Package 'standaRdized'

April 3, 2019

Type Package

Title Standardized Index Calculation

Version 1.0

Date 2019-04-03

Author Willem Maetens

Maintainer Willem Maetens <w.maetens@vmm.be>

Depends xts, FAdist, glogis

Imports stats, fitdistrplus, lmomco, goftest

Suggests grDevices, graphics

Description This package provides functions for the calculation of daily Standardized Index values such as the Standardized Precipitation Index (SPI), Standardized Precipitation Evapotranspiration Index (SPEI) or Standardized Streamflow Index (SSI)).

License MIT + file LICENSE

RoxygenNote 6.1.1

R topics documented:

standaRdized-package *Standardized Index Calculation*

Description

Functions and methods for the calculation of Standardized Index values

Details

This package provides functions for the calculation of daily Standardized Index values such as the Standardized Precipitation Index (SPI), Standardized Precipitation Evapotranspiration Index (SPEI) or Standardized Streamflow Index (SSI)).

Author(s)

Willem Maetens

Maintainer: Willem Maetens <w.maetens@vmm.be>

fit.distribution *Calculate distribution parameters and statistics*

Description

This function calculates the distribution parameters and a number of statistics for a numeric vector and returns them as a named vector. This function is used internally by the standardized.index function or can be used to provide input to it.

Usage

```
fit.distribution(data, distr, method = c("mle", "lmom"), na.thres = 10)
```

Arguments

data	vector of data
distr	character string specifying the distribution, see details
method	distribution fitting method, see details
na.thres	maximum percentage of NA values allowed in data, default = 10%

Details

Supported distributions are: gamma ('gamma'), 3-parameter gamma ('gamma3'), Weibull ('weibull'), 3-parameter Weibull ('weibull3'), Generalized Extreme Value ('gev'), and Generalized Logistic ('glogis'). Supported distribution fitting methods are: Maximum Likelihood Estimation ('mle', default for 'gamma','weibull','gev' and 'glogis') and L-Moments ('lmom', default for 'gamma3' and 'weibull3'). 'mle' is not supported for distributions 'gamma3' and 'weibull3'. For distr = 'glogis' and method = 'lmom', the 'glo' distribution from package 'lmomco' is used, and its parameters are returned.

Value

a named vector containing:

- fitted distribution parameters, parameters are distribution-specific
- prob.zeroempirical probability of zeros in the data, used in SI calculation with with distributions not including zero such as gamma and Weibull
- n.obsthe total number of observations in the data (including NA values)
- n.nathe number of NA values in the data
- pct.nathe percentage of NA values in the data
- ks.pvalp-value for a two-sided Kolmogorov-Smirnov test that data comes form the fitted distribution
- ad.pvalp-value for a two-sided Anderson-Darling test that data comes form the fitted distribution The data (without NA values) and distr, method and na.thres settings are added to the result as addditional attributes.

See Also

standardized.index, fitplot

Examples

```
data(Ukkel_RR)
# calculate the total rainfall for all months June
monthly.precipitation <- apply.monthly(x=Ukkel_RR,FUN=sum)
data <- c(coredata(monthly.precipitation[format(index(monthly.precipitation),'%m')=='06']))
# fit gamma distribution to the data
fit <- fit.distribution(data=data,distr='gamma',method='mle')
fitplot(x=fit,main='June precipitation',xlab='precipitation (mm)')
# fit gev distribution to the data
fit <- fit.distribution(data=data,distr='gev',method='mle')
fitplot(x=fit,main='June precipitation',xlab='precipitation (mm)')
# fit glogis distribution to the data
fit <- fit.distribution(data=data,distr='glogis',method='mle')
fitplot(x=fit,main='June precipitation',xlab='precipitation (mm)')
```

---

fitplot *Goodness-of-fit Plot*

Description

Function to generate a goodness-of-fit plot for output of the fit.distribution function.

Usage

```
fitplot(x, date = NULL, main = NULL, xlab = 'data', filename = NULL)
```

Arguments

x	output object of the fit.distribution function
date	reference date for the distribution parameters (optional)
main	plot title (optional)
xlab	X-axis label for plot (optional)
filename	if provided, the plot is written to a .png file with this filename, otherwise console output is used

Value

A figure displaying goodness-of-fit information.

See Also

standardized.index, fitplot

Examples

```
data(Ukkel_RR)
# calculate the total rainfall for all months June
monthly.precipitation <- apply.monthly(x=Ukkel_RR,FUN=sum)
data <- c(coredata(monthly.precipitation[format(index(monthly.precipitation),'%m')=='06']))
# fit gamma distribution to the data?pn
fit <- fit.distribution(data=data,distr='gamma',method='mle')
# goodness of fit plot
fitplot(x=fit,main='June precipitation',xlab='precipitation (mm)')
```

---

fprint *Formatted xts Printing*

Description

Print the xts attributes, data head and data tail of xts objects or xts objects in lists in a formatted manner.

