Package: respeciate (via r-universe)

June 6, 2026

Title Speciation profiles for gases and aerosols

Version 0.4.2

Date 2026-01-10

Description Access to the air pollutant emission profiles in US EPA SPECIATE (v5.2) and EU JRC SPECIEUROPE archives. More details in Simon et al (2010) doi:10.5094/APR.2010.026 and Pernigotti et al (2016) doi:10.1016/j.apr.2015.10.007, respectively.

Type Package

License MIT + file LICENSE

URL https://github.com/atmoschem/respeciate

BugReports https://github.com/atmoschem/respeciate/issues

LazyData yes

Depends R (>= 3.5.0)

RoxygenNote 7.3.3

Encoding UTF-8

Imports data.table, lattice, latticeExtra

Config/pak/sysreqs libjpeg-dev libpng-dev

Repository https://atmoschem.r-universe.dev

Date/Publication 2026-01-11 22:56:09 UTC

RemoteUrl https://github.com/atmoschem/respeciate

RemoteRef HEAD

RemoteSha a756c33bc9c94002ec35adc1013f4126b05dcd91

Contents

`respeciate.generics` *respeciate.generics*

Description

Generic functions for use with `respeciate` object classes.

Usage

```
as.respeciate(x, ...)

## Default S3 method:
as.respeciate(x, ...)

## S3 method for class 'respeciate'
print(x, n = 6, ...)

## S3 method for class 'rsp_pls'
print(x, n = NULL, ...)

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

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

## S3 method for class 'respeciate'
```

```
summary(object, ...)

## S3 method for class 'respeciate'
merge(x, y, ...)
```

Arguments

x	the respeciate object to be printed, plotted, etc.
...	any extra arguments, mostly ignored except by plot which passes them to rsp_plot_profile and merge with passes them to merge.
n	when plotting or printing a multi-profile object, the maximum number of profiles to report.
object	like x but for summary.
y	a second data set, typically a data.frame or a respeciate object, to be merged with x

Value

These generic functions/methods generate typical outputs for respeciate data sets and models: When supplied a data.frame or similar, as.respeciate attempts to coerce it into a respeciate object.

When supplied a respeciate object, print manages its appearance. When supplied a respeciate object, plot provides a basic plot output. This is currently wrapper for the respeciate function rsp_plot_profile.

When supplied a respeciate object, summary generates a summary table of profile information.

When supplied a respeciate object and a second respeciate-like object, e.g. data.frame, respeciate object, etc, merge attempts to merge them using common data columns. You can refine the merge operation using additional arguments.

Note

respeciate objects revert to data.frames when not doing anything package-specific, so you can still use them like data.frames with other packages. This is useful if you have other ideas how to plot more complex (multiple-profile, multiple-species) data sets, and want to use graphics packages like lattice or ggplot2.

---

rsp *Getting archived profiles*

Description

Getting source profile(s) from the local respeciate archives.

Usage

```
rsp(..., include.refs = FALSE, source = "all")

rsp_profile(...)
```

Arguments

`...`	The function assumes all inputs (except `include.refs` and `source`) are profile identifiers: namely, `PROFILE_CODE` and `Species.Id` in SPECIATE and SPECIEUROPE, respectively, or potential sources of profile information and requests these form the local respeciate archives. Typically, simple objects like character and numeric vectors, as assumed to be profile identifiers and composite data-types like `respeciate` or `data.frame` objects are assumed to contain a column named `.profile.id`, the `respeciate` equivalent of `PROFILE_CODE` and `Species.Id`. All recovered identifiers are requested and unrecognized ids (and duplicates) are ignored.
`include.refs`	logical, if profile reference information should be included when extracting the requested profile(s) from the archive, default `FALSE`.
`source`	character, the local archive to request a profile from: `'us'` US EPA SPECIATE, `'eu'` EU JRC SPECIEUROPE, or `'all'` (the default) both.

Value

`rsp_profile` or the short-hand `rsp` return an object of `respeciate` class, a `data.frame` containing one or more profile from the local respeciate archive.

Note

The option `include.refs` adds profile source reference information to the returned `respeciate` data set. The default option is to not include these because some SPECIATE profiles have several associated references and including these replicates records, once per reference. `respeciate` code is written to handle this but if you are developing own methods or code and include references in any profile build you may be biasing some analyses in favor of those multiple-reference profile unless you check and account such cases.

References

For SPECIATE:

Simon, H., Beck, L., Bhave, P.V., Divita, F., Hsu, Y., Luecken, D., Mobley, J.D., Pouliot, G.A., Reff, A., Sarwar, G. and Strum, M., 2010. The development and uses of EPA SPECIATE database. Atmospheric Pollution Research, 1(4), pp.196-206.

