## Package ‘simmr’

October 14, 2022

<table>
<thead>
<tr>
<th>Type</th>
<th>Package</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>A Stable Isotope Mixing Model</td>
</tr>
<tr>
<td>Version</td>
<td>0.4.5</td>
</tr>
<tr>
<td>Date</td>
<td>2021-02-28</td>
</tr>
<tr>
<td>Author</td>
<td>Andrew Parnell</td>
</tr>
<tr>
<td>BugReports</td>
<td><a href="https://github.com/andrewcparnell/simmr/issues">https://github.com/andrewcparnell/simmr/issues</a></td>
</tr>
<tr>
<td>Language</td>
<td>en-US</td>
</tr>
<tr>
<td>Maintainer</td>
<td>Andrew Parnell <a href="mailto:andrew.parnell@mu.ie">andrew.parnell@mu.ie</a></td>
</tr>
<tr>
<td>Description</td>
<td>Fits Stable Isotope Mixing Models (SIMMs) and is meant as a longer term replacement to the previous widely-used package SIAR. SIMMs are used to infer dietary proportions of organisms consuming various food sources from observations on the stable isotope values taken from the organisms' tissue samples. However SIMMs can also be used in other scenarios, such as in sediment mixing or the composition of fatty acids. The main functions are simmr_load and simmr_mcmc. The two vignettes contain a quick start and a full listing of all the features. The methods used are detailed in the papers Parnell et al 2010 <a href="">doi:10.1371/journal.pone.0009672</a>, and Parnell et al 2013 <a href="">doi:10.1002/env.2221</a>.</td>
</tr>
<tr>
<td>Depends</td>
<td>R (&gt;= 3.5.0), R2jags, ggplot2</td>
</tr>
<tr>
<td>Imports</td>
<td>MASS, compositions, boot, reshape2, graphics, stats, viridis, bayesplot, checkmate</td>
</tr>
<tr>
<td>Encoding</td>
<td>UTF-8</td>
</tr>
<tr>
<td>License</td>
<td>GPL (&gt;= 2)</td>
</tr>
<tr>
<td>LazyData</td>
<td>TRUE</td>
</tr>
<tr>
<td>Suggests</td>
<td>knitr, markdown, readxl, testthat, covr</td>
</tr>
<tr>
<td>VignetteBuilder</td>
<td>knitr</td>
</tr>
<tr>
<td>NeedsCompilation</td>
<td>no</td>
</tr>
<tr>
<td>RoxygenNote</td>
<td>7.1.1</td>
</tr>
<tr>
<td>Repository</td>
<td>CRAN</td>
</tr>
<tr>
<td>Date/Publication</td>
<td>2021-02-27 14:00:03 UTC</td>
</tr>
</tbody>
</table>
R topics documented:

combine_sources .......................................................... 2
compare_groups .............................................................. 4
compare_sources ............................................................ 6
geese_data ................................................................. 8
geese_data_day1 ............................................................. 9
plot.simmr_input .......................................................... 10
plot.simmr_output ......................................................... 12
posterior_predictive ......................................................... 14
print.simmr_input ........................................................ 15
print.simmr_output ......................................................... 16
prior_viz ................................................................. 16
simmr .................................................................. 18
simmr_data_1 ............................................................... 19
simmr_data_2 ............................................................... 20
simmr_elicit ............................................................... 20
simmr_load ............................................................... 22
simmr_mcmc ............................................................... 25
simmr_mcmc_tdf .......................................................... 30
square_data ............................................................... 33
summary.simmr_output .................................................. 34
summary.simmr_output_tdf .............................................. 36

Index 38

combine_sources  Combine the dietary proportions from two food sources after running simmr

Description

This function takes in an object of class simmr_output and combines two of the food sources. It works for single and multiple group data.

Usage

```
combine_sources(
  simmr_out,
  to_combine = simmr_out$input$source_names[1:2],
  new_source_name = "combined_source"
)
```
Arguments

- **simmr_out**: An object of class `simmr_output` created from `simmr_mcmc`.
- **to_combine**: The names of exactly two sources. These should match the names given to `simmr_load`.
- **new_source_name**: A name to give to the new combined source.

Details

Often two sources either (1) lie in similar location on the iso-space plot, or (2) are very similar in phylogenetic terms. In case (1) it is common to experience high (negative) posterior correlations between the sources. Combining them can reduce this correlation and improve precision of the estimates. In case (2) we might wish to determine the joint amount eaten of the two sources when combined. This function thus combines two sources after a run of `simmr_mcmc` (known as a posteriori combination). The new object can then be called with `plot.simmr_input` or `plot.simmr_output` to produce iso-space plots of summaries of the output after combination.

Value

A new `simmr_output` object

Author(s)

Andrew Parnell <andrew.parnell@mu.ie>

See Also

See `simmr_mcmc` and the associated vignette for examples.

Examples

```r
## Not run:
# The data
data(geese_data)

# Load into simmr
simmr_1 <- with(
geese_data_day1,
simmr_load(
mixtures = mixtures,
source_names = source_names,
source_means = source_means,
source_sds = source_sds,
correction_means = correction_means,
correction_sds = correction_sds,
concentration_means = concentration_means
)
)

# Plot
```
compare_groups

Compare dietary proportions for a single source across different groups

Description

This function takes in an object of class `simmr_output` and creates probabilistic comparisons for a given source and a set of at least two groups.

