6] Lipid_extraction: Were lipids chemically extracted prior to analysis (Y/N/NA)?
- [, 7] Lipid_extraction_method: Solvent used to extract lipids
- [, 8] Exchange: Was a correction for exchangeable H made (Y/N/NA)?
- [, 9] Exchange_method: Method used to correct for exchangeable H
- [, 10] Exchange_T: Was H exchange carried out at ambient or high temperature (Ambient/High/NA)?
- [, 11] H_cal: Reference scale used to calibrate H isotope data, see [stds](#) object hstds

- [, 12] O_cal: Reference scale used to calibrate O isotope data, see [stds](#) object ostds
- [, 13] Std_powdered: Were calibration standards powdered (Y/N/NA)?
- [, 14] Drying: Did the study document how samples were fully dried and transferred dry to instrument (Y/N/NA)?
- [, 15] Analysis_method: Instrument configuration used for analysis
- [, 16] Analysis_type: What elements were analyzed for stable isotope ratios (H/O/H₂O)?
- [, 17] Source_comments: Data source comments

See Also

[subOrigData](#)

Examples

```
library(terra)

class(knownOrig$sites)
class(knownOrig$samples); class(knownOrig$sources)

summary(knownOrig$samples)

print(knownOrig$sources[, 1:2])

plot(wrld_simpl, border = "grey")
points(knownOrig$sites)
```

naMap

data: North America boundary map

Description

Simplified spatial polygon layer representing the boundary of North America.

Usage

```
naMap
```

Format

```
SpatVector
```

Examples

```
library(terra)

plot(naMap)
```

oddsRatio	<i>Odds ratio of points or regions</i>
-----------	--

Description

Calculate ratio of odds for two locations (points or polygons)

Usage

```
oddsRatio(pdR, inputP)
```

Arguments

pdR	SpatRaster of probability density maps, e.g., as produced by pdRaster.
inputP	SpatVector points object of length 1 or 2 or polygons object of length 2

Examples

```
library(terra)

# load hydrogen isotope data for human hair in North America
d = subOrigData(group = "Modern human", mask = naMap, genplot = FALSE)

# rescale from environmental isoscape to tissue isoscape
r = calRaster(d, d2h_lrNA, naMap, genplot = FALSE)

# four unknown-origin examples
id = c("A", "B", "C", "D")
d2H = c(-110, -90, -105, -102)
un = data.frame(id, d2H)

# posterior probabilities
pp = pdRaster(r, un, mask = naMap, genplot = FALSE)

# SpatialPolygons for two regions of interest
s1 = states[states$STATE_ABBR == "UT",]
s2 = states[states$STATE_ABBR == "NM",]
plot(naMap)
plot(s1, border = "red", add = TRUE)
plot(s2, border = "blue", add = TRUE)

# Get odds ratio for two regions using SpatialPolygon method
s12 = rbind(s1, s2)
oddsRatio(pp, s12)

# Create SpatialPoints for two points of interest
p1 = c(-112, 40)
p2 = c(-105, 33)
p12 = vect(rbind(p1, p2), crs = "WGS84")
```

```

points(p12, pch = 21, bg = "light blue")

# Get odds ratio for two points using SpatialPoints method
oddsRatio(pp, p12)

```

pdRaster

Probability of origin surfaces

Description

Calculate posterior probabilities of origin for a sample based on its isotope ratio.

Usage

```
pdRaster(r, unknown, prior = NULL, mask = NULL, genplot = TRUE, outDir = NULL)
```

Arguments

r	SpatRaster with two layers, rescale object (see calRaster), or isoStack object. For user-generated raster objects, the first layer must be the substrate-specific isoscape (mean prediction) and the second the isoscape prediction uncertainty (1 standard deviation).
unknown	data.frame, refTrans object, or list of two or more refTrans objects. For user-created data.frame, first column should contain unique sample IDs, and subsequent columns should contain sample isotope values for one or more isotopes to be used in assignment.
prior	SpatRaster. Optional raster layer with prior probabilities, which has the same projection, resolution and extent as r.
mask	SpatVector. This polygon mask will constrain the assignment area. If this is not provided, a default of mask of the extent of r is used.
genplot	logical. Plot results in R.
outDir	character string. Directory path to which output will be saved. If NULL no files are written.