Usage

```
fprint(x, nlines = 10)
```

Arguments

x	an xts object to be printed
nlines	approximate number of object rows to print (divided over head and tail of object) with a minimum of 4, if negative, the entire object is printed

See Also

xts

Examples

```
data(Ukkel_RR)
fprint(Ukkel_RR)
```

`get.aggregated.value` *Get Aggregated Value*

Description

Function to get the aggregated value for a single aggregation period

Usage

```
get.aggregated.value(date, data, agg.length, agg.fun = "sum",
  agg.na.thres = 10, agg.interpolation = c("none", "linear", "mean",
  "zeros"), period.warn = TRUE)
```

Arguments

`date`	date for which the aggregated value is determined
`data`	an xts object containing daily data
`agg.length`	length of the aggregation period in days
`agg.fun`	function on x to apply to the aggregation data, default is 'sum'
`agg.na.thres`	threshold for the percentage of na values allowed in the aggregation period data, default = 10
`agg.interpolation`	interpolation type for missing values in individual aggregation period data before applying agg.fun: 'none' (default, NA's are removed), 'linear', 'mean', or 'zeros'
`period.warn`	if TRUE, a warning is when the requested aggregation period contains dates not included in data

Details

This function determines the dates in an aggregation period for standardized index calculation, extracts the corresponding data and applies an aggregation function to the data (default is 'sum', but any function on x can be passed).

Value

A numeric value giving the aggregated value for the aggregation period

See Also

standardized.index, get.reference.values

Examples

```
data(Ukkel_RR)
date <- as.Date('2018-07-01')
# get the aggregated value values for the 30 day-period preceding date
get.aggregated.value(date = date, data = Ukkel_RR, agg.length=30)
```

---

get.reference.values *Get Reference Values for Aggregation Periods*

Description

Function to get the reference values for aggregation periods

Usage

```
get.reference.values(date, ref.data, agg.length, agg.fun = "sum",
  ref.years = NULL, ref.length = 30, agg.na.thres = 10,
  agg.interpolation = c("none", "linear", "mean", "zeros"))
```

Arguments

date	date for which the reference values from other years are determined
ref.data	an xts object containing daily reference data
agg.length	length of the aggregation period in days
agg.fun	function on x to apply to the aggregation data, default is 'sum'
ref.years	years to be used as reference period, default (NULL) is to use all years in ref.data. If ref.years is set to NA, ref.length is used to determine the reference period.
ref.length	if ref.years is null, the ref.length number of years (default = 30) preceding (but not including) the index.out timestamp are used as reference period
agg.na.thres	threshold for the percentage of na values allowed in the aggregation period data, default = 10
agg.interpolation	interpolation type for missing values in individual aggregation period data before applying agg.fun: 'none' (default, NA's are removed), 'linear', 'mean', or 'zeros'

Details

This function determines the dates in each of the aggregation periods for standardized index calculation, extracts the corresponding data from ref.data and applies an aggregation function to the data of each individual aggregation period (default is 'sum', but any function on x can be passed). Reference periods are set by using the ref.years argument to pass specific years to be used as reference period (e.g. for 1981 to 2010, pass seq(1981,2010)). When ref.years is set to NULL, all possible years in ref.data are used. Alternatively, ref.years can be set to NA, in which case ref.length (default = 30) will determine the length of the reference period preceding, but not including, the date for which the reference values are being determined. Warnings will be generated when the requested reference period falls outside ref.data, or when expected aggregation period dates are not present in ref.data.

Value

A named vector with reference period data.

See Also

`standardized.index`, codefit.distribution

Examples

```
data(Ukkel_RR)
date <- as.Date('2018-07-01')
# get all reference values for the 30 day-period preceding date
get.reference.values(date = date, ref.data = Ukkel_RR, agg.length=30)
# get 1981-2010 reference values for the 30 day-period preceding date
get.reference.values(date = date, ref.data = Ukkel_RR, agg.length=30, ref.years=seq(1981,2010))
# get the previos 30 years' reference values for the 30 day-period preceding date
get.reference.values(date = date, ref.data = Ukkel_RR, agg.length=30, ref.years=NA, ref.length=30)
```

---

glo *Generalized Logistic Distribution*

Description

Generalized Logistic Distribution

Usage

```
pglo(q, xi, alpha, kappa)

rglo(n, xi, alpha, kappa)
```

Arguments

q	vector of quantiles
xi	distribution location parameter
alpha	distribution scale parameter
kappa	distribution shape parameter
n	number of observatins

Details

Generalized Logistic Distribution CDF definition as a wrapper for cdfglo analogous to other pdistr functions.

Value

pglo gives the distribution function and rglo generates random deviates.