Usage

```r
compare_groups(
  simmr_out,
  source_name = simmr_out$input$source_names[1],
  groups = 1:2,
)```
compare_groups

  plot = TRUE
)

Arguments

  simmr_out An object of class simmr_output created from simmr_mcmc.
  source_name The name of a source. This should match the names exactly given to simmr_load.
  groups The integer values of the group numbers to be compared. At least two groups must be specified.
  plot A logical value specifying whether plots should be produced or not.

Details

When two groups are specified, the function produces a direct calculation of the probability that one group is bigger than the other. When more than two groups are given, the function produces a set of most likely probabilistic orderings for each combination of groups. The function produces boxplots by default and also allows for the storage of the output for further analysis if required.

Value

If there are two groups, a vector containing the differences between the two groups proportions for that source. If there are multiple groups, a list containing the following fields:

- Ordering The different possible orderings of the dietary proportions across groups
- out_all The dietary proportions for this source across the groups specified as columns in a matrix

Author(s)

Andrew Parnell <andrew.parnell@mu.ie>

See Also

See simmr_mcmc for complete examples.

Examples

```r
## Not run:
data(geese_data)
simmr_in <- with(
geese_data,
simmr_load(
immutes = mixtures,
source_names = source_names,
source_means = source_means,
source_sds = source_sds,
correction_means = correction_means,
correction_sds = correction_sds,
concentration_means = concentration_means,
group = groups
```
compare_sources

Compare dietary proportions between multiple sources

Description

This function takes in an object of class simmr_output and creates probabilistic comparisons between the supplied sources. The group number can also be specified.

Usage

```r
compare_sources(
  simmr_out,
  source_names = simmr_out$input$source_names,
  group = 1,
  source = "Zostera", groups = 1:2)
```
**compare_sources**

```r
plot = TRUE
```

### Arguments

- **simmr_out** An object of class `simmr_output` created from `simmr_mcmc`.
- **source_names** The names of at least two sources. These should match the names exactly given to `simmr_load`.
- **group** The integer values of the group numbers to be compared. If not specified assumes the first or only group.
- **plot** A logical value specifying whether plots should be produced or not.

### Details

When two sources are specified, the function produces a direct calculation of the probability that the dietary proportion for one source is bigger than the other. When more than two sources are given, the function produces a set of most likely probabilistic orderings for each combination of sources. The function produces boxplots by default and also allows for the storage of the output for further analysis if required.

### Value

If there are two sources, a vector containing the differences between the two dietary proportion proportions for these two sources. If there are multiple sources, a list containing the following fields:

- **Ordering** The different possible orderings of the dietary proportions across sources
- **out_all** The dietary proportions for these sources specified as columns in a matrix

### Author(s)

Andrew Parnell <andrew.parnell@mu.ie>

### See Also

See `simmr_mcmc` for complete examples.

### Examples

```r
## Not run:
data(geese_data_day1)
simmr_1 <- with(
  geese_data_day1,
simmr_load(
    mixtures = mixtures,
    source_names = source_names,
    source_means = source_means,
    source_sds = source_sds,
    correction_means = correction_means,
    correction_sds = correction_sds,
```
geese_data

Geese stable isotope mixing data set

Description

A real Geese data set with 251 observations on 2 isotopes, with 4 sources, and with corrections/trophic enrichment factors (TEFs or TDFs), and concentration dependence means. Taken from Inger et al (2016). See link for paper

Usage

geese_data
**geese_data_day1**

**Format**

A list with the following elements

- `mixtures` A two column matrix containing delta 13C and delta 15N values respectively
- `source_names` A character vector of the food source names
- `tracer_names` A character vector of the tracer names (d13C, d15N, d34S)
- `source_means` A matrix of source mean values for the tracers in the same order as `mixtures` above
- `source_sds` A matrix of source sd values for the tracers in the same order as `mixtures` above
- `correction_means` A matrix of TEFs mean values for the tracers in the same order as `mixtures` above
- `correction_sds` A matrix of TEFs sd values for the tracers in the same order as `mixtures` above
- `concentration_means` A matrix of concentration dependence mean values for the tracers in the same order as `mixtures` above...

@seealso `simmr_mcmc` for an example where it is used

**Source**


**Description**

A real Geese data set with 9 observations on 2 isotopes, with 4 sources, and with corrections/trophic enrichment factors (TEFs or TDFs), and concentration dependence means. Taken from Inger et al (2016). See link for paper

**Usage**

`geese_data_day1`

**Format**

A list with the following elements

- `mixtures` A two column matrix containing delta 13C and delta 15N values respectively
- `source_names` A character vector of the food source names
- `tracer_names` A character vector of the tracer names (d13C, d15N, d34S)
- `source_means` A matrix of source mean values for the tracers in the same order as `mixtures` above
- `source_sds` A matrix of source sd values for the tracers in the same order as `mixtures` above
- `correction_means` A matrix of TEFs mean values for the tracers in the same order as `mixtures` above
correction_sds  A matrix of TEFs sd values for the tracers in the same order as mixtures above
concentration_means A matrix of concentration dependence mean values for the tracers in the same order as mixtures above ...
@seealso simmr_mcmc for an example where it is used

Source

---

**plot.simmr_input**  
*Plot the simmr_input data created from simmr_load*

---

### Description
This function creates iso-space (AKA tracer-space or delta-space) plots. They are vital in determining whether the data are suitable for running in a SIMM.

### Usage
```r
## S3 method for class 'simmr_input'
plot(
x,
  tracers = c(1, 2),
  title = "Tracers plot",
  xlab = colnames(x$mixtures)[tracers[1]],
  ylab = colnames(x$mixtures)[tracers[2]],
  sigmas = 1,
  group = 1:x$n_groups,
  mix_name = "Mixtures",
  ggargs = NULL,
  colour = TRUE,
  ...;
)
```

### Arguments
- **x**  
  An object created via the function `simmr_load`
- **tracers**  
  The choice of tracers to plot. If there are more than two tracers, it is recommended to plot every pair of tracers to determine whether the mixtures lie in the mixing polygon defined by the sources
- **title**  
  A title for the graph
- **xlab**  
  The x-axis label. By default this is assumed to be delta-13C but can be made richer if required. See examples below.
- **ylab**  
  The y-axis label. By default this is assumed to be delta-15N in per mil but can be changed as with the x-axis label
**plot.simmr_input**

sigmas  
The number of standard deviations to plot on the source values. Defaults to 1.
group  
Which groups to plot. Can be a single group or multiple groups
mix_name  
A optional string containing the name of the mixture objects, e.g. Geese.
ggargs  
Extra arguments to be included in the ggplot (e.g. axis limits)
colour  
If TRUE (default) creates a plot. If not, puts the plot in black and white
...  
Not used

**Details**

It is desirable to have the vast majority of the mixture observations to be inside the convex hull defined by the food sources. When there are more than two tracers (as in one of the examples below) it is recommended to plot all the different pairs of the food sources. See the vignette for further details of richer plots.

**Author(s)**

Andrew Parnell <andrew.parnell@mu.ie>

**See Also**

See `plot.simmr_output` for plotting the output of a simmr run. See `simmr_mcmc` for running a simmr object once the iso-space is deemed acceptable.

**Examples**

