Package: fMRIscrub (via r-universe)

October 29, 2024

Type Package

Title Scrubbing and Other Data Cleaning Routines for fMRI

Version 0.13.0

Maintainer Amanda Mejia <mandy.mejia@gmail.com>

Description Data-driven fMRI denoising with projection scrubbing (Pham et al (2022) <doi:10.1016/j.neuroimage.2023.119972>). Also includes routines for DVARS (Derivatives VARianceS) (Afyouni and Nichols (2018) <doi:10.1016/j.neuroimage.2017.12.098>), motion scrubbing (Power et al (2012) <doi:10.1016/j.neuroimage.2011.10.018>), aCompCor (anatomical Components Correction) (Muschelli et al (2014) <doi:10.1016/j.neuroimage.2014.03.028>), detrending, and nuisance regression. Projection scrubbing is also applicable to other outlier detection tasks involving high-dimensional data.

Depends R (>= 3.5.0)

License GPL-3

Encoding UTF-8

Imports MASS, cellWise, e1071, fMRItools (>= 0.2.2), pesel, robustbase, stats, expm, utils, gamlss

Suggests corpcor, cowplot, ciftiTools, gifti, knitr, rmarkdown, RNifti, ggplot2, fastICA, oro.nifti, testthat (>= 3.0.0), covr

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.3

URL https://github.com/mandymejia/fMRIscrub

BugReports https://github.com/mandymejia/fMRIscrub/issues

LazyData true

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

RemoteUrl https://github.com/mandymejia/fmriscrub

RemoteRef HEAD

RemoteSha 7ee8d6defb35b7414ae74bbcd38fa5b58dc73966

Contents

artifact_images . . . 2
Dat1 . . . 3
Dat2 . . . 4
DVARS . . . 4
FD . . . 5
flags_to_nuis_spikes . . . 7
fMRIscrub . . . 7
high_kurtosis . . . 8
leverage . . . 8
plot.scrub_DVARS . . . 9
plot.scrub_FD . . . 10
plot.scrub_projection . . . 10
pscrub . . . 11
robdist . . . 14
rob_stabilize . . . 17
scrub . . . 18
scrub_xifti . . . 18
SHASH_out . . . 19
SHASH_to_normal . . . 20
summary.scrub_DVARS . . . 21
summary.scrub_FD . . . 21
summary.scrub_projection . . . 22

Index **23**

`artifact_images` *Artifact images*

Description

Visualize artifact patterns from the results of `pscrub`. Requires `pscrub(..., get_dirs=TRUE)`.

Usage

```
artifact_images(psx, idx = NULL, use_dt = TRUE)
```

Arguments

`psx`	A `"scrub_projection"` object containing projection scrubbing results.
`idx`	The timepoints or column indexes for which to compute artifact images. If `NULL` (default), use the outlying timepoints.
`use_dt`	If detrended components are available (the "U" matrix of PCA or "M" matrix of ICA), should they be used to compute the artifact images? Otherwise, use the non-detrended components. Default: `TRUE`.

Details

Computes two types: "mean" artifact images based on a weighted sum of the projection directions, with weights determined by the scores for each component at the flagged timepoint, and "top" artifact images based on the projection direction with the greatest score at the flagged timepoint.

Value

A list of three: `idx`, the timepoints for which the artifact images were computed; `mean`, the "mean" artifact images; and `top`, the "top" artifact images. The row names of the `top` artifact images matrix give the index of the top component ("V" in PCA and "S" in ICA) at each timepoint.

`Dat1` *First Example Time Series from the ABIDE*

Description

A sagittal slice from the fMRI time series for subject 0050048. The scan was obtained at the University of Pittsburgh School of Medicine. The scan has been pre-processed with slice time correction, rigid body realignment estimation, spatial normalization to MNI space, and linear detrending. Subject 0050048 was a typically-developing 11-year-old male. The scan has many artifacts. A mask was applied to vectorize the spatial dimensions.

Usage

```
Dat1
```

Format

A numeric matrix of 193 time points by 4675 voxels

Details

Source: http://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html

References

- 1.Di Martino, A. et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry 19, 659–667 (2014).