For SPECIEUROPE:

Pernigotti, D., Belis, C.A., Spano, L., 2016. SPECIEUROPE: The European data base for PM source profiles. Atmospheric Pollution Research, 7(2), pp.307-314. DOI: https://doi.org/10.1016/j.apr.2015.10.007

See Also

SPECIATE and SPECIEUROPE regarding data sources; and, rsp_find_profile and rsp_find_species regarding archive searching.

Examples

```
## Not run:
x <- rsp_profile(8833, 8850)
plot(x)
## End(Not run)
```

---

rsp.average *Data averaging multiple profile data sets*

Description

Functions to build composite respeciate profiles

rsp_average_profile generates an average composite of a supplied multi-profile respeciate object.

Usage

```
rsp_average_profile(rsp, code = NULL, name = NULL, method = 1, ...)
```

Arguments

rsp	A respeciate object, a data.frame of respeciate profiles.
code	required character, the unique profile code to assign to the average profile.
name	character, the profile name to assign to the average profile. If not supplied, this defaults to a collapsed list of the codes of all the profiles averaged.
method	numeric, the averaging method to apply: Currently only 1 (default) mean(rsp).
...	additional arguments, currently ignored

Value

rsp_average_profile returns a single profile average version of the supplied respeciate profile.

Note

In development function; arguments and outputs likely to be subject to change.

This is one of the very few respeciate functions that modifies the WEIGHT_PERCENT column of the respectiate data.frame.

rsp.build *Building respeciate-like Objects*

Description

rsp function(s) to reconfigure data.frames (and similar object classes) for use with data and functions in respeciate.

Usage

```
rsp_build_x(x, profile_id, profile_name, species_name, species_id, value, ...)

rsp_build_simx(m, n = 1, ...)
```

Arguments

x	data.frame or similar (i.e. something that can be coerced into a data.frame using as.data.frame) to be converted into a respeciate object.
profile_name, profile_id	(character) The names of the columns in x containing profile names and identifiers, respectively. If not already named according to respeciate conventions, at least one of these will need to be assigned.
species_name, species_id	(character) The names of the columns in x containing species name and identifiers, respectively. If not already named according to respeciate conventions, at least one of these will need to be assigned.
value	(character) The name of the column in x containing measurement values. If not already named according to respeciate conventions, this will need to be assigned.
...	(any other arguments) currently ignored.
m	respeciate data set of source profiles intended to be used as the source profiles (or M) matrix when building a simulated data set for use with a PLS model (see rsp_pls_x)
n	a numeric object, e.g. a vector, matrix, data.frame or a similar object that can be coerced into a data.frame of suitable dimensions for use as the source strength matrix (N) to build a simulated data set for use with a PLS model (see rsp_pls_x).

Value

rsp_builds attempt to build and return a respeciate-like object that can be directly compared with data from respeciate.

rsp_build_x is the standard object builder.

rsp_build_simx builds a simulation of an x data set based on the 'linear combination of profiles' model applied in conventional source apportionment. (See below and rsp_pls_x)

Note

If you want to compare your data with profiles in the `respeciate` archive, you need `respeciate` conventions when assigning species names and identifiers. We are working on options to improve on this (and very happy to discuss if anyone has ideas), but current best suggestion is: (1) identify the `respeciate` species code for each of the species in your data set, and (2) assign these as `species_code` when `rsp_building`. The function will then associate the `species_name` from `respeciate` species records.

`rsp.cluster` *Profile cluster analysis methods*

Description

Functions for studying similarities (or dissimilarities) within respeciate data sets

`rsp_distance_profile` calculates the statistical distance between respeciate profiles, and clusters profiles according to nearness.

Usage

```
rsp_distance_profile(rsp, output = c("plot", "report"))
```

Arguments

`rsp`	A `respeciate` object, a `data.frame` of respeciate profiles.
`output`	Character vector, required function output: `'report'` the calculated distance matrix; `'plot'` a heat map of that distance matrix.

Value

Depending on the `output` option, `sp_distance_profile` returns one or more of the following: the correlation matrix, a heat map of the correlation matrix.

Note

Please note: function in development; structure and arguments may be subject to change.

rsp.combine *combining respeciate profiles*

Description

Functions to combining respeciate data sets.

`rsp_lbind` binds two or more `respeciate`-like objects. The default option is to stack the supplied data sets (e.g. `respeciate`, `data.frame`, etc) like `rbindlist` in `data.table` (or row_bind in dplyr). This matches columns by name before stacking the supplied data sets.

Usage

```
rsp_lbind(...)
```

Arguments

`...`	(various) This function is intended to be quite flexible. All supplied arguments are tested and handled as follows: `respeciate`-like objects are passed to `data.table::rbindlist` as a list to rbind using `data.table` methods; Any other arguments that are valid `rbindlist` arguments are passed on 'as is'; And, anything else is (hopefully) ignored.

Value

`rsp_lbind` attempts to return a single stacked version of the supplied data sets. If it is successful, the (stacked) data set is typically returned as a `respeciate` object or a `data.frame` with a warning if it is missing columns `respeciate` expects.