```r
# A simple example with 10 observations, 4 food sources and 2 tracers
data(geese_data_day1)
simmr_1 <- with(
  geese_data_day1,
  simmr_load(
    mixtures = mixtures,
    source_names = source_names,
    source_means = source_means,
    source_sds = source_sds,
    correction_means = correction_means,
    correction_sds = correction_sds,
    concentration_means = concentration_means
  )
)

# Plot
plot(simmr_1)

### A more complicated example with 30 obs, 3 tracers and 4 sources
data(simmr_data_2)
simmr_3 <- with(
  simmr_data_2,
  simmr_load(
    mixtures = mixtures,
  )
)```
source_names = source_names,
source_means = source_means,
source_sds = source_sds,
correction_means = correction_means,
correction_sds = correction_sds,
concentration_means = concentration_means
)
)

# Plot 3 times - first default d13C vs d15N
plot(simmr_3)
# Now plot d15N vs d34S
plot(simmr_3, tracers = c(2, 3))
# and finally d13C vs d34S
plot(simmr_3, tracers = c(1, 3))
# See vignette('simmr') for fancier x-axis labels

# An example with multiple groups - the Geese data from Inger et al 2006
data(geese_data)
simmr_4 <- with(geese_data,
  simmr_load(
    mixtures = mixtures,
    source_names = source_names,
    source_means = source_means,
    source_sds = source_sds,
    correction_means = correction_means,
    correction_sds = correction_sds,
    concentration_means = concentration_means,
    group = groups
  )
)

# Print
simmr_4

# Plot
plot(simmr_4,
  xlab = expression(paste(delta^13, " C (\u2030)", sep = "")),
  ylab = expression(paste(delta^15, " N (\u2030)", sep = "")),
  title = "Isospace plot of Inger et al Geese data"
)="#'

---

**plot.simmr_output**  Plot different features of an object created from `simmr_mcmc`.

**Description**

This function allows for 4 different types of plots of the simmr output created from `simmr_mcmc`. The types are: histogram, kernel density plot, matrix plot (most useful) and boxplot. There are some minor customisation options.
Usage

```r
## S3 method for class 'simmr_output'
plot(
  x,
  type = c("isospace", "histogram", "density", "matrix", "boxplot"),
  group = 1,
  binwidth = 0.05,
  alpha = 0.5,
  title = if (length(group) == 1) { "simmr output plot" } else {
    paste("simmr output plot: group", group) },
  ggargs = NULL,
  ...
)
```

Arguments

- **x**: An object of class `simmr_output` created via `simmr_mcmc`
- **type**: The type of plot required. Can be one or more of `histogram`, `density`, `matrix`, or `boxplot`
- **group**: Which group(s) to plot.
- **binwidth**: The width of the bins for the histogram. Defaults to 0.05
- **alpha**: The degree of transparency of the plots. Not relevant for matrix plots
- **title**: The title of the plot.
- **ggargs**: Extra arguments to be included in the ggplot (e.g. axis limits)
- **...**: Currently not used

Details

The matrix plot should form a necessary part of any SIMM analysis since it allows the user to judge which sources are identifiable by the model. Further detail about these plots is provided in the vignette.

Author(s)

Andrew Parnell <andrew.parnell@mu.ie>

See Also

See `simmr_mcmc` for creating objects suitable for this function, and many more examples. See also `simmr_load` for creating simmr objects, `plot.simmr_input` for creating isospace plots, `summary.simmr_output` for summarising output.

Examples

```r
## Not run:
# A simple example with 10 observations, 2 tracers and 4 sources
```
# The data
data(geese_data)

# Load into simmr
simmr_1 <- with(geese_data_day1, simmr_load(mixtures = mixtures, source_names = source_names, source_means = source_means, source_sds = source_sds, correction_means = correction_means, correction_sds = correction_sds, concentration_means = concentration_means
)

# Plot
plot(simmr_1)

# MCMC run
simmr_1_out <- simmr_mcmc(simmr_1)

# Plot
plot(simmr_1_out) # Creates all 4 plots
plot(simmr_1_out, type = "boxplot")
plot(simmr_1_out, type = "histogram")
plot(simmr_1_out, type = "density")
plot(simmr_1_out, type = "matrix")

## End(Not run)

---

posterior_predictive  
*Plot the posterior predictive distribution for a simmr run*

### Description

This function takes the output from `simmr_mcmc` and plots the posterior predictive distribution to enable visualisation of model fit. The simulated posterior predicted values are returned as part of the object and can be saved for external use.

### Usage

`posterior_predictive(simmr_out, group = 1, prob = 0.5, plot_ppc = TRUE)`

### Arguments

- `simmr_out`  
  A run of the simmr model from `simmr_mcmc`

- `group`  
  Which group to run it for (currently only numeric rather than group names)
**print.simmr_input**

The probability interval for the posterior predictives. The default is 0.5 (i.e. 50pc intervals)

**plot_ppc**

Whether to create a bayesplot of the posterior predictive or not.

**Examples**

```r
## Not run:
data(geese_data_day1)
simmr_1 <- with(
geese_data_day1,
simmr_load(
  mixtures = mixtures,
  source_names = source_names,
  source_means = source_means,
  source_sds = source_sds,
  correction_means = correction_means,
  correction_sds = correction_sds,
  concentration_means = concentration_means
)
)

# Plot
plot(simmr_1)

# Print
simmr_1

# MCMC run
simmr_1_out <- simmr_mcmc(simmr_1)

# Prior predictive
post_pred <- posterior_predictive(simmr_1_out)

## End(Not run)
```

---

**Description**

Print simmr input object

**Usage**

```r
## S3 method for class 'simmr_input'
print(x, ...)
```

**Arguments**

- `x` An object of class simmr_input
- `...` Other arguments (not supported)
print.simmr_output  

Description:
A neat presentation of your simmr object.