Details

If more than one isotope marker is to be used for multivariate assignment, r must be an [isoStack](#) object and the number of isoscapes in that object must be equal to the number of isotope-value columns or [refTrans](#) objects included in unknown. Isoscapes and unknown sample values will be matched based on order, so it is critical that the values appear in the same order in these two input objects.

Value

SpatRaster including a probability density surface for each individual in unknown. If outDir is not NULL, writes individual rasters in GeoTIFF format and a single PDF file with images for each probability density raster to the designated directory.

See Also[calRaster refTrans isoStack](#)**Examples**

```
# load hydrogen isotope data for human hair in North America
d = subOrigData(group = "Modern human", mask = naMap, genplot = FALSE)

# rescale from environmental isoscape to tissue isoscape
r = calRaster(d, d2h_lrNA, naMap, genplot = FALSE)

# sample to assign
id = "smile"
d2H = -80
un = data.frame(id, d2H)

# posterior probability surface
pp = pdRaster(r, un, mask = naMap)
```

`plot.isoStack`*Plot method for stacked isoscapes*

Description

Plot the output from [isoStack](#).

Usage

```
## S3 method for class 'isoStack'
plot(x, ...)
```

Arguments

<code>x</code>	An isoStack object
<code>...</code>	Other arguments to be passed to plot

See Also[isoStack](#)**Examples**

```
#stack H and Sr isoscapes
h_s = isoStack(d2h_lrNA, sr_MI)

#plot isoStack
plot(h_s)
```

`plot.QA`*Plot quality metrics for geographic assignments*

Description

Plot the output from [QA](#), including spatial precision, bias, sensitivity and odds ratio of known locations for validation samples.

Usage

```
## S3 method for class 'QA'  
plot(x, ..., outDir = NULL)
```

Arguments

<code>x</code>	One or more QA objects
<code>...</code>	Other arguments to be passed to plot
<code>outDir</code>	character string. Directory path to which output will be saved. If NULL no files are written.

References

Ma, C. et al. (2020) assignR : An R package for isotope-based geographic assignment. *Methods in Ecology and Evolution* **11** 996–1001. doi:10.1111/2041210X.13426.

Vander Zanden, H. B. et al. (2014) Contrasting assignment of migratory organisms to geographic origins using long-term versus year-specific precipitation isotope maps. *Methods in Ecology and Evolution* **5** 891–900. doi:10.1111/2041210X.12229

See Also

[QA](#)

Examples

```
# extract some known-origin data  
d1 = subOrigData(taxon = "Buteo lagopus")  
  
# run quality assessment based on precipitation hydrogen isotopes and  
# known-origin samples; small values of valiStation and valiTime  
# are used in example to reduce run time  
  
# first with one example  
# gives warning because a small number of samples are available  
qa1 = QA(isoscape = d2h_lrNA, known = d1, valiStation = 1,  
         valiTime = 2, by = 10, mask = naMap, name = "Buteo")  
  
# plot the qa result  
plot(qa1)
```

```
# now compare with a second data set
d2 = subOrigData(taxon = "Charadrius montanus")
qa2 = QA(isoscape = d2h_lrNA, known = d2, valiStation = 1,
        valiTime = 2, by = 10, mask = naMap, name = "Charadrius")
plot(qa1, qa2)
```

plot.wDist

Plot weighted distance and bearing distributions

Description

Plot the output from `wDist`, including weighted kernel density distributions for distance and bearing of travel.

Usage

```
## S3 method for class 'wDist'
plot(x, ..., bin = 20, pty = "both", index = c(1:5))
```

Arguments

<code>x</code>	A <code>wDist</code> object
<code>...</code>	Other arguments to be passed to <code>plot</code>
<code>bin</code>	numeric. Bin width used to generate rose plot of travel bearings, in degrees. Must be a factor of 360.
<code>pty</code>	character. Type of plot to produce. Must be one of "dist", "bear", or "both".
<code>index</code>	numeric. Which items in <code>x</code> to plot? Numeric vector of up to 5 integers. Values in excess of 5 or exceeding the length of <code>x</code> will be ignored.

Details

For the default `pty`, two plot panels will be printed to the active graphical device showing the distance and bearing distributions for (up to) the first five samples in `wd`. If more than five items exist in `wd`, those beyond the fifth will be ignored and a message returned.