See Also

cdfglo, fit.distribution

Examples

```
data(Ukkel_RR)
# calculate the total rainfall for all months June
monthly.precipitation <- xts::apply.monthly(x=Ukkel_RR,FUN=sum)
data <- c(coredata(monthly.precipitation[format(index(monthly.precipitation),'\%m')=='06']))
# fit generalized logistic distribution to the data
fit <- fit.distribution(data=data,distr='glo',method='lmom')
fit
```

---

standardized.index *Standardized Index Calculation*

Description

Function to calculate a Standardized Index (SPI, SPEI, SSI,...) for a time series.

Usage

```
standardized.index(data, index.out, agg.length, agg.fun = "sum",
  ref.data = data, distr = c("gamma", "gamma3", "weibull", "weibull3",
  "gev", "glogis"), method = c("mle", "lmom"), params = NULL,
  ks.thres = NULL, ad.thres = NULL, ref.years = NULL,
  ref.length = 30, ref.na.thres = 10, agg.na.thres = ref.na.thres,
  agg.interpolation = c("none", "linear", "mean", "zeros"), digits = 2,
  output.attrs = c("some", "all"))
```

Arguments

data	an xts object containing observed daily data
index.out	vector of dates for which the standardized index is to be calculated
agg.length	length of the aggregation period in days
agg.fun	function on x to apply to the aggregation data, default is 'sum'
ref.data	an xts object containing daily reference data, default is data itself
distr	name of the distribution to be fitted, see details
method	distribution fitting method, 'mle' (default) for maximum likelihood estimation, 'lmom' for L-moments
params	xts containing the reference distribution parameters for each index.out, if not specified (NULL) or missing for specific dates, they are calculated, in which case ref.data should be specified
ks.thres	threshold p-value for the Kolmogrov-Smirnov test, if rejected, the value is set to NA, default= NULL (test not applied)
ad.thres	threshold p-value for the Anderson-Darling test, if rejected, the value is set to NA, default= NULL (test not applied)
ref.years	years to be used as reference period, default is to use all years in ref.data. If NULL, ref.length is used
ref.length	if ref.years is null, the ref.length number of years (default = 30) preceding (but not including) the index.out timestamp are used as reference period
ref.na.thres	threshold for the percentage of NA values allowed in reference period data, default = 10%
agg.na.thres	threshold for the percentage of na values allowed in the aggregation period data, default = ref.na.thres
agg.interpolation	interpolation type for missing values in aggregation data for an individual reference period element: 'none' (default, NA's are removed), 'linear', 'mean', or 'zeros'
digits	number of digits for rounding the resulting standardized index values, default = 2
output.attrs	determines what is attached as xtsAttributes: 'some' (default) adds function settings, and 'all' adds function settings, data and ref.data attributes and the reference values for index.out, aggregation values for index.out, and an xts object containing the fitted parameters

Details

The argument distr can be either: 'gamma' for the gamma distribution, 'gamma3' for the 3-parameter gamma distribution, 'weibull' for the 'weibull distribution, 'weibull3' for the 'weibull distribution, 'gev' for the Generalized Extreme Value distribution, or 'glo' for the shifted log-logistic distribution. The fit.distribution function is used internally to calculate distribution fits, alternatively the parameters for the reference distribution can be supplied via the params argument.

Value

An xts object containing the standardized index values.

See Also

fit.distribution, get.reference.values, fprint

Examples

```
data(Ukkel_RR)
# since this is rainfall data, we are calculating the SPI
# calculate SPI-1 for July 2011, which is approximated by setting agg.length to 30 days
SPI_1 <- standardized.index(data=Ukkel_RR,agg.length=30,index.out=index(Ukkel_RR['2011-07']))
fprint(SPI_1)
# calculate SPI-3 for July 2011, which is approximated by setting agg.length to 90 days
SPI_3 <- standardized.index(data=Ukkel_RR,agg.length=90,index.out=index(Ukkel_RR['2011-07']))
fprint(SPI_3)
```

Ukkel_RR *Ukkel Daily Precipitation*

Description

Daily precipitation (mm/day) for Ukkel from 1898 to 2002.

Format

An xts object on 1880-01-01/2019-02-28 containing:

['value']	num	Precipitation	(mm/day)

With xts attributes:

$ name :	chr	"Ukkel"
$ country :	chr	"Belgium"
$ element :	chr	"RR"
$ unit :	chr	"mm"
$ longitude:	num	4.36638889
$ latitude :	num	50.8
$ elevation:	num	100
$ source :	chr	"ECA\&D (ecad.eu)"

Details

xts object containing daily rainfall at the Ukkel station in Belgium

Source

Royal Meteorological Institute of Belgium (RMI) via https://www.ecad.eu/

Examples

```
data(Ukkel_RR)
str(Ukkel_RR)
fprint(Ukkel_RR)
```

Index