## S3 method for class 'simmr_output'
print(x, ...)

Arguments:
- x: An object of class simmr_output
- ...: Other arguments (not supported)

Value:
Returns a neat summary of the object

See Also:
simmr_mcmc for creating simmr_output objects

prior_viz  

Description:
This function takes the output from simmr_mcmc and plots the prior distribution to enable visual inspection. This can be used by itself or as part of posterior_predictive to visually evaluate the influence of the prior on the posterior distribution.

Usage:
prior_viz(
  simmr_out,
  group = 1,
  plot = TRUE,
  include_posterior = TRUE,
  n_sims = 10000,
  ggargs = NULL
)
prior_viz

Arguments

- **simmr_out**: A run of the simmr model from `simmr_mcmc`
- **group**: Which group to run it for (currently only numeric rather than group names)
- **plot**: Whether to create a density plot of the prior or not. The simulated prior values are returned as part of the object
- **include_posterior**: Whether to include the posterior distribution on top of the priors. Defaults to TRUE
- **n_sims**: The number of simulations from the prior distribution
- **ggargs**: Extra arguments to be included in the ggplot (e.g. axis limits)

Examples

```r
## Not run:
data(geese_data_day1)
simmr_1 <- with(
  geese_data_day1,
  simmr_load(
    mixtures = mixtures,
    source_names = source_names,
    source_means = source_means,
    source_sds = source_sds,
    correction_means = correction_means,
    correction_sds = correction_sds,
    concentration_means = concentration_means
  )
)

# Plot
plot(simmr_1)

# Print
simmr_1

# MCMC run
simmr_1_out <- simmr_mcmc(simmr_1)

# Prior predictive
prior <- prior_viz(simmr_1_out)
head(prior)
summary(prior)

## End(Not run)
```
Description

This package runs a simple Stable Isotope Mixing Model (SIMM) and is meant as a longer term replacement to the previous function SIAR. These are used to infer dietary proportions of organisms consuming various food sources from observations on the stable isotope values taken from the organisms’ tissue samples. However SIMMs can also be used in other scenarios, such as in sediment mixing or the composition of fatty acids. The main functions are `simmr_load` and `simmr_mcmc`. The help files contain examples of the use of this package. See also the vignette for a longer walkthrough.

Details

An even longer term replacement for properly running SIMMs is MixSIAR, which allows for more detailed random effects and the inclusion of covariates.

Author(s)

Andrew Parnell <andrew.parnell@mu.ie>

References


Examples

```r
## Not run:
# A first example with 2 tracers (isotopes), 10 observations, and 4 food sources
data(geese_data_day1)
simmr_in <- with(
geese_data_day1,
simmr_load(
  mixtures = mixtures,
  source_names = source_names,
  source_means = source_means,
  source_sds = source_sds,
  correction_means = correction_means,
  correction_sds = correction_sds,
  concentration_means = concentration_means
)
)
```
# Plot
plot(simmr_in)

# MCMC run
simmr_out <- simmr_mcmc(simmr_in)

# Check convergence - values should all be close to 1
summary(simmr_out, type = "diagnostics")

# Look at output
summary(simmr_out, type = "statistics")

# Look at influence of priors
prior_viz(simmr_out)

# Plot output
plot(simmr_out, type = "histogram")

## End(Not run)

---

**simmr_data_1**

**A simple fake stable isotope mixing data set**

**Description**

A simple fake data set with 10 observations on 2 isotopes, with 4 sources, and with corrections/trophic enrichment factors (TEFs or TDFs), and concentration dependence means

**Usage**

simmr_data_1

**Format**

A list with the following elements

- **mixtures**: A two column matrix containing delta 13C and delta 15N values respectively
- **source_names**: A character vector of the food source names
- **tracer_names**: A character vector of the tracer names (d13C and d15N)
- **source_means**: A matrix of source mean values for the tracers in the same order as `mixtures` above
- **source_sds**: A matrix of source sd values for the tracers in the same order as `mixtures` above
- **correction_means**: A matrix of TEFs mean values for the tracers in the same order as `mixtures` above
- **correction_sds**: A matrix of TEFs sd values for the tracers in the same order as `mixtures` above
- **concentration_means**: A matrix of concentration dependence mean values for the tracers in the same order as `mixtures` above...

@seealso *simmr_mcmc* for an example where it is used
simmr_data_2  A 3-isotope fake stable isotope mixing data set

Description

A fake data set with 30 observations on 3 isotopes, with 4 sources, and with corrections/trophic enrichment factors (TEFs or TDFs), and concentration dependence means.

Usage

simmr_data_2

Format

A list with the following elements

- **mixtures**  A three column matrix containing delta 13C, delta 15N, and delta 34S values respectively
- **source_names**  A character vector of the food source names
- **tracer_names**  A character vector of the tracer names (d13C, d15N, d34S)
- **source_means**  A matrix of source mean values for the tracers in the same order as `mixtures` above
- **source_sds**  A matrix of source sd values for the tracers in the same order as `mixtures` above
- **correction_means**  A matrix of TEFs mean values for the tracers in the same order as `mixtures` above
- **correction_sds**  A matrix of TEFs sd values for the tracers in the same order as `mixtures` above
- **concentration_means**  A matrix of concentration dependence mean values for the tracers in the same order as `mixtures` above...

@seealso simmr_mcmc for an example where it is used

simmr_elicit  Function to allow informative prior distribution to be included in simmr

Description

The main `simmr_mcmc` function allows for a prior distribution to be set for the dietary proportions. The prior distribution is specified by transforming the dietary proportions using the centralised log ratio (CLR). The `simmr_elicit` and `simmr_elicit` functions allows the user to specify prior means and standard deviations for each of the dietary proportions, and then finds CLR-transformed values suitable for input into `simmr_mcmc`. 
Usage

`simmr_elicit`

```r
simmr_elicit(
  n_sources,
  proportion_means = rep(1/n_sources, n_sources),
  proportion_sds = rep(0.1, n_sources),
  n_sims = 1000
)
```

Arguments

- `n_sources`: The number of sources required
- `proportion_means`: The desired prior proportion means. These should sum to 1. Should be a vector of length `n_sources`
- `proportion_sds`: The desired prior proportions standard deviations. These have no restricted sum but should be reasonable estimates for a proportion.
- `n_sims`: The number of simulations for which to run the optimisation routine.