Dat2 *Second Example Time Series from the ABIDE*

Description

A sagittal slice from the fMRI time series for subject 0051479. The scan was obtained at the California Institute of Technology. The scan has been pre-processed with slice time correction, rigid body realignment estimation, spatial normalization to MNI space, and linear detrending. Subject 0051479 was a typically-developing 20-year-old female. The scan has few visible artifacts. A mask was applied to vectorize the spatial dimensions.

Usage

```
Dat2
```

Format

A numeric matrix of 145 time points by 4679 voxels

Details

Source: http://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html

References

- 1.Di Martino, A. et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry 19, 659–667 (2014).

DVARS *DVARS*

Description

Computes the DSE decomposition and DVARS-related statistics. Based on code from github.com/asoroosh/DVARS .

Usage

```
DVARS(
  X,
  normalize = TRUE,
  cutoff_DPD = 5,
  cutoff_ZD = qnorm(1 - 0.05/nrow(as.matrix_ifti(X))),
  verbose = FALSE
)
```

Arguments

`X`	a T by N numeric matrix representing an fMRI run. There should not be any missing data (`NA` or `NaN`).
`normalize`	Normalize the data? Default: `TRUE`. Normalization removes constant-zero voxels, scales by 100 / the median of the mean image, and then centers each voxel on its mean. To replicate Afyouni and Nichols' procedure for the HCP MPP data, since the HCP scans are already normalized to 10,000, just divide the data by 100 and center the voxels on their means: `Y <- Y/100; DVARS(t(Y - apply(Y, 1, mean)))` where `Y` is the V by T data matrix. Note that while voxel centering doesn't affect DVARS, it does affect DPD and ZD.
`cutoff_DPD, cutoff_ZD`	Numeric outlier cutoffs. Timepoints exceeding these cutoffs will be flagged as outliers.
`verbose`	Should occasional updates be printed? Default is `FALSE`.

Value

A list with components

measure A data.frame with T rows, each column being a different variant of DVARS.

measure_info "DVARS"

outlier_cutoff The outlier cutoff value(s).

outlier_flag A logical data.frame with T rows, where `TRUE` indicates suspected outlier presence.

References

- Afyouni, S. & Nichols, T. E. Insight and inference for DVARS. NeuroImage 172, 291-312 (2018).

`FD` *Framewise Displacement*

Description

Calculate Framewise Displacement (FD)

Usage

```
FD(
  X,
  trans_units = c("mm", "cm", "in"),
  rot_units = c("deg", "rad", "mm", "cm", "in"),
  brain_radius = NULL,
  detrend = FALSE,
  lag = 1,
  cutoff = 0.3
)
```

Arguments

X	An N by 6 matrix in which the first three columns represent the translational RPs (trans_units), and the second three columns represent the rotational RPs (rot_units). If rot_units measures an angle, it will be converted to trans_units by measuring displacement on a sphere of radius brain_radius trans_units. Alternatively, this can be the file path to an N by 6 matrix which can be read with read.table (fields separated by white-space; no header).
trans_units	"mm" for millimeters (default), "cm" for centimeters, or "in" for inches.
rot_units	"deg" for degrees (default), "rad" for radians, or one of the trans_units options.
brain_radius	If rot_units measures an angle, the rotational RPs are transformed to a spatial measurement representing the displacement on a sphere of radius brain_radius trans_units. If brain_radius is NULL (default), it will be set to 50 mm.
detrend	Detrend each RP with the DCT before computing FD? Default: FALSE. Can be a number of DCT bases to use, or TRUE to use 4.
lag	The difference of indices between which to calculate change in position. Default: 1 (the previous timepoint). Changing this argument sets $\Delta x_i = x_{i-lag} - x_i$ (and similarly for the other RPs).
cutoff	FD values higher than this will be flagged. Default: .3.

Details

The FD formula is taken from Power et. al. (2012):

$$FD_i = |\Delta x_i| + |\Delta y_i| + |\Delta z_i| + |\Delta \alpha_i| + |\Delta \beta_i| + |\Delta \gamma_i|$$

where i is the timepoint; x, y and z are the translational realignment parameters (RPs); α, β and γ are the rotational RPs; and $\Delta x_i = x_{i-1} - x_i$ (and similarly for the other RPs).