See Also

[wDist](#)

Examples

```
# load hydrogen isotope data for human hair in North America
d = subOrigData(group = "Modern human", mask = naMap, niter = 100)

# rescale from environmental isoscape to tissue isoscape
r = calRaster(known = d, isoscape = d2h_lrNA, mask = naMap)

# four unknown-origin examples
id = c("A", "B", "C", "D")
d2H = c(-110, -90, -105, -102)
un = data.frame(id,d2H)

# posterior probabilities
pp = pdRaster(r, unknown = un, mask = naMap)

# random collection locations
sites = d$data[sample(seq(length(d$data)), 4),]

# generate a wDist object
wd = wDist(pp, sites)

# plot distributions
plot(wd)

# plot bearing distribution for sample B with a finer bin size
plot(wd, bin = 5, pty = "bear", index = 2)
```

Description

How well does a given isoscape and/or known origin data set constrain the geographic origin of samples? Uses iterative re-sampling of known origin data to evaluate sample assignments and reports a suite of quality metrics.

Usage

```
QA(known, isoscape, bySite = TRUE, valiStation = 1, valiTime = 50,
    recal = TRUE, by = 2, prior = NULL, mask = NULL, setSeed = TRUE,
    name = NULL)
```

Arguments

known subOrigData, list of subOrigData, or SpatVector. Known-origin tissue isotope data from the subOrigData function or provided by user. User-provided data must be formatted as subOrigData objects (see [subOrigData](#)) or a SpatVector (see Details).

<code>isoscape</code>	SpatRaster with two layers or <code>isoStack</code> object. For user-generated raster objects, the first layer must be the isoscape (mean prediction) and the second the isoscape prediction uncertainty (1 standard deviation).
<code>bySite</code>	logical. Resample known by site (TRUE) or by sample (FALSE)?
<code>valiStation</code>	numeric. How many sites or samples from known are withheld for validation? Must be two or more smaller than the length of known.
<code>valiTime</code>	numeric. How many times do you want to randomly draw validation samples and run the validation? Must be an integer greater than one.
<code>recal</code>	logical. Recalibrate the isoscape(s) using the known-origin data? If FALSE, isoscape should be a calibrated product appropriate to the samples, and a single iteration is run for each sample in known; parameters <code>bySite</code> , <code>valiStation</code> , and <code>valiTime</code> are ignored.
<code>by</code>	integer. Threshold increment to use in evaluating assignment performance. Must be between 1 and 25.
<code>prior</code>	SpatRaster. Optional layer with prior probabilities, which has the same projection, resolution and extent as <code>isoscape</code> .
<code>mask</code>	SpatVector. Constrains the area of the analysis. If this is not provided, the entire area of <code>isoscape</code> is used.
<code>setSeed</code>	logical. Do you want to set <code>.seed()</code> when you randomly draw validation stations? “TRUE” gives the same sequence of random draws each time the function is called.
<code>name</code>	character. Useful for identifying the QA output in subsequent plotting.

Details

If `known` is a user-provided `SpatVector`, the first data field must include the measured value for the first (or only) isotope marker and the second the one standard deviation uncertainty on that value. Subsequent fields must include the same information for all other isotope markers included in the analysis, and these markers must appear in the same order as in `isoscape`. A user-provided `SpatVector` must include a field named “`Site_ID`” containing unique values for each sampling site to support the “`bySite`” option, otherwise use `bySite = FALSE`.

Value

Returns an object of class “QA”.

<code>val_stations</code>	numeric. An X*Y data.frame of validation station IDs for all <code>valiTime</code> . X = <code>valiTime</code> and Y = <code>valiStation</code> .
<code>pd_val</code>	numeric. An X*Y data.frame containing the posterior probability density for the validation stations. X = <code>valiTime</code> and Y = <code>valiStation</code> .
<code>prption_byArea</code>	numeric. An X*Y data.frame showing the proportion of validation individuals for which the known origin is contained within the top 0.00 to 1.00 area quantile (with increment of <code>by / 100</code> ; Y = <code>ceiling(100 / by) + 1</code>). X = <code>valiTime</code> .
<code>prption_byProb</code>	numeric. An X*Y data.frame showing the proportion of validation individuals for which the known origin is contained within the top 0.00 to 1.00 probability quantile (with increment of <code>by / 100</code> ; Y = <code>ceiling(100 / by) + 1</code>). X = <code>valiTime</code> .

precision	list. The length of the list is <code>valiTime</code> . Each element is an X*Y matrix showing the proportional area of the total assignment surface covered by the assignment region at a given probability quantile from 0.00 to 1.00 (with increment of <code>by / 100</code> ; <code>X = ceiling(100 / by) + 1</code>) for each validation sample (<code>Y = valiStation</code>).
random_prob_density	Random probability of assignment to any given grid cell on the assignment surface (i.e. 1 divided by the total number of grid cells).
name	character. Name assigned to the QA object.
by	integer. Value of <code>by</code> used.