References

Dowle M, Srinivasan A (2023). data.table: Extension of 'data.frame'. R package version 1.14.8, https://CRAN.R-project.org/package=data.table.

rsp.cor *Species correlations*

Description

Functions for studying relationships between species in respeciate data sets.

`rsp_cor_species` generates a by-species correlation matrix of the supplied respeciate data sets.

Usage

```
rsp_cor_species(
  rsp,
  min.n = 3,
  cols = c("#80FFFF", "#FFFFFF", "#FF80FF"),
  na.col = "#CFCFCF",
  heatmap.args = TRUE,
  key.args = TRUE,
  report = "silent"
)
```

Arguments

rsp	respeciate object, a data.frame of respeciate profiles.
min.n	numeric (default 3), the minimum number of species measurements needed in a profile for the function to use it in correlation calculations. Here, it should be noted that this does not guarantee the three matched pairs of measurements needed to calculate a correlation coefficient because not all profiles contain all species, so there may still be insufficient overlap on a case-by-case basis.
cols	a series of numeric, character or other class values that can be translated into a color gradient, used to color valid cases when generating plots and color keys, default c("#80FFFF", "#FFFFFF", "#FF80FF") equivalent to cm.colors output.
na.col	numeric, character or other class that can be translated into a single color, used to color NAs when generating plots and color keys, default grey "#CFCFCF".
heatmap.args	logical or list, heat map settings. Options include TRUE (default) to generate the heat map without modification; FALSE to not plot it; or a list of heat map options to alter the plot default appearance. The plot, a standard heat map with the dendrograms removed, is generated using heatmap, so see associated documentation for valid options.
key.args	logical or list, color key settings if plotting the correlation matrix heat map. Options include TRUE (default) to generate the key without modification; FALSE to not include the key; or a list of options to alter the key appearance.
report	logical or character, the required function output. Options include: 'silent' (default), to return the correlation matrix invisibly; TRUE to return the matrix (visibly); and, FALSE to not return it.

Value

By default rsp_cor_species invisibly returns the calculated correlation matrix a plots it as a heat map, but arguments including heatmap and report can be used to modify function outputs.

rsp.eu *Quick access to common SPECIAEUROPE subsets.*

Description

rsp_eu and rsp_eu_ functions are quick access wrappers to commonly requested SPECIEUROPE subsets.

Usage

```
rsp_eu()

rsp_eu_pm10()

rsp_eu_pm2.5()
```

Value

rsp_eu and rsp_eu_functions typically return a respeciate data.frame of the requested profiles:

rsp_eu() returns all profiles in the local version of SPECIEUROPE

rsp_eu_pm10 returns all SPECIEUROPE profiles classified as PM10 (using Particle.Size=="PM10"), rsp_eu_pm10 for PM2.5 and so on...

See Also

SPECIEUROPE

rsp.export *Exporting respeciate objects*

Description

rsp function(s) to export respeciate (and respeciate-like) objects to other software

Usage

```
rsp_export_esat(
  rsp,
  file.name = "file",
  index = "row.count",
  unc = 0.15,
  bad.values = "fill.1",
  output = c("con.csv", "unc.csv"),
  overwrite = FALSE,
  ...
)
```

Arguments

rsp	(respeciate or similar, e.g. a data-frame set up for use with respeciate), the data-set to export.
file.name	(character), the file name of the exported file or files. See also output and Details below.
index	(character), the name of rsp column to use as the output file(s) index or (default) 'row.count' a row number counter.
unc	(various), if numeric, the scaling factor to apply to concentration values when hole filling uncertainties, else 'eu.rsp', in which case it tries to recover values from the SPECIEUROPE meta information.
bad.values	(character), handling method to use if bad values are found in the supplied data.
output	(character), the file types to export. See also Details below.
overwrite	(character), overwrite file if it already exists.
...	other arguments, currently ignored.

Details

rsp_build_esat makes files that can be used as inputs with ESAT. output options: 'con.csv' and 'unc.csv' (both required by ESAT).

Value

rsp_exports attempt to build and save files suitable for use outside r.

---

rsp.find *Information about data sets currently in respeciate*

Description

Functions that provide respeciate source information. rsp_find_profile searches the currently installed respeciate data sets for profile records. rsp_find species searches the currently installed respeciate data sets for species records.

Usage

```
rsp_find_profile(
  ...,
  by = "keywords",
  partial = TRUE,
  source = "all",
  ref = NULL
)

rsp_profile_info(...)
```

```
rsp_find_species(
  ...,
  by = ".species",
  partial = TRUE,
  source = "all",
  ref = NULL
)

rsp_species_info(...)
```

Arguments

`...`	character(s), any search term(s) to use when searching the local respeciate archive for relevant records using `rsp_find_profile` or `rsp_find_species`.
`by`	character, the section of the archive to search, by default `'keywords'` for `rsp_find_profile` and `'.species'` for `sp_find_species`.
`partial`	logical, if `TRUE` (default) `rsp_find_profile` and `rsp_find_species` use partial matching.
`source`	character, the data set to search: `'us'` US EPA SPECIATE; `'eu'` JRC SPECIEUROPE; or, `'all'` (default) both archives.
`ref`	any `respeciate` object, `data.frame` or similar that profile or species information can be extracted from.