Value

A list with components

measure A length N vector of FD values in trans_units.

measure_info "FD"

outlier_cutoff `cutoff`

outlier_flag A length-N logical vector, where `TRUE` indicates suspected outlier presence.

References

- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59, 2142-2154 (2012).

`flags_to_nuis_spikes` *Flags to nuisance spikes*

Description

Convert flagged volumes to corresponding one-hot encoded vectors which can be used for nuisance regression.

Usage

```
flags_to_nuis_spikes(flags, n_time)
```

Arguments

`flags`	Numeric vector of integers indicating the indices of the flagged volumes. Or, a logical vector of length `n_time` where `TRUE` values indicate the flagged volumes.
`n_time`	The length of the vectors to obtain. For nuisance regression, this is the length of the BOLD data. The highest index in `flags` should not exceed `n_time`.

Value

A numeric matrix of ones and zeroes. The number of rows will be `n_time` and the number of columns will be the number of flags. Each column will have a `1` at the flag index, and `0` elsewhere.

`fMRIscrub` *fMRIscrub: fMRI scrubbing and other data cleaning routines*

Description

See `help(package="fMRIscrub")` for a list of functions.

high_kurtosis *Which components have high kurtosis?*

Description

The kurtosis cutoff is a high quantile (default 0.99) of the sampling distribution of kurtosis for Normal iid data of the same length as the components; it is estimated by simulation or calculated from the theoretical asymptotic distribution if the components are long enough.

Usage

```
high_kurtosis(Comps, kurt_quantile = 0.99, n_sim = 5000, min_1 = FALSE)
```

Arguments

Comps	A matrix; each column is a component. For PCA, this is the U matrix. For ICA, this is the M matrix.
kurt_quantile	components with kurtosis of at least this quantile are kept.
n_sim	The number of simulation data to use for estimating the sampling distribution of kurtosis. Only used if a new simulation is performed. (If $n < 1000$ and the quantile is 90%, a pre-computed value is used instead. If $n > 1000$, the theoretical asymptotic distribution is used instead.)
min_1	Require at least one component to be selected? In other words, if no components meet the quantile cutoff, should the component with the highest kurtosis be returned? Default: FALSE.

Details

The components should not have any strong low-frequency trends, because trends can affect kurtosis in unpredictable ways unrelated to outlier presence.

Value

A logical vector indicating whether each component has high kurtosis.

leverage *Leverage*

Description

Computes the leverage of each observation in the PC score (U) or IC mixing (M) matrix for projection scrubbing. Can threshold to flag potential outliers.

Usage

```
leverage(Comps, are_norm = FALSE, median_cutoff = NULL)
```

Arguments

Comps	The n by Q PC score matrix/IC mixing matrix.
are_norm	Assume the columns of Comps are orthogonal and have 2-norms equal to 1? Speeds up the computation.
median_cutoff	The outlier cutoff, in multiples of the median leverage. Default: NULL (do not compute outliers).

Value

A list with entries "meas" (the leverage values), "cut" (the leverage cutoff value) and "flag" (logical vector indicating the outliers). If !is.null(median_cutoff), "cut" and "flag" are omitted.

---

plot.scrub_DVARS *Plot a "scrub_DVARS" object*

Description

Plot a "scrub_DVARS" object

Usage

```
## S3 method for class 'scrub_DVARS'
plot(x, title = NULL, ...)
```

Arguments

x	The "scrub_DVARS" object
title	(Optional) If provided, will add a title to the plot.
...	Additional arguments to ggplot, e.g. main, sub, xlab, ylab, legend.position

Value

A ggplot

plot.scrub_FD *Plot a "scrub_FD" object*

Description

Plot a "scrub_FD" object

Usage

```
## S3 method for class 'scrub_FD'
plot(x, title = NULL, ...)
```

Arguments

x	The "scrub_FD" object
title	(Optional) If provided, will add a title to the plot.
...	Additional arguments to ggplot, e.g. main, sub, xlab, ylab, legend.position