Note

See Ma et al. (2020) for methodological details.

References

Ma, C. et al. (2020) `assignR` : An R package for isotope-based geographic assignment. *Methods in Ecology and Evolution* **11** 996–1001. doi:10.1111/2041210X.13426.

Vander Zanden, H. B. et al. (2014) Contrasting assignment of migratory organisms to geographic origins using long-term versus year-specific precipitation isotope maps. *Methods in Ecology and Evolution* **5** 891–900. doi:10.1111/2041210X.12229

See Also

[plot.QA](#)

Examples

```
# extract some known-origin data
d1 = subOrigData(taxon = "Buteo lagopus")

# run quality assessment based on precipitation hydrogen isotopes and
# known-origin samples; small values of valiStation and valiTime
# are used in example to reduce run time

# first with one example
# gives warning because a small number of samples are available
qa1 = QA(known = d1, isoscape = d2h_lrNA, valiTime = 2, by = 10,
        mask = naMap, name = "Buteo")

# plot the qa result
plot(qa1)

# now compare with a second data set
d2 = subOrigData(taxon = "Charadrius montanus")
qa2 = QA(known = d2, isoscape = d2h_lrNA, valiTime = 2, by = 10,
        mask = naMap, name = "Charadrius")
plot(qa1, qa2)
```

qtlRaster

*Sample assignment using thresholds***Description**

Selects the grid cells of probability density rasters with the highest probability and returns rasters with these cell values set to 1. Cells are selected based on the user-specified quantile threshold so that the most-probable cells representing a given fraction of the assignment area or posterior probability are returned.

Usage

```
qtlRaster(pdR, threshold, thresholdType = "area", genplot = TRUE, outDir = NULL)
```

Arguments

pdR	SpatRaster. Probability density maps for individual samples, e.g., as output by pdRaster.
threshold	numeric from 0 to 1. Quantile to be selected.
thresholdType	character. Either "area" (default) or "prob". If "area", the most probable cells constituting threshold percent of the assignment area are selected. If "prob", the most probable cells constituting threshold percent of the posterior probability are selected.
genplot	logical. Plot results in R.
outDir	character string. Directory path to which output will be saved. If NULL no files are written.

Value

SpatRaster including a binary assignment surface for each individual in pdR. If outDir is not NULL, writes individual rasters in GeoTIFF format and a single PDF file with images for each raster to the designated directory.

Examples

```
library(terra)

# load hydrogen isotope data for human hair in North America
d = subOrigData(group = "Modern human", mask = naMap, genplot = FALSE)

# rescale from environmental isoscape to tissue isoscape
r = calRaster(d, d2h_lrNA, naMap, genplot = FALSE)

# four unknown-origin examples
id = c("A", "B", "C", "D")
d2H = c(-110, -90, -105, -102)
un = data.frame(id, d2H)
```

```
# posterior probabilities
pp = pdRaster(r, un, mask = naMap, genplot = FALSE)

# assign to most probable 10 percent of area
## Not run: qtlRaster(pp, threshold = 0.1)

# assign to most probable 10 percent of probability distribution
qtlRaster(pp, threshold = 0.1, thresholdType = "prob")
```

refTrans	<i>Transform reference scale of data</i>
----------	--

Description

This function conducts transformations to convert isotope measurements between reference scales.

Usage

```
refTrans(samples, marker = "d2H", ref_scale = "VSMOW_H", niter = 5000)
```

Arguments

samples	data.frame. Must include a field with data to be transformed, analytical reproducibility of sample data (1 standard deviation), and original reference scale for calibration of data. These fields must be named <i>marker</i> , <i>marker.sd</i> , and <i>marker_cal</i> , respectively, where <i>marker</i> is “d2H” or “d18O”. Values for the cal field should correspond to <i>Calibration</i> codes found in stds tables <i>hstds</i> and <i>ostds</i> .
marker	character string. Column name for isotopic data to be extracted, either “d2H” or “d18O”.
ref_scale	character string. Text identifier for reference scale to which all isotope values will be transformed. See stds .
niter	integer. Number of random samples used to propagate uncertainty in calibration hierarchy transformations.