Value

`rsp_profile_info` returns a `data.frame` of profile information, as a `respeciate` object. `rsp_species_info` returns a `data.frame` of species information as a `respeciate` object.

References

For SPECIATE:

Simon, H., Beck, L., Bhave, P.V., Divita, F., Hsu, Y., Luecken, D., Mobley, J.D., Pouliot, G.A., Reff, A., Sarwar, G. and Strum, M., 2010. The development and uses of EPA SPECIATE database. Atmospheric Pollution Research, 1(4), pp.196-206.

For SPECIEUROPE:

Pernigotti, D., Belis, C.A., Spano, L., 2016. SPECIEUROPE: The European data base for PM source profiles. Atmospheric Pollution Research, 7(2), pp.307-314. DOI: https://doi.org/10.1016/j.apr.2015.10.007

See Also

SPECIATE and SPECIEUROPE

Examples

```
## Not run:
profile <- "Ethanol"
pr <- rsp_find_profile(profile)
```

```
pr

species <- "Ethanol"
sp <- rsp_find_species(species)
sp
## End(Not run)
```

`rsp.id` *rsp_id_ functions to identify common species groups for grouping and subsetting respeciate profiles*

Description

`rsp_id_` functions generate a vector of assignment terms and can be used to subset or condition a supplied (re)SPECIATE `data.frame`.

Most commonly, the `rsp_id_` functions accept a single input, a respeciate `data.frame` and return a logical vector of length `nrow(x)`, identifying species of interest as `TRUE`. So, for example, they can be used when `subset`ting in the form:

`subset(rsp, rsp_id_nalkane(rsp))`

... to extract just n-alkane records from a supplied `respeciate` object `rsp`.

However, some accept additional arguments. For example, `rsp_id_copy` also accepts a reference data set, `ref`, and a column identifier, `by`, and tests `rsp$by %in% unique(ref$by)`.

Usage

```
rsp_id_copy(rsp, ref = NULL, by = ".species.id")

rsp_id_nalkane(rsp)

rsp_id_btex(rsp)

rsp_id_pah16(rsp)
```

Arguments

`rsp`	a `respeciate` object, a `data.frame` of respeciate profiles.
`ref`	(`rsp_id_copy` only) a second `respeciate` object, to be used as reference when subsetting (or conditioning) `rsp`.
`by`	(`rsp_id_copy` only) character, the name of the column in `ref` to copy when subsetting (or conditioning) `rsp`.

Value

`rsp_id_copy` outputs can be modified but, by default, it identifies all species in the supplied reference data set.

`rsp_id_nalkane` identifies (straight chain) C1 to C40 n-alkanes.

`rsp_id_btex` identifies the BTEX group of aromatic hydrocarbons (benzene, toluene, ethyl benzene, and M-, O- and P-xylene).

---

rsp.info *Information about data sets currently in respeciate*

Description

Functions that provide respeciate source information. `rsp_info` generates a brief version report for the currently installed respeciate data sets.

Usage

```
rsp_info()
```

Value

`rsp_info` provides a brief version information report on the currently installed respeciate archive.

References

For SPECIATE:

Simon, H., Beck, L., Bhave, P.V., Divita, F., Hsu, Y., Luecken, D., Mobley, J.D., Pouliot, G.A., Reff, A., Sarwar, G. and Strum, M., 2010. The development and uses of EPA SPECIATE database. Atmospheric Pollution Research, 1(4), pp.196-206.

For SPECIEUROPE:

Pernigotti, D., Belis, C.A., Spano, L., 2016. SPECIEUROPE: The European data base for PM source profiles. Atmospheric Pollution Research, 7(2), pp.307-314. DOI: https://doi.org/10.1016/j.apr.2015.10.007

See Also

`SPECIATE` and `SPECIEUROPE`

Examples

```
## Not run:
rsp_info()

## End(Not run)
```

rsp.match *Find nearest matches from reference set of profiles*

Description

rsp_match_profile compares a supplied respeciate profile (or similar data set) and a reference set of supplied profiles and attempts to identify nearest matches on the basis of similarity.