Details

The function takes the desired proportion means and standard deviations, and fits an optimised least squares to the means and standard deviations in turn to produced CLR-transformed estimates for use in `simmr_mcmc`. Using prior information in SIMMs is highly desirable given the restricted nature of the inference. The prior information might come from previous studies, other experiments, or other observations of e.g. animal behaviour.

Due to the nature of the restricted space over which the dietary proportions can span, and the fact that this function uses numerical optimisation, the procedure will not match the target dietary proportion means and standard deviations exactly. If this problem is severe, try increasing the `n_sims` value.

Value

A list object with two components

- `mean`: The best estimates of the mean to use in `control.prior` in `simmr_mcmc`
- `sd`: The best estimates of the standard deviations to use in `control.prior` in `simmr_mcmc`

Author(s)

Andrew Parnell <andrew.parnell@mu.ie>

Examples

```r
## Not run:
# Data set: 10 observations, 2 tracers, 4 sources
data(geese_data_day1)
simmr_1 <- with(
  geese_data_day1,
```
```r
simmr_load(
  mixtures = mixtures,
  source_names = source_names,
  source_means = source_means,
  source_sds = source_sds,
  correction_means = correction_means,
  correction_sds = correction_sds,
  concentration_means = concentration_means
)

# MCMC run
simmr_1_out <- simmr_mcmc(simmr_1)

# Look at the prior influence
prior_viz(simmr_1_out)

# Summary
summary(simmr_1_out, "quantiles")
# A bit vague:
# 2.5%  25%  50%  75%  97.5%
# Source A 0.029 0.115 0.203 0.312 0.498
# Source B 0.146 0.232 0.284 0.338 0.453
# Source C 0.216 0.255 0.275 0.296 0.342
# Source D 0.032 0.123 0.205 0.299 0.465

# Now suppose I had prior information that:
# proportion means = 0.5,0.2,0.2,0.1
# proportion sds = 0.08,0.02,0.01,0.02
prior <- simmr_elicit(4, c(0.5, 0.2, 0.2, 0.1), c(0.08, 0.02, 0.01, 0.02))

simmr_1a_out <- simmr_mcmc(simmr_1, prior_control = list(means = prior$mean, sd = prior$sd))

#' # Look at the prior influence now
prior_viz(simmr_1a_out)

summary(simmr_1a_out, "quantiles")
# Much more precise:
# 2.5%  25%  50%  75%  97.5%
# Source A 0.441 0.494 0.523 0.553 0.610
# Source B 0.144 0.173 0.188 0.204 0.236
# Source C 0.160 0.183 0.196 0.207 0.228
# Source D 0.060 0.079 0.091 0.105 0.135

## End(Not run)
```

---

**simmr_load**  
*Function to load in simmr data and check for errors*
**Description**

This function takes in the mixture data, food source means and standard deviations, and (optionally) correction factor means and standard deviations, and concentration proportions. It performs some (non-exhaustive) checking of the data to make sure it will run through simmr. It outputs an object of class `simmr_input`.

**Usage**

```r
simmr_load(
mixtures,
source_names,
source_means,
source_sds,
correction_means = NULL,
correction_sds = NULL,
concentration_means = NULL,
group = NULL
)
```

**Arguments**

- **mixtures**: The mixture data given as a matrix where the number of rows is the number of observations and the number of columns is the number of tracers (usually isotopes)
- **source_names**: The names of the sources given as a character string
- **source_means**: The means of the source values, given as a matrix where the number of rows is the number of sources and the number of columns is the number of tracers
- **source_sds**: The standard deviations of the source values, given as a matrix where the number of rows is the number of sources and the number of columns is the number of tracers
- **correction_means**: The means of the correction values, given as a matrix where the number of rows is the number of sources and the number of columns is the number of tracers. If not provided these are set to 0.
- **correction_sds**: The standard deviations of the correction values, given as a matrix where the number of rows is the number of sources and the number of columns is the number of tracers. If not provided these are set to 0.
- **concentration_means**: The means of the concentration values, given as a matrix where the number of rows is the number of sources and the number of columns is the number of tracers. These should be between 0 and 1. If not provided these are all set to 1.
- **group**: A grouping variable. These can be a character or factor variable

**Details**

For standard stable isotope mixture modelling, the mixture matrix will contain a row for each individual and a column for each isotopic value. `simmr` will allow for any number of isotopes and any
number of observations, within computational limits. The source means/sds should be provided for each food source on each isotope. The correction means (usually trophic enrichment factors) can be set as zero if required, and should be of the same shape as the source values. The concentration dependence means should be estimated values of the proportion of each element in the food source in question and should be given in proportion format between 0 and 1. At present there is no means to include concentration standard deviations.

Value

An object of class simmr_input with the following elements:

- `mixtures` The mixture data
- `source_means` Source means
- `source_sds` Source standard deviations
- `correction_means` Correction means
- `correction_sds` Correction standard deviations
- `concentration_means` Concentration dependence means
- `n_obs` The number of observations
- `n_tracers` The number of tracers/isotopes
- `n_sources` The number of sources
- `n_groups` The number of groups

Author(s)

Andrew Parnell <andrew.parnell@mu.ie>

See Also

See `simmr_mcmc` for complete examples.