Value

Returns an object of class “refTrans”.

data	data.frame. Formatted identically to input object <i>samples</i> , with values for the data and data uncertainty fields replaced with transformed values.
chains	list. Each item is a character string vector containing the hierarchy of calibrations used in the transformation for a set of selected samples. See stds .

References

Magozzi, S. et al. (in press) Calibration chain transformation to improve the comparability of organic hydrogen and oxygen isotope data. *Methods in Ecology and Evolution*

Examples

```
# Some fake sample data
s = data.frame("d2H" = seq(-100, -10, by=10), "d2H.sd" = rep(2), "d2H_cal" = rep("01dUT_H_1"))

# Transform to VSMOW-SLAP scale using default arguments
d1 = refTrans(s)

# Transformed values
d1$data$d2H

# error - target scale not valid for marker
## Not run: d2 = refTrans(s, ref_scale = "VSMOW_0")
```

sr_MI

data: low resolution locally weathered Sr isoscape for Michigan

Description

Modeled $87\text{Sr}/86\text{Sr}$ value of the local rock weathering flux, obtained from waterisotopes.org and aggregated to 10 km resolution.

Usage

```
sr_MI
```

Format

SpatRaster with two layers. The first layer is the mean prediction and the second is 1 standard deviation (here estimated as 1 percent of the modeled mean)

References

Bataille, C. P. and Bowen, G. J. (2012) Mapping $87\text{Sr}/86\text{Sr}$ variations in bedrock and water for large scale provenance studies. *Chemical Geology*, **304–305**, 39–52.

Examples

```
library(terra)

plot(sr_MI)
```

states *data: outline of US states*

Description

Outline map of the of lower 48 U.S. states.

Usage

```
states
```

Format

SpatVector

Examples

```
library(terra)

plot(states)
```

stds *data: documentation of keratin H and O isotope standard calibrations*

Description

This data object contains information on keratin H and O isotope standard materials and calibrations used across multiple laboratories since the year 2000.

Usage

```
data("stds")
```

Format

list.

hstds data.frame with 18 fields.

- [, 1] Calibration: Calibration code
- [, 2] High_ID: Identifier for high-value standard
- [, 3] High_material: Description of high-value standard material
- [, 4] High: Mean hydrogen isotope value of high-value standard
- [, 5] High_sd: Standard deviation of calibration data for high-value standard
- [, 6] High_n: Number of calibration data for high-value standard
- [, 7] High_se: Standard error of the calibrated mean for high-value standard

- [, 8] Low_ID: Identifier for low-value standard
- [, 9] Low_material: Description of low-value standard material
- [, 10] Low: Mean hydrogen isotope value of low-value standard
- [, 11] Low_sd: Standard deviation of calibration data for low-value standard
- [, 12] Low_n: Number of calibration data for low-value standard
- [, 13] Low_se: Standard error of the calibrated mean for low-value standard
- [, 14] Ref_scale: Calibration scale against which the values for this calibration are anchored
- [, 15] Citation_val: Source for the calibrated values
- [, 16] Citation_cal: Source for the methodology used for this calibration
- [, 17] Treatment: Description of calibration procedure
- [, 18] H_calibration_comments: Comments

ostds data.frame with 18 fields.

- [, 1] Calibration: Calibration code
- [, 2] High_ID: Identifier for high-value standard
- [, 3] High_material: Description of high-value standard material
- [, 4] High: Mean oxygen isotope value of high-value standard
- [, 5] High_sd: Standard deviation of calibration data for high-value standard
- [, 6] High_n: Number of calibration data for high-value standard
- [, 7] High_se: Standard error of the calibrated mean for high-value standard
- [, 8] Low_ID: Identifier for low-value standard
- [, 9] Low_material: Description of low-value standard material
- [, 10] Low: Mean oxygen isotope value of low-value standard
- [, 11] Low_sd: Standard deviation of calibration data for low-value standard
- [, 12] Low_n: Number of calibration data for low-value standard
- [, 13] Low_se: Standard error of the calibrated mean for low-value standard
- [, 14] Ref_scale: Calibration scale against which the values for this calibration are anchored
- [, 15] Citation_val: Source for the calibrated values
- [, 16] Citation_cal: Source for the methodology used for this calibration
- [, 17] Treatment: Description of calibration procedure
- [, 18] O_calibration_comments: Comments

ham matrix. $n \times n$ symmetric, where n is the number of calibrations represented here and in `stds$hstds`.

oam matrix. $n \times n$ symmetric, where n is the number of calibrations represented here and in `stds$ostds`.