Value

A ggplot

plot.scrub_projection *Plot a "scrub_projection" object*

Description

Plot a "scrub_projection" object

Usage

```
## S3 method for class 'scrub_projection'
plot(x, title = NULL, ...)
```

Arguments

x	The "scrub_projection" object
title	(Optional) If provided, will add a title to the plot.
...	Additional arguments to ggplot, e.g. main, sub, xlab, ylab, legend.position

Value

A ggplot

pscrub *Projection scrubbing*

Description

Projection scrubbing is a data-driven method for identifying artifact-contaminated volumes in fMRI. It works by identifying component directions in the data likely to represent patterns of burst noise, and then computing a composite measure of outlyingness based on leverage within these directions, at each volume. The projection can be PCA, ICA, or "fused PCA." Projection scrubbing can also be used for other outlier detection tasks involving high-dimensional data.

Usage

```
pscrub(
  X,
  projection = c("ICA", "PCA"),
  nuisance = "DCT4",
  center = TRUE,
  scale = TRUE,
  comps_mean_dt = FALSE,
  comps_var_dt = FALSE,
  PESEL = TRUE,
  kurt_quantile = 0.99,
  get_dirs = FALSE,
  full_PCA = FALSE,
  get_outliers = TRUE,
  cutoff = 4,
  seed = 0,
  verbose = FALSE
)
```

Arguments

X	Wide numeric data matrix (T observations by V variables, $T << V$). If X represents an fMRI run, T should be the number of timepoints and V should be the number of vertices/voxels. Projection scrubbing will measure the outlyingness of each row in X.
projection	One of the following: "ICA" (default) or "PCA".
nuisance	Nuisance signals to regress from each column of X. Should be specified as a design matrix: a T by N numeric matrix where N represents the number of nuisance signals. Or can be "DCT4" (default), which will create a matrix with a constant column (the intercept term) and four DCT bases. This default nuisance regression will have the effect of demeaning and detrending the data by removing low-frequency components. To not perform any nuisance regression set this argument to NULL, 0, or FALSE. Detrending is highly recommended for time-series data, especially if there are many time points or evolving circumstances affecting the data. Additionally, if

kurtosis is being used to select the projection directions, trends can induce positive or negative kurtosis, contaminating the connection between high kurtosis and outlier presence. Detrending should not be used with non-time-series data because the observations are not temporally related.

Additional nuisance regressors can be specified like so: `cbind(1, fMRItools::dct_bases(nrow(x), 4), more_nuisance)`.

`center`, `scale`
: Center the columns of the data by their medians, and scale the columns of the data by their median absolute deviations (MADs)? Default: `TRUE`. Centering is necessary for computing the projections, so if `center` is `FALSE`, the data must already be centered.

  Note that centering and scaling occur after nuisance regression, so even if `center` is `FALSE`, the data will be centered on the means if the nuisance regression included an intercept term, as it does by default.

`comps_mean_dt`, `comps_var_dt`
: Stabilize the mean and variance of each projection component's timecourse prior to computing kurtosis and leverage? These arguments should be `TRUE`, `FALSE` (default), or the number of DCT bases to use for detrending (`TRUE` will use 4). Note that these arguments affect the projection components and not the data itself. Also, if variance-stabilizing but not mean-stabilizing, the components must already be expected to be mean-stabilized, for example if the data was rigorously detrended; otherwise, the results will be invalid.

  Slow-moving mean and variance patterns in the components will interfere with the roles of kurtosis and leverage in identifying outliers. While `nuisance` can be used to detrend the data, this nuisance regression is estimated *non-robustly*, since a robust model takes too long to estimate at each data location. On the other hand, `comps_mean_dt` and `comps_var_dt` can be used to apply a *robust* nuisance regression at each component, since there are much fewer components than original data locations. Thus, even if the data has been detrended with `nuisance` it may be helpful to detrend the components with `comps_mean_dt`; furthermore, the data nuisance regression does not address the potential existence of variance patterns in the components.