Usage

```
rsp_match_profile(
  rsp,
  ref,
  matches = 10,
  rescale = 5,
  min.n = NULL,
  method = "sid * srd",
  self.test = FALSE,
  ...,
  output = "summary"
)
```

Arguments

rsp	A respeciate object or similar data.frame containing a species profile to be compared with profiles in ref. If rsp contains more than one profile, these are averaged (using rsp_average_profile), and the average compared.
ref	A respeciate object, a data.frame containing a multiple species profiles, to be used as reference library when identifying nearest matches for rsp.
matches	Numeric (default 10), the maximum number of profile matches to report.
rescale	Numeric (default 5), the data scaling method to apply before comparing rsp and profiles in ref: options 0 to 5 handled by rsp_rescale.
min.n	Numeric (or NULL), the minimum number of paired species measurements required for a match to be assessed. The larger min.n, the greater the required rsp and ref profile overlap, so the better the matching confidence for paired cases but also the more likely that a sparse but relevant ref profile may be missing. The default option, NULL, is 65% of the number of species in rsp or 6 if larger.
method	Character (default 'sid * srd'), the ranking metric used to rank profile matches. The function calculates several matching metrics: 'pd', the Pearson's Distance (1 - Pearson's correlation coefficient), 'srd', like pd but using the Spearman Ranked data correlation coefficient, and 'sid', the Standardized Identity Distance (See References). All the metrics tend to zero for better matches, and the method can be any character string that can be evaluated from any of these, e.g., 'pd', 'srd', 'sid', and combinations thereof.

self.test	Logical (default FALSE). The match process self-tests by adding rsp to ref, which should generate an ideal (nearness = 0) score. Setting self.test to TRUE retains this as an extra record.
...	Additional arguments, typically ignore but sometimes used for function development. Currently, testing rm.reps (logical) option to remove what appear to be replicate profile matches from the result set. This is based on the assumption that identical 'pd' and 'sid' scores identical identical ref profiles (or identical overlaps with rsp) but is not validated, so handle with care...
output	Character, output options, including: 'summary' (the default) a data.frame of the requested best matches, ranked according to the method used; 'data' the full data set used to make plots; 'plot' the associated output from rsp_plot_match; or, a combination of these.

Value

By default rsp_match_profile returns a fit report summary: a data.frame of up to matches fit reports for the nearest matches to profiles from the reference profile data set, ref. (See also output above for other options). If several options are requested, earlier options are report (e.g. using print or plot) and only the final option is returned.

References

Distance metrics are based on recommendations by Belis et al (2015) and as implemented in Mooibroek et al (2022):

Belis, C.A., Pernigotti, D., Karagulian, F., Pirovano, G., Larsen, B.R., Gerboles, M., Hopke, P.K., 2015. A new methodology to assess the performance and uncertainty of source apportionment models in intercomparison exercises. Atmospheric Environment, 119, 35–44. https://doi.org/10.1016/j.atmosenv.2015.08.002.

Mooibroek, D., Sofowote, U.M. and Hopke, P.K., 2022. Source apportionment of ambient PM10 collected at three sites in an urban-industrial area with multi-time resolution factor analyses. Science of The Total Environment, 850, p.157981. http://dx.doi.org/10.1016/j.scitotenv.2022.157981.

See Also

rsp_plot_match

---

rsp.pad *Meta-data padding respeciate data sets*

Description

Functions for padding respeciate objects.

rsp_pad pads a supplied respeciate profile data set with profile and species meta-data.

Usage

```
rsp_pad(rsp, pad = "standard", drop.nas = TRUE)
```

Arguments

rsp	A respeciate object, a data.frame of respeciate profiles.
pad	character, type of meta data padding, current options 'profile', 'species', 'weight', 'reference', 'standard' (default; all but 'reference'), and 'all' (all).
drop.nas	logical, discard any rows where the respeciate species amount column .pc.weight is NA, default TRUE.

Value

rsp_pad returns supplied respeciate data set, with requested additional profile and species meta-data added as additional data.frame columns. See Note.

Note

Some data handling can remove respeciate meta-data, and rsp_pads provide a quick rebuild/repair. For example, rsp_dcasting to a (by-species or by-profile) widened form strips some meta-data, and padding is used as part of the rsp_melt_wide to re-add this meta-data when returning the data set to its standard long form.

See Also

rsp_pad

---

rsp.plot *plotting respeciate source profiles*

Description

General plots for respeciate objects.

rsp_plot functions generate plots for supplied respeciate data sets.

Usage

```
rsp_plot_profile(
  rsp,
  id,
  multi.profile = "group",
  order = TRUE,
  log = FALSE,
  ...,
  silent = FALSE,
  output = "default"
)

rsp_plot_species(
```

```
  rsp,
  id,
  multi.species = "group",
  order = FALSE,
  log = FALSE,
  ...,
  silent = FALSE,
  output = "default"
)

rsp_plot_match(
  rsp,
  ref = NULL,
  plot.type = 2,
  log = FALSE,
  ...,
  output = "plot"
)
```

Arguments

rsp	A respeciate object, a data.frame of respeciate profiles.
id	numeric, the indices of profiles or species to use when plotting with rsp_plot_profile or rsp_plot_species, respectively. For example, rsp_plot_profile(rsp, id=1:6) plots first 6 profiles in respeciate object rsp.
multi.profile	character, how rsp_plot_profile should handle multiple profiles, e.g. 'group' or 'panel' (default group).
order	logical, order the species in the profile(s) by relative abundance before plotting.
log	logical, log y scale when plotting.
...	any additional arguments, typically passed on the lattice plotting functions.
silent	logical, hide warnings when generating plots (default FALSE)
output	character, output method, one of: 'plot' to return just the requested plot; 'data' to return just the data; and, c('plot', 'data') to plot then return the data invisibly (default).
multi.species	character, like multi.profile in rsp_plot_profile but for species in rsp_plot_species.
ref	respeciate or similar data set of profiles, used by rsp_match_plot as a reference when comparing with rsp. See rsp_match_profile for further details and other matching arguments.
plot.type	numeric, option if the rsp_plot... function includes different plot reports.