Examples

```r
# A simple example with 10 observations, 2 tracers and 4 sources
# data(geese_data_day1)
simmr_1 <- with(
  geese_data_day1, 
simmr_load(
    mixtures = mixtures,
    source_names = source_names,
    source_means = source_means,
    source_sds = source_sds,
    correction_means = correction_means,
    correction_sds = correction_sds,
    concentration_means = concentration_means
  )
)```


**simmr_mcmc**

Run a simmr_input object through the main simmr Markov chain Monte Carlo (MCMC) function

---

**Description**

This is the main function of simmr. It takes a simmr_input object created via simmr_load, runs an MCMC to determine the dietary proportions, and then outputs a simmr_output object for further analysis and plotting via summary.simmr_output and plot.simmr_output.

**Usage**

```r
simmr_mcmc(
  simmr_in,
  prior_control = list(means = rep(0, simmr_in$n_sources), sd = rep(1, simmr_in$n_sources)),
  mcmc_control = list(iter = 10000, burn = 1000, thin = 10, n.chain = 4)
)
```

**Arguments**

- `simmr_in`: An object created via the function simmr_load
- `prior_control`: A list of values including arguments named means and sd which represent the prior means and standard deviations of the dietary proportions in centralised log-ratio space. These can usually be left at their default values unless you wish to include to include prior information, in which case you should use the function simmr_elicit.
- `mcmc_control`: A list of values including arguments named iter (number of iterations), burn (size of burn-in), thin (amount of thinning), and n.chain (number of MCMC chains).

**Details**

If, after running simmr_mcmc the convergence diagnostics in summary.simmr_output are not satisfactory, the values of iter, burn and thin in mcmc_control should be increased by a factor of 10.

**Value**

An object of class simmr_output with two named top-level components:

- `input`: The simmr_input object given to the simmr_mcmc function
- `output`: A set of MCMC chains of class mcmc.list from the coda package. These can be analysed using the summary.simmr_output and plot.simmr_output functions.
Author(s)
Andrew Parnell <andrew.parnell@mu.ie>

References

See Also
simmr_load for creating objects suitable for this function, plot.simmr_input for creating isospace plots, summary.simmr_output for summarising output, and plot.simmr_output for plotting output.

Examples
## Not run:
## See the package vignette for a detailed run through of these 4 examples

# Data set 1: 10 obs on 2 isos, 4 sources, with tefs and concdep
data(geese_data_day1)
simmr_1 <- with(
  geese_data_day1,
  simmr_load(
    mixtures = mixtures,
    source_names = source_names,
    source_means = source_means,
    source_sds = source_sds,
    correction_means = correction_means,
    correction_sds = correction_sds,
    concentration_means = concentration_means
  )
)

# Plot
plot(simmr_1)

# Print
simmr_1

# MCMC run
simmr_1_out <- simmr_mcmc(simmr_1)

# Print it
print(simmr_1_out)

# Summary
summary(simmr_1_out, type = "diagnostics")
summary(simmr_1_out, type = "correlations")
summary(simmr_1_out, type = "statistics")
ans <- summary(simmr_1_out, type = c("quantiles", "statistics"))

# Plot
plot(simmr_1_out, type = "boxplot")
plot(simmr_1_out, type = "histogram")
plot(simmr_1_out, type = "density")
plot(simmr_1_out, type = "matrix")

# Compare two sources
compare_sources(simmr_1_out, source_names = c("Zostera", "Enteromorpha"))

# Compare multiple sources
compare_sources(simmr_1_out)

# A version with just one observation
data(geese_data_day1)
simmr_2 <- with(
  geese_data_day1,
  simmr_load(
    mixtures = mixtures[1,, drop = FALSE],
    source_names = source_names,
    source_means = source_means,
    source_sds = source_sds,
    correction_means = correction_means,
    correction_sds = correction_sds,
    concentration_means = concentration_means
  )
)

# MCMC run - automatically detects the single observation
simmr_2_out <- simmr_mcmc(simmr_2)

# Print it
print(simmr_2_out)

# Summary
summary(simmr_2_out)
summary(simmr_2_out, type = "diagnostics")
ans <- summary(simmr_2_out, type = c("quantiles"))

# Plot
plot(simmr_2_out)
plot(simmr_2_out, type = "boxplot")
plot(simmr_2_out, type = "histogram")
plot(simmr_2_out, type = "density")
plot(simmr_2_out, type = "matrix")
# Data set 2: 3 isotopes (d13C, d15N and d34S), 30 observations, 4 sources
data(simmr_data_2)
simmr_3 <- with(simmr_data_2,
  simmr_load(
    mixtures = mixtures,
    source_names = source_names,
    source_means = source_means,
    source_sds = source_sds,
    correction_means = correction_means,
    correction_sds = correction_sds,
    concentration_means = concentration_means
  )
)

# Get summary
print(simmr_3)

# Plot 3 times
plot(simmr_3)
plot(simmr_3, tracers = c(2, 3))
plot(simmr_3, tracers = c(1, 3))

# See vignette('simmr') for fancier axis labels

# MCMC run
simmr_3_out <- simmr_mcmc(simmr_3)

# Print it
print(simmr_3_out)

# Summary
summary(simmr_3_out)
summary(simmr_3_out, type = "diagnostics")
summary(simmr_3_out, type = "quantiles")
summary(simmr_3_out, type = "correlations")

# Plot
plot(simmr_3_out)
plot(simmr_3_out, type = "boxplot")
plot(simmr_3_out, type = "histogram")
plot(simmr_3_out, type = "density")
plot(simmr_3_out, type = "matrix")

# Data set 4 - identified by Fry (2014) as a failing of SIMMs
# See the vignette for more interpretation of these data and the output

# The data
data(square_data)
```r
simmr_4 <- with(
    square_data,
    simmr_load(
        mixtures = mixtures,
        source_names = source_names,
        source_means = source_means,
        source_sds = source_sds
    )
)

# Get summary
print(simmr_4)

# Plot
plot(simmr_4)

# MCMC run - needs slightly longer
simmr_4_out <- simmr_mcmc(simmr_4)

# Print it
print(simmr_4_out)

# Summary
summary(simmr_4_out)
ans <- summary(simmr_4_out, type = c("quantiles", "statistics"))

# Plot
plot(simmr_4_out)
plot(simmr_4_out, type = "boxplot")
plot(simmr_4_out, type = "histogram")
plot(simmr_4_out, type = "density")
plot(simmr_4_out, type = "matrix") # Look at the massive correlations here