  Overall, for fMRI we recommend enabling `comps_mean_dt` and `comps_var_dt` unless the data has been cleaned not only with a low-pass filter like DCT nuisance regression, but also with anatomical CompCor, ICA-FIX, or a similar data-driven strategy that takes into account common sources of artifactual mean and variance trends such as motion and physiological cycles.

`PESEL`
: Use `pesel` to select the number of components? Default: `TRUE`. Otherwise, use the number of principal components with above-average variance.

`kurt_quantile`
: What quantile cutoff should be used to select the components? Default: `0.99`. Use `0` to select all high-variance components regardless of kurtosis value.

  We model each component as a length T vector of Normal iid random variables, for which the distribution of kurtosis values can be approximated. The quantile is estimated based on this distribution.

`get_dirs`
: Should the projection directions be returned? This is the V matrix in PCA and S matrix in ICA. The default is `FALSE` to save memory. However, `get_dirs==TRUE` is required for `artifact_images`.

full_PCA	Only applies to the PCA projection. Return the full SVD? Default: FALSE (return only the high-variance components).
get_outliers	Should outliers be flagged based on cutoff? Default: TRUE.
cutoff	Median leverage cutoff value. Default: 4.
seed	Set a seed right before the call to pesel::pesel or ica::icaimax? If NULL, do not set a seed. If numeric (default: 0), will use as the seed.
verbose	Should occasional updates be printed? Default: FALSE.

Details

Refer to the projection scrubbing vignette for a demonstration and an outline of the algorithm: vignette("projection_scrubbing", package="fMRIscrub")

Value

A "pscrub" object, i.e. a list with components

measure A numeric vector of leverage values.

outlier_cutoff The numeric outlier cutoff value (cutoff times the median leverage).

outlier_flag A logical vector where TRUE indicates where leverage exceeds the cutoff, signaling suspected outlier presence.

mask A length P numeric vector corresponding to the data locations in X. Each value indicates whether the location was masked:

- **0** The data location was not masked out.
- **-1** The data location was masked out, because it had at least one NA or NaN value.
- **-2** The data location was masked out, because it was constant.

PCA This will be a list with components:

- **U** The T by Q PC score matrix.
- **D** The standard deviation of each PC.
- **V** The P by Q PC directions matrix. Included only if get_dirs.
- **highkurt** The length Q logical vector indicating scores of high kurtosis.
- **U_dt** Detrended components of U. Included only if components were mean- or variance-detrended.
- **highkurt** The length Q logical vector indicating detrended scores of high kurtosis.
- **nPCs_PESEL** The number of PCs selected by PESEL.
- **nPCs_avgvar** The number of above-average variance PCs.

where Q is the number of PCs selected by PESEL or of above-average variance (or the greater of the two if both were used). If PCA was not used, all entries except nPCs_PESEL and/or nPCs_avgvar will not be included, depending on which method(s) was used to select the number of PCs.

ICA If ICA was used, this will be a list with components:

- **S** The P by Q source signals matrix. Included only if get_dirs
- **M** The T by Q mixing matrix.
- **highkurt** The length Q logical vector indicating mixing scores of high kurtosis.

M_dt Detrended components of `M`. Included only if components were mean- or variance-detrended.

highkurt The length `Q` logical vector indicating detrended mixing scores of high kurtosis. Included only if components were mean- or variance-detrended.

References

- Mejia, A. F., Nebel, M. B., Eloyan, A., Caffo, B. & Lindquist, M. A. PCA leverage: outlier detection for high-dimensional functional magnetic resonance imaging data. Biostatistics 18, 521-536 (2017).
- Pham, D., McDonald, D., Ding, L., Nebel, M. B. & Mejia, A. Less is more: balancing noise reduction and data retention in fMRI with projection scrubbing. (2022).

Examples

```
library(fastICA)
n_voxels = 2e3
n_timepoints = 35
X = matrix(rnorm(n_timepoints*n_voxels), ncol = n_voxels)

psx = pscrub(X)
```

`robdist` *Robust distance scrubbing*

Description

Scrubbing with robust distance.