Value

rsp_plot graph, plot, etc usually as a trellis object.

Note

These functions are currently in development, so may change.

References

Most respeciate plots make extensive use of lattice and latticeExtra code:

Sarkar D (2008). *Lattice: Multivariate Data Visualization with R*. Springer, New York. ISBN 978-0-387-75968-5, http://lmdvr.r-forge.r-project.org.

Sarkar D, Andrews F (2022). *latticeExtra: Extra Graphical Utilities Based on Lattice*. R package version 0.6-30, https://CRAN.R-project.org/package=latticeExtra.

They also incorporate ideas from loa:

Ropkins K (2023). *loa: various plots, options and add-ins for use with lattice*. R package version 0.2.48.3, https://CRAN.R-project.org/package=loa.

---

rsp.pls *Positive Least Squares models*

Description

Functions for Positive Least Squares (PSL) fitting of respeciate profiles

rsp_pls_x builds PSL models for supplied profile(s) using the nls function, the 'port' algorithm and a lower limit of zero for all model outputs to enforce the positive fits. The modeled profiles are typically from an external source, e.g. a measurement campaign, and are fit as a linear additive series of reference profiles, here typically from respeciate, to provide a measure of source apportionment based on the assumption that the profiles in the reference set are representative of the mix that make up the modeled sample. The pls_ functions work with rsp_pls_x outputs, and are intended to be used when refining and analyzing these PLS models. See also pls_plots for PLS model plots.

Usage

```
rsp_pls_x(x, m, power = 1, ...)

pls_report(pls)

pls_test(pls)

pls_fit_species(
  pls,
  species,
  power = 1,
  refit.profile = TRUE,
  as.marker = FALSE,
  drop.missing = FALSE,
  ...
)

pls_refit_species(
  pls,
```

```
  species,
  power = 1,
  refit.profile = TRUE,
  as.marker = FALSE,
  drop.missing = FALSE,
  ...
)

pls_rebuild(
  pls,
  species,
  power = 1,
  refit.profile = TRUE,
  as.marker = FALSE,
  drop.missing = FALSE,
  ...
)
```

Arguments

x	A respeciate object, a data.frame of profiles in standard long form, intended for PLS modelling.
m	A respeciate object, a data.frame of profiles also in standard long form, used as the set of candidate source profiles when fitting x.
power	A numeric, an additional factor to be added to weightings when fitting the PLS model. This is applied in the form weight^power, and increasing this, increases the relative weighting of the more heavily weighted measurements. Values in the range 1 - 2.5 are sometimes helpful.
...	additional arguments, typically ignored or passed on to nls.
pls	A rsp_pls_x output, intended for use with pls_ functions.
species	for pls_fit_species, a data.frame of measurements of an additional species to be fitted to an existing PLS model, or for pls_refit_species a character vector of the names of species already included in the model to be refit. Both are multiple-species wrappers for pls_rebuild, a general-purpose PLS fitter than only handles single species.
refit.profile	(for pls_fit_species, pls_refit_species and pls_rebuild) logical. When fitting a new species (or refitted an existing species), all other species in the reference profiles are held 'as is' and added species is fit to the source contribution time-series of the previous PLS model. By default, the full PLS model is then refit using the revised m source profile to generate a PLS model based on the revised source profiles (i.e., m + new species or m + refit species). However, this second step can be omitted using refit.profile=FALSE if you want to use the supplied species as an indicator rather than a standard member of the apportionment model.
as.marker	for pls_rebuild, pls_fit_species and pls_refit_species, logical, default FALSE, when fitting (or refitting) a species, treat it as source marker.

drop.missing — for pls_rebuild, pls_fit_species and pls_refit_species, logical, default FALSE, when building or rebuilding a PLS model, discard cases where species is missing.

Value

rsp_pls_x returns a list of nls models, one per profile/measurement set in x. The pls_ functions work with these outputs. pls_report generates a data.frame of model outputs, and is used of several of the other pls_ functions. pls_fit_species, pls_refit_species and pls_fit_parent return the supplied rsp_pls_profile output, updated on the basis of the pls_ function action. pls_plots (documented separately) produce various plots commonly used in source apportionment studies.