# Data set 5 - Multiple groups Geese data from Inger et al 2006
# Do this in raw data format - Note that there's quite a few mixtures!
data(geese_data)
simmr_5 <- with(
    geese_data,
    simmr_load(
        mixtures = mixtures,
        source_names = source_names,
        source_means = source_means,
        source_sds = source_sds,
        correction_means = correction_means,
        correction_sds = correction_sds,
        concentration_means = concentration_means,
        group = groups
    )
)
```
# Plot
plot(simmr_5,
     xlab = expression(paste(delta^13, " C (\u2030)", sep = "")),
     ylab = expression(paste(delta^15, " N (\u2030)", sep = "")),
     title = "Isospace plot of Inger et al Geese data"
)

# Run MCMC for each group
simmr_5_out <- simmr_mcmc(simmr_5)

# Summarise output
summary(simmr_5_out, type = "quantiles", group = 1)
summary(simmr_5_out, type = "quantiles", group = c(1, 3))
summary(simmr_5_out, type = c("quantiles", "statistics"), group = c(1, 3))

# Plot - only a single group allowed
plot(simmr_5_out, type = "boxplot", group = 2, title = "simmr output group 2")
plot(simmr_5_out, type = c("density", "matrix"), grp = 6, title = "simmr output group 6")

# Compare sources within a group
compare_sources(simmr_5_out, source_names = c("Zostera", "U.lactuca"), group = 2)
compare_sources(simmr_5_out, group = 2)

# Compare between groups
compare_groups(simmr_5_out, source = "Zostera", groups = 1:2)
compare_groups(simmr_5_out, source = "Zostera", groups = 1:3)
compare_groups(simmr_5_out, source = "U.lactuca", groups = c(4:5, 7, 2))

## End(Not run)

---

`simmr_mcmc_tdf`  

Estimate correction factors from stable isotope data with known dietary proportions

### Description

This function runs a slightly different version of the main `simmr_mcmc` function with the key difference that it estimates the correction factors (sometimes called trophic enrichment or trophic discrimination factors; TEFs/TDFs) for a given set of dietary proportions.

### Usage

```r
simmr_mcmc_tdf(
    simmr_in,
    p = matrix(rep(1/simmr_in$n_sources, simmr_in$n_sources), ncol = simmr_in$n_sources, nrow = simmr_in$n_obs, byrow = TRUE),
    prior_control = list(c_mean_est = rep(2, simmr_in$n_tracers), c_sd_est = rep(2,
```
simmr_mcmc_tdf

simmr_in$n_tracers)),
mcmc_control = list(iter = 10000, burn = 1000, thin = 10, n.chain = 4)
)

Arguments

simmr_in An object created via the function simmr_load
p The known dietary proportions for the feeding study. Dietary proportions should be given per individual (even if they are all identical)

prior_control A list of values including arguments named means and sd which represent the prior means and standard deviations of the correction factors. These can usually be left at their default values unless you wish to include to include prior information on them.
mcmc_control A list of values including arguments named iter (number of iterations), burn (size of burn-in), thin (amount of thinning), and n.chain (number of MCMC chains).

Details

The idea is that this code can be used for feeding studies where an organism is fed a known proportional diet with a view to estimating the correction factors to be used in a later stable isotope mixing model when the organisms are observed in the field.

The main argument of the function is an object created from simmr_load which contains mixture data on a number of tracers and food source means and standard deviations. Any correction factors included in this object will be ignored. The known dietary proportions should be provided for each individual (i.e. should be a matrix with the same number of rows as mix). It is advisable to have multiple different dietary proportion values as part of the feeding experimental design.

The output of the function is a posterior distribution on the correction factors for each food source. Just like the output from simmr_mcmc, this should be checked for convergence. Examples are included below to help assist with this check and further plots.

If, after running simmr_mcmc_tdf the convergence diagnostics in summary.simmr_output_tdf are not satisfactory, the values of iter, burn and thin in mcmc_control should be increased by e.g. a factor of 10.

Value

An object of class simmr_tdf with two named top-level components:

input The simmr_input object given to the simmr_mcmc function
output A set of MCMC chains of class mcmc.list from the coda package. These can be analysed using summary.simmr_output_tdf

Author(s)

Andrew Parnell <andrew.parnell@mu.ie>
References


See Also

`simmr_load` for creating objects suitable for this function, `simmr_mcmc` for estimating dietary proportions, `plot.simmr_input` for creating isospace plots, `summary.simmr_output_tdf` for summarising output.

Examples

```r
## Not run:
## Example of estimating TDFs for a simple system with known dietary proportions

# Data set 1: 10 obs on 2 isos, 4 sources, with tefs and concdepm
# Assume p = c(0.25, 0.25, 0.25, 0.25)

# The data
data(simmr_data_1)
# Load into simmr
simmr_tdf <- with(simmr_data_1, simmr_load(
  mixtures = mixtures,
  source_names = source_names,
  source_means = source_means,
  source_sds = source_sds,
  correction_means = correction_means,
  correction_sds = correction_sds,
  concentration_means = concentration_means
))