Usage

```
robdist(
  X,
  RD_cutoff = 4,
  RD_quantile = 0.99,
  trans = c("none", "robust-YJ", "SHASH"),
  bootstrap_n = 1000,
  bootstrap_alpha = 0.01,
  projection = c("ICA", "PCA"),
  nuisance = "DCT4",
  center = TRUE,
  scale = TRUE,
  comps_mean_dt = FALSE,
  comps_var_dt = FALSE,
  PESEL = TRUE,
  kurt_quantile = 0.99,
  get_dirs = FALSE,
```

```
  full_PCA = FALSE,
  get_outliers = TRUE,
  cutoff = 4,
  seed = 0,
  skip_dimred = FALSE,
  verbose = FALSE
)
```

Arguments

X	Wide numeric data matrix (T observations by V variables, $T << V$). If X represents an fMRI run, T should be the number of timepoints and V should be the number of vertices/voxels. Projection scrubbing will measure the outlyingness of each row in X.
RD_cutoff	Default: 4.
RD_quantile	Quantile cutoff...?
trans	Apply a transformation prior to univariate outlier detection? Three options: "none" (default), "robust-YJ", and "SHASH".
bootstrap_n	Use bootstrapping to estimate the robust distance null distribution? If so, set this to the number of bootstraps. Default: 100. Use 0 (or FALSE), to use an empirical quantile instead.
bootstrap_alpha	If using bootstrap (bootstrap > 0), this is the level of the bootstrap CI. Default: 0.99.
projection	One of the following: "ICA" (default) or "PCA".
nuisance	Nuisance signals to regress from each column of X. Should be specified as a design matrix: a T by N numeric matrix where N represents the number of nuisance signals. Or can be "DCT4" (default), which will create a matrix with a constant column (the intercept term) and four DCT bases. This default nuisance regression will have the effect of demeaning and detrending the data by removing low-frequency components. To not perform any nuisance regression set this argument to NULL, 0, or FALSE. Detrending is highly recommended for time-series data, especially if there are many time points or evolving circumstances affecting the data. Additionally, if kurtosis is being used to select the projection directions, trends can induce positive or negative kurtosis, contaminating the connection between high kurtosis and outlier presence. Detrending should not be used with non-time-series data because the observations are not temporally related. Additional nuisance regressors can be specified like so: cbind(1, fMRItools::dct_bases(nrow(x), 4), more_nuisance).
center, scale	Center the columns of the data by their medians, and scale the columns of the data by their median absolute deviations (MADs)? Default: TRUE. Centering is necessary for computing the projections, so if center is FALSE, the data must already be centered. Note that centering and scaling occur after nuisance regression, so even if center is FALSE, the data will be centered on the means if the nuisance regression included an intercept term, as it does by default.

comps_mean_dt, comps_var_dt
: Stabilize the mean and variance of each projection component's timecourse prior to computing kurtosis and leverage? These arguments should be TRUE, FALSE (default), or the number of DCT bases to use for detrending (TRUE will use 4). Note that these arguments affect the projection components and not the data itself. Also, if variance-stabilizing but not mean-stabilizing, the components must already be expected to be mean-stabilized, for example if the data was rigorously detrended; otherwise, the results will be invalid.

  Slow-moving mean and variance patterns in the components will interfere with the roles of kurtosis and leverage in identifying outliers. While nuisance can be used to detrend the data, this nuisance regression is estimated *non-robustly*, since a robust model takes too long to estimate at each data location. On the other hand, comps_mean_dt and comps_var_dt can be used to apply a *robust* nuisance regression at each component, since there are much fewer components than original data locations. Thus, even if the data has been detrended with nuisance it may be helpful to detrend the components with comps_mean_dt; furthermore, the data nuisance regression does not address the potential existence of variance patterns in the components.

  Overall, for fMRI we recommend enabling comps_mean_dt and comps_var_dt unless the data has been cleaned not only with a low-pass filter like DCT nuisance regression, but also with anatomical CompCor, ICA-FIX, or a similar data-driven strategy that takes into account common sources of artifactual mean and variance trends such as motion and physiological cycles.