Note

This implementation of PLS applies the following modeling constraints:

1. It generates a model of x that is positively constrained linear product of the profiles in m, so outputs can only be zero or more. Although the model is generated using nls, which is a Nonlinear Least Squares (NLS) model, the fitting term applied in this case is linear.

2. The model is fit in the form:

$$X_{i,j} = \sum_{k=1}^{K} N_{i,k} * M_{k,j} + e_{i,j}$$

Where X is the data set of measurements, input x in rsp_pls_x, M (m) is data set of reference profiles, and N is the data set of source contributions, the source apportion solution, to be solved by minimising e, the error terms.

3. The number of species in x must be more than the number of profiles in m to reduce the likelihood of over-fitting.

---

rsp.pls.plot *Plots for use with respeciate profile Positive Least Squares models*

Description

The pls_plot functions are intended for use with PLS models built using rsp_pls_profile (documented separately). They generate some plots commonly used with source apportionment model outputs.

Usage

```
pls_plot(pls, plot.type = 1, ..., output = "default")

pls_plot_profile(pls, plot.type = 1, log = FALSE, ..., output = "default")

pls_plot_species(pls, id, plot.type = 1, ..., output = "default")
```

Arguments

pls	A rsp_pls_profile output, intended for use with pls_ functions.
plot.type	numeric, the plot type if multiple options are available.
...	other arguments, typically passed on to the associated lattice plot.
output	character, output method, one of: 'plot' to return just the requested plot; 'data' to return just the data; and, c('plot', 'data') to plot then return the data invisibly (default).
log	(for pls_plot_profile only) logical, if TRUE this applies 'log' scaling to the primary Y axes of the plot.
id	numeric or character identifying the species or profile to plot. If numeric, these are treated as indices of the species or profile, respectively, in the PLS model; if character, species is treated as the name of species and profile is treated as the profile code. Both can be concatenated to produce multiple plots and the special case id = -1 is a short cut to all species or profiles, respectively.

Value

pls_plots produce various plots commonly used in source apportionment studies.

---

rsp.rescale *rescaling respeciate profiles*

Description

Functions for rescaling respeciate data sets

rsp_rescale rescales the percentage weight records in a supplied respeciate profile data set. This can be by profile or species subsets, and rsp_rescale_profile and rsp_rescale_species provide short-cuts to these options.

Usage

```
rsp_rescale(rsp, method = 2, by = "species")

rsp_rescale_profile(rsp, method = 1, by = "profile")

rsp_rescale_species(rsp, method = 2, by = "species")
```

Arguments

rsp	A respeciate object, a data.frame of respeciate profiles.
method	numeric, the rescaling method to apply: 1 x/total(x); 2 x/mean(x); 3 x-min(x)/max(x)-min(x); 4 x-mean(x)/sd(x); 5 x/max(x). The alternative 0 returns the records to their original values.
by	character, when rescaling x with rsp_rescale, the data type to group and rescale, currently 'species' (default) or 'profile'.

Value

rsp_rescale and rsp_rescale return the respeciate profile with the percentage weight records rescaled using the requested method. See Note.

Note

Data sometimes needs to be normalised, e.g. when applying some statistical analyses. Rather than modify source information in SPECIATE and SPECIEUROPE, respeciate creates a duplicate column .value which is modified by operations like sp_rescale_profile and sp_rescale_species. This means rescaling is always applied to the source information, rather than rescaling an already rescaled value, and the EPA records are retained unaffected. So, the original source information can be easily recovered.

References

Dowle M, Srinivasan A (2023). data.table: Extension of 'data.frame'. R package version 1.14.8, https://CRAN.R-project.org/package=data.table.

---

rsp.reshape *Reshaping respeciate data sets*

Description

Functions for reshaping respeciate profiles

rsp_dcast and rsp_melt_wide reshape supplied respeciate profile(s). rsp_dcast converts these from their supplied long form to a widened form, dcasting the data set by either species or profiles depending on the widen setting applied. rsp_dcast_profile, rsp_dcast_profile_id, rsp_dcast_species and rsp_dcast_species_id are wrappers for these options. rsp_melt_wide attempts to return a previously widened data set to the original long form.

Usage

```
rsp_dcast(rsp, widen = "species")

rsp_dcast_profile(rsp, widen = "profile")

rsp_dcast_profile_id(rsp, widen = "profile.id")

rsp_dcast_species(rsp = rsp, widen = "species")

rsp_dcast_species_id(rsp = rsp, widen = "species.id")

rsp_melt_wide(rsp, pad = FALSE, drop.nas = FALSE)
```

Arguments

rsp	A respeciate object, a data.frame of respeciate profiles in standard long form or widened form using rsp_dcast and rsp_melt_wide, respectively.
widen	character, when widening rsp with rsp_dcast, the data type to dcast, currently 'species' (default), 'species.id', 'profile' or 'profile.id'. See Note.
pad	logical or character, when melting a previously widened data set, should output be re-populated with species and/or profile meta-data, discarded when widening. This is currently handled by rsp_pad. The default FALSE does not pad, TRUE pads, applies standard settings, so does not include profile sources reference meta-data. (See rsp_pad for other options).
drop.nas	logical, when melting a previously widened data set, should output be stripped of any rows containing empty weight/value columns. Because not all profile contains all species, the dcast/melt process can generate empty rows, and this step attempt account for that when working with standard reSPECIATE profiles. It is, however, sometimes useful to check first, e.g. when building profiles yourself.