# Plot
plot(simmr_tdf)

# MCMC run
simmr_tdf_out <- simmr_mcmc_tdf(simmr_tdf, p = matrix(rep(1 / simmr_tdf$n_sources, simmr_tdf$n_sources), ncol = simmr_tdf$n_sources, nrow = simmr_tdf$n_obs, byrow = TRUE)
```
# Summary
summary(simmr_tdf_out, type = "diagnostics")
summary(simmr_tdf_out, type = "quantiles")

# Now put these corrections back into the model and check the
# iso-space plots and dietary output
simmr_tdf_2 <- with(
  simmr_data_1,
  simmr_load(
    mixtures = mixtures,
    source_names = source_names,
    source_means = source_means,
    source_sds = source_sds,
    correction_means = simmr_tdf_out$c_mean_est,
    correction_sds = simmr_tdf_out$c_sd_est,
    concentration_means = concentration_means
  )
)

# Plot with corrections now
plot(simmr_tdf_2)

simmr_tdf_2_out <- simmr_mcmc(simmr_tdf_2)
summary(simmr_tdf_2_out, type = "diagnostics")
plot(simmr_tdf_2_out, type = "boxplot")

## End(Not run)

---

**square_data**

An artificial data set used to indicate effect of priors

**Description**

A fake box data set identified by Fry (2014) as a failing of SIMMs See the link for more interpretation of these data and the output

**Usage**

`square_data`

**Format**

A list with the following elements

- **mixtures** A two column matrix containing delta 13C and delta 15N values respectively
- **source_names** A character vector of the food source names
- **tracer_names** A character vector of the tracer names (d13C, d15N)
source_means  A matrix of source mean values for the tracers in the same order as mixtures above
source_sds   A matrix of source sd values for the tracers in the same order as mixtures above
correction_means  A matrix of TEFs mean values for the tracers in the same order as mixtures above
correction_sds  A matrix of TEFs sd values for the tracers in the same order as mixtures above
concentration_means  A matrix of concentration dependence mean values for the tracers in the same order as mixtures above ...
               @seealso simmr_mcmc for an example where it is used

Source


summary.simmr_output  Summarises the output created with simmr_mcmc

Description

Produces textual summaries and convergence diagnostics for an object created with simmr_mcmc. The different options are: 'diagnostics' which produces Brooks-Gelman-Rubin diagnostics to assess MCMC convergence, 'quantiles' which produces credible intervals for the parameters, 'statistics' which produces means and standard deviations, and 'correlations' which produces correlations between the parameters.

Usage

## S3 method for class 'simmr_output'
summary(
  object,
  type = c("diagnostics", "quantiles", "statistics", "correlations"),
  group = 1,
  ... )

Arguments

object  An object of class simmr_output produced by the function simmr_mcmc

type    The type of output required. At least none of 'diagnostics', 'quantiles', 'statistics', or 'correlations'.
group   Which group or groups the output is required for.
...     Not used
Details

The quantile output allows easy calculation of 95 per cent credible intervals of the posterior dietary proportions. The correlations, along with the matrix plot in `plot.simmr_output` allow the user to judge which sources are non-identifiable. The Gelman diagnostic values should be close to 1 to ensure satisfactory convergence.

When multiple groups are included, the output automatically includes the results for all groups.

Value

A list containing the following components:

- `gelman` The convergence diagnostics
- `quantiles` The quantiles of each parameter from the posterior distribution
- `statistics` The means and standard deviations of each parameter
- `correlations` The posterior correlations between the parameters

Note that this object is reported silently so will be discarded unless the function is called with an object as in the example below.

Author(s)

Andrew Parnell <andrew.parnell@mu.ie>

See Also

See `simmr_mcmc` for creating objects suitable for this function, and many more examples. See also `simmr_load` for creating simmr objects, `plot.simmr_input` for creating isospace plots, `plot.simmr_output` for plotting output.

Examples

```r
## Not run:
# A simple example with 10 observations, 2 tracers and 4 sources

# The data
data(geese_data_day1)
simmr_1 <- with(
geese_data_day1,
simmr_load(
mixtures = mixtures,
source_names = source_names,
source_means = source_means,
source_sds = source_sds,
correction_means = correction_means,
correction_sds = correction_sds,
concentration_means = concentration_means
)
)

# Plot
```
plot(simmr_1)

# MCMC run
simmr_1_out <- simmr_mcmc(simmr_1)

# Summarise
summary(simmr_1_out) # This outputs all the summaries
summary(simmr_1_out, type = "diagnostics") # Just the diagnostics
# Store the output in an
ans <- summary(simmr_1_out,
   type = c("quantiles", "statistics")
)

## End(Not run)

---

**summary.simmr_output_tdf**

*Summarises the output created with simmr_mcmc_tdf*

### Description

Produces textual summaries and convergence diagnostics for an object created with `simmr_mcmc_tdf`. The different options are: 'diagnostics' which produces Brooks-Gelman-Rubin diagnostics to assess MCMC convergence, 'quantiles' which produces credible intervals for the parameters, 'statistics' which produces means and standard deviations, and 'correlations' which produces correlations between the parameters.

### Usage

```r
## S3 method for class 'simmr_output_tdf'
summary(
  object,
  type = c("diagnostics", "quantiles", "statistics", "correlations"),
  ...
)
```

### Arguments

- `object`: An object of class `simmr_output_df` produced by the function `simmr_mcmc_tdf`
- `type`: The type of output required. At least none of 'diagnostics', 'quantiles', 'statistics', 'correlations'
- `...`: Not used

### Details

The quantile output allows easy calculation of 95 per cent credible intervals of the posterior dietary proportions. The Gelman diagnostic values should be close to 1 to ensure satisfactory convergence. Multiple groups are not currently supported for estimating TDFs.
Value

An list containing the following components:

gelman        The convergence diagnostics
quantiles     The quantiles of each parameter from the posterior distribution
statistics    The means and standard deviations of each parameter
correlations  The posterior correlations between the parameters

Note that this object is reported silently so will be discarded unless the function is called with an object as in the example below.

Author(s)

Andrew Parnell <andrew.parnell@mu.ie>

See Also

See `simmr_mcmc_tdf` for creating objects suitable for this function, and many more examples. See also `simmr_load` for creating simmr objects, `plot.simmr_input` for creating isospace plots, `plot.simmr_output` for plotting output.
Index

* datasets
  geese_data, 8
  geese_data_day1, 9
  simmr_data_1, 19
  simmr_data_2, 20
  square_data, 33
* multivariate
  simmr, 18

combine_sources, 2
compare_groups, 4
compare_sources, 6

geese_data, 8
geese_data_day1, 9

plot.simmr_input, 3, 10, 13, 26, 32, 35, 37
plot.simmr_output, 3, 11, 12, 25, 26, 35, 37
posterior_predictive, 14, 16
print.simmr_input, 15
print.simmr_output, 16
prior_viz, 16

simmr, 18
simmr_data_1, 19
simmr_data_2, 20
simmr_elicit, 20, 20, 25
simmr_load, 3, 5, 7, 10, 13, 18, 22, 25, 26, 31, 32, 35, 37
simmr_mcmc, 3, 5, 7, 9–14, 16–21, 24, 25, 25, 30–32, 34, 35
simmr_mcmc_tdf, 30, 31, 36, 37
square_data, 33
summary.simmr_output, 13, 25, 26, 34
summary.simmr_output_tdf, 31, 32, 36