References

Magozzi, S. et al. (in press) Calibration chain transformation to improve the comparability of organic hydrogen and oxygen isotope data. *Methods in Ecology and Evolution*

Examples

```
library(graphics)
data("stds")

print(stds$hstds[, 1:5])
print(stds$ostds[, 1:5])
image(stds$ham)
image(stds$oam)
```

subOrigData	<i>Extract known origin sample data</i>
-------------	---

Description

This function subsets the known-origin isotope dataset included in this package and conducts optional transformations to convert isotope measurements to a common reference scale.

Usage

```
subOrigData(marker = "d2H", taxon = NULL, group = NULL, dataset = NULL,
            age_code = NULL, mask = NULL, ref_scale = "VSMOW_H", niter = 5000, genplot = TRUE)
```

Arguments

marker	character string. Column name for isotopic data to be extracted, either “d2H” or “d18O”.
taxon	character string or string vector. Species name(s) for data to be extracted.
group	character string or string vector. Taxonomic groups for data to be extracted.
dataset	integer or integer vector. Dataset_ID(s) for data to be extracted. See knownOrig feature <i>sources</i> .
age_code	character string or string vector. Animal age code for data to be extracted.
mask	SpatVector. Polygon layer used to constrain the geographic area from which data are extracted. If not provided, global.
ref_scale	character string. Text identifier for reference scale to which all isotope values will be transformed. See stds . Use “NULL” to obtain untransformed values on the originally reported scale.
niter	integer. Number of random samples used to propagate uncertainty in calibration hierarchy transformations.
genplot	logical. Plot results in R.

Value

Returns an object of class “subOrigData”, formatted for use in [calRaster](#) or [QA](#) functions.

data	SpatVector including one point feature for each selected sample. Data fields are described in knownOrig feature <i>samples</i> .
sources	data.frame. Information for all data sources for the selected samples. Fields are described in knownOrig feature <i>sources</i>
chains	list. Each item is a character string vector containing the hierarchy of calibrations used in the transformation for a set of selected samples. See stds .
marker	character string. The isotopic marker specified in the call to subOrigData

References

Magozzi, S. et al. (in press) Calibration chain transformation to improve the comparability of organic hydrogen and oxygen isotope data. *Methods in Ecology and Evolution*

Examples

```
## WITHOUT mask
# extract d2H data for Jackdaw, Partridge and Willow Grouse, transformed
# to the VSMOW/SLAP H reference scale by default
d1 = subOrigData(taxon = c("Danaus plexippus", "Setophaga ruticilla",
  "Turdus migratorius"), niter = 100)
summary(d1)

# extract d2H data for insects and passerine birds without transformation
d2 = subOrigData(group = c("Insect","Passerine"), ref_scale = NULL, genplot = FALSE)
summary(d2)

# extract d180 data for all humans, transformed to the VSMOW/SLAP 0 reference scale
d3 = subOrigData(marker = "d180",
  group = c("Modern human", "Indigenous human"), ref_scale = "VSMOW_0",
  niter = 100, genplot = FALSE)
summary(d3)

# extract d2H data for humans using taxon, transformed to the VSMOW/SLAP H reference scale
d4 = subOrigData(marker = "d2H", taxon = "Homo sapiens", ref_scale =
  "VSMOW_H", niter = 100, genplot = FALSE)
summary(d4)

## WITH mask
# error - no samples found
## Not run: d5 = subOrigData(taxon = "Turdus philomelos", mask = naMap)
# this works OK
d6 = subOrigData(taxon = c("Danaus plexippus", "Setophaga ruticilla",
  "Turdus migratorius"), mask = naMap, genplot = FALSE)
```

unionP	<i>Union probability of origin</i>
--------	------------------------------------

Description

Probabilities that at least one individual came from each location in the assignment area (union of probabilities)

Usage

```
unionP(pdR)
```

Arguments

pdR SpatRaster of probability density maps, e.g., as produced by pdRaster.