PESEL
: Use pesel to select the number of components? Default: TRUE. Otherwise, use the number of principal components with above-average variance.

kurt_quantile
: What quantile cutoff should be used to select the components? Default: 0.99. Use 0 to select all high-variance components regardless of kurtosis value.

  We model each component as a length T vector of Normal iid random variables, for which the distribution of kurtosis values can be approximated. The quantile is estimated based on this distribution.

get_dirs
: Should the projection directions be returned? This is the V matrix in PCA and S matrix in ICA. The default is FALSE to save memory. However, get_dirs==TRUE is required for artifact_images.

full_PCA
: Only applies to the PCA projection. Return the full SVD? Default: FALSE (return only the high-variance components).

get_outliers
: Should outliers be flagged based on cutoff? Default: TRUE.

cutoff
: Median leverage cutoff value. Default: 4.

seed
: Set a seed right before the call to pesel::pesel or ica::icaimax? If NULL, do not set a seed. If numeric (default: 0), will use as the seed.

skip_dimred
: Skip dimension reduction? Default: FALSE.

verbose
: Should occasional updates be printed? Default: FALSE.

Value

A "robdist" object, i.e. a list with components

lwr_50 ...

lwr_80 ...

B_quant ...

Examples

```
library(fastICA)
n_voxels = 2e3
n_timepoints = 35
X = matrix(rnorm(n_timepoints*n_voxels), ncol = n_voxels)

rdx = robdist(X)
```

`rob_stabilize` *Stabilize the center and scale of a timeseries robustly*

Description

Stabilize the center and scale of a timeseries using robust regression of DCT bases on the first and second moments.

Usage

```
rob_stabilize(
  x,
  center = TRUE,
  scale = TRUE,
  lmrob_method = "MM",
  rescale = TRUE
)
```

Arguments

`x`	The timeseries to stabilize.
`center, scale`	Center and scale? Default: `TRUE` for both. If scaling but not centering, the data must already be centered; otherwise, the results will be invalid. Can also be the number of DCT bases to use for robust stabilization of center/scale; `TRUE` will use 4.
`lmrob_method`	The `lmrob_method` argument to `robustbase::lmrob`.
`rescale`	After stabilizing `x`, re-center and re-scale to the original mean and variance? Default: `TRUE`.

Value

the timeseries with its center and scale stabilized

scrub *Data-driven scrubbing*

Description

Performs projection scrubbing or DVARS scrubbing, and optionally thresholds to identify artifactual time points.

Usage

```
scrub(X, method = c("projection", "DVARS"), ...)
```

Arguments

X	A T by V numeric matrix representing an fMRI run. There should not be any missing data (NA or NaN).
method	"projection" (default) or "DVARS"
...	Additional arguments to the specific scrubbing function: see pscrub or DVARS.

Value

A list with components

measure A length T vector or data.frame with T rows, giving the outlyingness measure(s)

measure_info Describes the outlyingness measure(s)

outlier_cutoff The outlier cutoff value(s).

outlier_flag A length T vector or data.frame with T rows, where TRUE indicates suspected outlier presence.

scrub_xifti *Scrub fMRI data in CIFTI format*

Description

Performs projection scrubbing or DVARS scrubbing, and optionally thresholds to identify artifactual time points. Requires ciftiTools and the Connectome Workbench.

Usage

```
scrub_xifti(
  X,
  method = c("projection", "DVARS"),
  brainstructures = c("left", "right"),
  ...
)
```

Arguments

X	Path to a CIFTI file, or a "xifti" object.
method	"projection" or "DVARS"
brainstructures	Character vector indicating which brain structure(s) to use: "left" (left cortical surface), "right" (right cortical surface) and/or "subcortical" (subcortical and cerebellar gray matter). Can also be "all" (obtain all three brain structures). Default: c("left", "right") (excludes the subcortex).
...	Additional arguments to each specific scrubbing function: pscrub or DVARS.

Value

A list with components

measure A length T vector or data.frame with T rows, giving the outlyingness measure(s)

measure_info Describes the outlyingness measure(s)

outlier_cutoff The outlier cutoff value(s).

outlier_flag A length T vector or data.frame with T rows, where TRUE indicates suspected outlier presence.