Value

rsp_dcast returns the wide form of the supplied respeciate profile. rsp_melt_wide returns the (standard) long form of a previously widened profile.

Note

Conventional long-to-wide reshaping of data, or dcasting, can be slow and memory inefficient. So, respeciate uses the data.table::dcast method. The rsp_dcast_species method, applied using widen='species', is effectively:

```
dcast(..., .profile.id+.profile~.species, value.var=".value")
```

And, the alternative widen='profile':

```
dcast(..., .species.id+.species~.profile, value.var=".value")
```

respeciate uses a local version of the SPECIATE and SPECIEUROPE weight measurements .value, so the EPA and JCR source information can easily be recovered. See also rsp_rescale_profile.

References

Dowle M, Srinivasan A (2023). _data.table: Extension of 'data.frame'_. R package version 1.14.8, <https://CRAN.R-project.org/package=data.table>.

---

rsp.us *Quick access to common SPECIATE subsets.*

Description

rsp_us_ functions are quick access wrappers to commonly requested SPECIATE subsets.

Usage

```
rsp_us_gas()

rsp_us_other()

rsp_us_pm()

rsp_us_pm.ae6()

rsp_us_pm.ae8()

rsp_us_pm.cr1()

rsp_us_pm.simplified()
```

Value

`rsp_us_` functions typically return a `respeciate` `data.frame` of the requested profiles.

For example:

`rsp_us_gas()` returns all gaseous profiles in SPECIATE (`PROFILE_TYPE == 'GAS'`).

`rsp_us_pm` returns all particulate matter (PM) profiles in SPECIATE not classified as a special PM type (`PROFILE_TYPE == 'PM'`).

The special PM types are subsets profiles intended for special applications, and these include `rsp_us_pm.ae6` (type `PM-AE6`), `rsp_us_pm.ae8` (type `PM-AE8`), `rsp_us_pm.cr1` (type `PM-CR1`), and `rsp_us_pm.simplified` (type `PM-Simplified`).

`rsp_us_other` returns all profiles classified as other in SPECIATE (`PROFILE_TYPE == 'OTHER'`).

See Also

SPECIATE

SPECIATE *SPECIATE*

Description

the SPECIATE data set is a local version of the EPA's SPECIATE repository of organic gas and particulate matter (PM) speciation profiles of air pollution sources.

Currently using version 5.4 as of 2025-11-18.

Usage

```
SPECIATE
```

Format

A (13 long) 'list' object

PROFILES The main `data.frame` of profile-specific meta-data, with one row per profile, key term `PROFILE_CODE`.

SPECIES The main `data.frame` of individual record meta-data, with one row per species in each profile, key terms `PROFILE_CODE` and `SPECIES_ID` linking `PROFILES` and `SPECIES_PROPERTIES`.

SPECIES_PROPERTIES The main `data.frame` of species-specific meta-data, with one row per species, key term `SPECIES_ID`.

PROFILE_REFERENCE The `data.frame` linking profile and reference meta-data, one row per references per profile, key terms `PROFILE_CODE` and `REF_Code`.

REFERENCES The main `data.frame` of references for profile source meta-data, one row per reference, key term `REF_Code`.

And others Currently not documented.

Source

https://www.epa.gov/air-emissions-modeling/speciate

References

Simon, H., Beck, L., Bhave, P.V., Divita, F., Hsu, Y., Luecken, D., Mobley, J.D., Pouliot, G.A., Reff, A., Sarwar, G. and Strum, M., 2010. The development and uses of EPA SPECIATE database. Atmospheric Pollution Research, 1(4), pp.196-206.

SPECIEUROPE *SPECIEUROPE*

Description

The SPECIEUROPE data set is a local version of the European Commission (EC) Joint Research Centre JRC's repository of particulate matter (PM) speciation profiles of European air pollutant sources.

Currently using version 3.0 as of 2025-11-19.

Usage

```
SPECIEUROPE
```

Format

A (3 long) 'list' object

source The main SPECIEUROPE data set

ref The source citation, to be used whenever this data is used.

website The SPECIEUROPE project website link

Source

https://source-apportionment.jrc.ec.europa.eu/

References

Pernigotti, D., Belis, C.A., Spano, L., 2016. SPECIEUROPE: The European data base for PM source profiles. Atmospheric Pollution Research, 7(2), pp.307-314. DOI: https://doi.org/10.1016/j.apr.2015.10.007

Index