

rstars package

rstars is the R package of the Sequential T-test Analysis of Regime Shifts developed by Rodionov (2004) and then implemented in R language by Stirnimann (*et al.*, 2019). It is an automatic function which allows the detection of shifts in the mean of a given sequence of observation in time without a priori hypothesis. Prewhitening settings are also supported: ordinary least square (OLS) technique to estimate the red noise, Marriott-Pope and Kendall (MPK) and Inverse Proportionality with 4 corrections (IP4) to correct OLS bias.

rstars package works with data frame composed by the first column as “Time” and the following columns as parameters, and each row is a new observation.

Note: This algorithm is built to work with complete (no gaps) time series and with same columns of the same dimension. The presence of NA observations compromises the run of the algorithm. If the NA observations are present, it is advised to remove the gaps.

Usage

```
rstars (data.timeseries = PDO, l.cutoff = 10, pValue = 0.05, Huber = 1, preWhitening = FALSE, OLS = FALSE, MPK = FALSE, IP4 = FALSE, SubsampleSize = ..., save.data = TRUE, show.plot = TRUE, FilteredData = TRUE, save.path = ..., timeseries = TRUE, Endfunction = FALSE)
```

Arguments

Data.timeseries	an R data frame composed by two columns minimum: the first must be “Time”; the following must contain the parameter vectors. Each row is an observation. No gaps, Null or Na value are permitted.
l.cutoff	a numeric value indicating the window where STARS performs the analysis.
pValue	p-value of the statistical analysis which the user wants to test the hypothesis with.
Huber	a numeric value indicating the parameter of Huber’s weight function (Huber, 1964)
preWhitening	logical value indicating if a prewhitening method should be applied to remove red noise from time series. FALSE as default.

SubsampleSize	numeric value indicating the subsample size required during the prewhitening method. SubsampleSize = (cut-off + 1) / 3 as default (Rodionov 2006).
OLS	logical value indicating if the ordinary least square (OLS) technique should be used to estimate the red noise (used only if preWhitening is specified).
MPK	logical value indicating if the Marriott-Pope and Kendall (MPK) should be used to correct OLS bias (used only if preWhitening is specified).
IP4	logical value indicating if the Inverse Proportionality with 4 corrections (IP4) should be used to correct OLS bias (used only if preWhitening is specified).
save.data	logical value indicating if data should be saved on .txt files.
save.path	where to save data. file.choose() as default.
show.plot	logical value indicating if plots should be printed. No autosave.
FilteredData	logical value indicating if filtered data should be shown on plots.
timeseries	if TRUE (default) the “Time” should be in format “yyyy-mm-dd HH-MM-SS”, “yyyy-mm-ss” or “HH-MM-SS”. If FALSE, the vector “Time” can be a different format. If FALSE, the data set is considered just as a data series. This setting was added to the first version of this package to avoid “parsing time” issues.
Endfunction	Logical value indicating if shifts at the end of the time series within the cut-off window should be removed. FALSE as default

Examples

```
#1
Time <- as.character(seq(as.Date("1964-04-01"), as.Date("2018-11-01"), by = "month"))

SST <- c(rnorm(length(Time)/2, mean = 1, 5), rnorm(length(Time)/2, mean = 6, 4) )

require(data.table)
myts <- data.frame(cbind(Time, SST))

myts$SST <- as.numeric(as.character(myts$SST))

rstars (myts, l.cutoff = 120, preWhitening = T, OLS = T)

#2 using PDO data set
rstars (l.cutoff = 10, preWhitening = T, IP4 = T, timeseries = F)

#3
year <- seq(1900, 2000, by = 1)
Jan <- rnorm(length(year), mean = 1, 5)
Feb <- rnorm(length(year), mean = 6, 9)
Mar <- rnorm(length(year), mean = -3, 4)

TS <- as.data.frame(cbind(year, Jan, Feb, Mar))

rstars (TS, l.cutoff = 10, timeseries = F)

#stars_citation()
```