Value

SpatRaster.

Examples

```
# load hydrogen isotope data for human hair in North America
d = subOrigData(group = "Modern human", mask = naMap, genplot = FALSE)

# rescale from environmental isoscape to tissue isoscape
r = calRaster(d, d2h_lrNA, naMap, genplot = FALSE)

# four unknown-origin examples
id = c("A", "B", "C", "D")
d2H = c(-110, -90, -105, -102)
un = data.frame(id, d2H)

# posterior probabilities
pp = pdRaster(r, un, mask = naMap, genplot = FALSE)

# probability that one or more individuals are from a given location
unionP(pp)
```

wDist	<i>Probability weighted distances and bearings</i>
-------	--

Description

Calculate the distance and bearing of migration for one or more samples, weighted by probabilities from a pdRaster analysis.

Usage

```
wDist(pdR, sites, maxpts = 1e5, bw = "sj")
```

Arguments

pdR	SpatRaster of n probability density maps, e.g., as produced by pdRaster.
sites	SpatVector object containing the collection locations for the n samples represented in pdR.
maxpts	numeric. Maximum number of grid cells at which to calculate bearing and distance.
bw	character or numeric. Smoothing bandwidth to be used in kernel density estimation. See bandwidth .

Details

pdR and sites must be of equal length and corresponding order, or if `length(sites) == 1 & nlyr(pdR) > 1` then the location in sites is recycled with a message. Names in the output object are taken from the names of the layers in pdR.

Distances and bearings are calculated on the WGS84 geoid using functions from the **terra** and **geosphere** package. These calculations can take a long time for large rasters. If maxpts is less than the number of grid cells in each pdR layer, calculations are carried out for maxpts randomly selected cells.

Bearing values correspond to the initial bearing from source to collection location, and are reported on a scale of -180 to +180 degrees. The statistical metrics are rectified so that values for distributions spanning due south are reported correctly. Both weighted bearing and distance distributions are often multimodal, and it is recommended to review the distribution densities to assess the representativeness of the statistics (e.g., using [plot.wDist](#)).

When algorithmic bandwidth selection is used weights are ignored for this step and warnings to this effect are suppressed.

Value

Returns an object of class “wDist”, a list of length n. Each item contains three named objects:

stats	named number. Statistics for the distance (dist, meters) and bearing (bear, degrees) between source and collection locations, including the weighted mean (wMean) and quantile (wXX) values.
d.dens	density. Weighted kernel density for the distance between source and collection locations (meters). See density .
b.dens	density. Weighted kernel density for the bearing between source and collection locations (degrees). See density .

Examples

```
# load hydrogen isotope data for human hair in North America
d = subOrigData(group = "Modern human", mask = naMap, genplot = FALSE)
```

```
# rescale from environmental isoscape to tissue isoscape
r = calRaster(d, d2h_lrNA, naMap, genplot = FALSE)

# four unknown-origin examples
id = c("A", "B", "C", "D")
d2H = c(-110, -90, -105, -102)
un = data.frame(id, d2H)

# posterior probabilities
pp = pdRaster(r, un, mask = naMap, genplot = FALSE)

# random collection locations
sites = d$data[sample(seq(length(d$data)), 4),]

# generate a wDist object
wd = wDist(pp, sites)

# structure of the wDist object
str(wd, 2)
```

wrlld_simpl

data: world boundary map

Description

Simplified spatial polygon layer representing the boundary of global continents.

Usage

```
wrlld_simpl
```

Format

SpatVector

Examples

```
library(terra)

plot(wrlld_simpl)
```

Index

* datasets

knownOrig, 8
stds, 22

assignR, 2

bandwidth, 27

c.wDist, 3

calRaster, 4, 7, 12, 13, 25

d2h_lrNA, 5

density, 27

getIsoscapes, 6

isoStack, 7, 12, 13, 17

jointP, 8

knownOrig, 8, 24, 25

naMap, 10

oddsRatio, 11

pdRaster, 5, 12

plot.isoStack, 13

plot.QA, 14, 18

plot.wDist, 15, 27

QA, 14, 16, 25

qtlRaster, 19

refTrans, 12, 13, 20

sr_MI, 21

states, 22

stds, 9, 10, 20, 22, 24, 25

subOrigData, 4, 10, 16, 24

unionP, 26

wDist, 3, 15, 26

wrld_simpl, 28