SHASH_out *Robust outlier detection based on SHASH distribution*

Description

A robust outlier detection based on modeling the data as coming from a SHASH distribution.

Usage

```
SHASH_out(x, maxit = 100, out_lim = 4, weight_init = NULL)
```

Arguments

x	The numeric vector in which to detect outliers.
maxit	The maximum number of iterations. Default: 10.
out_lim	SD threshold for outlier flagging. Default: 4.
weight_init	Initial weights. Default: NULL (no pre-determined outliers).

Value

A "SHASH_out" object, i.e. a list with components

out_idx Indices of the detected outliers.

x_norm The normalized data.

SHASH_coef Coefficients for the SHASH-to-normal transformation.

indx_iters TRUE for the detected outliers for each itertation.

last_iter Last iteration number.

converged Logical indicating whether the convergence criteria was satisfied or not.

Examples

```
x <- rnorm(100) + (seq(100)/200)
x[77] <- 13
SHASH_out(x)
```

SHASH_to_normal *SHASH to normal data transformation*

Description

Transform SHASH-distributed data to normal-distributed data.

Usage

```
SHASH_to_normal(x, mu, sigma, nu, tau, inverse = FALSE)
```

Arguments

x	Numeric vector of data to transform.
mu	Parameter that modulates the mean of x.
sigma	Parameter that modulates the variance of x. Must be greater than zero. This parameter is on the logarithm scale.
nu	Parameter that modulates the skewness of x.
tau	Parameter that modulates the tailweight of x. Must be greater than zero. This parameter is on the logarithm scale.
inverse	Transform normal data to SHASH instead? Default: FALSE.

Value

The transformed data.

summary.scrub_DVARS *Summarize a* "scrub_DVARS" *object*

Description

Summary method for class "scrub_DVARS"

Usage

```
## S3 method for class 'scrub_DVARS'
summary(object, ...)

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

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

Arguments

object	Object of class "scrub_DVARS".
...	further arguments passed to or from other methods.
x	Object of class "scrub_DVARS".

Value

A plot of the scrubbing results

summary.scrub_FD *Summarize a* "scrub_FD" *object*

Description

Summary method for class "scrub_FD"

Usage

```
## S3 method for class 'scrub_FD'
summary(object, ...)

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

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

Arguments

object	Object of class "scrub_FD".
...	further arguments passed to or from other methods.
x	Object of class "scrub_FD".

Value

A plot of the scrubbing results

---

summary.scrub_projection

Summarize a "scrub_projection" *object*

Description

Summary method for class "scrub_projection"

Usage

```
## S3 method for class 'scrub_projection'
summary(object, ...)

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

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

Arguments

object	Object of class "scrub_projection".
...	further arguments passed to or from other methods.
x	Object of class "scrub_projection".

Value

A plot of the scrubbing results

Index

∗ **datasets**
    Dat1, 3
    Dat2, 4

artifact_images, 2, *12*, *16*

Dat1, 3
Dat2, 4
DVARS, 4, *18, 19*

FD, 5
flags_to_nuis_spikes, 7
fMRIscrub, 7

high_kurtosis, 8

leverage, 8

pesel, *12*, *16*
plot.scrub_DVARS, 9
plot.scrub_FD, 10
plot.scrub_projection, 10
print.scrub_DVARS
        *(*summary.scrub_DVARS*)*, 21
print.scrub_FD *(*summary.scrub_FD*)*, 21
print.scrub_projection
        *(*summary.scrub_projection*)*, 22
print.summary.scrub_DVARS
        *(*summary.scrub_DVARS*)*, 21
print.summary.scrub_FD
        *(*summary.scrub_FD*)*, 21
print.summary.scrub_projection
        *(*summary.scrub_projection*)*, 22
pscrub, *2*, 11, *18, 19*

read.table, *6*
rob_stabilize, 17
robdist, 14

scrub, 18
scrub_xifti, 18
SHASH_out, 19
SHASH_to_normal, 20
summary.scrub_DVARS, 21
summary.scrub_FD, 21
summary.scrub_projection, 22