Package 'iki.dataclim'

February 20, 2015

Type Package

Title Consistency, Homogeneity and Summary Statistics of Climatological Data

Version 1.0

Date 2014-07-18

Author Boris Orlowsky

Maintainer Boris Orlowsky <boris@climate-babel.org>

URL www.climate-babel.org

Description The package offers an S4 infrastructure to store climatological station data of various temporal aggregation scales. In-built quality control and homogeneity tests follow the methodology from the European Climate Assessment & Dataset project. Wrappers for climate indices defined by the Expert Team on Climate Change Detection and Indices (ETCCDI), a quick summary of important climate statistics and climate diagram plots provide a fast overview of climatological characteristics of the station.

Depends R (>= 3.0.0)

Imports methods, zoo, lubridate, climdex.pcic (>= 1.1-1), PCICt

NeedsCompilation no

Repository CRAN

Date/Publication 2014-09-19 10:33:43

R topics documented:

`iki.dataclim-package` *Consistency, homogeneity, summary statistics and indices of climate data*

Description

The package offers an S4 infrastructure to store climatological station data of various temporal aggregation scales. In-built quality control and homogeneity tests follow the methodology from the European Climate Assessment & Dataset project. Wrappers for climate indices defined by the Expert Team on Climate Change Detection and Indices (ETCCDI), a quick summary of important climate statistics and climate diagram plots provide a fast overview of climatological characteristics of the station.

Details

Package:	iki.dataclim
Type:	Package
Version:	1.0
Date:	2014-07-18
License:	GPL-3
Depends:	methods

This package bundles part of the code developed for the **dataclim**-project, a collaboration between the German "Gesellschaft fuer Internationale Zusammenarbeit (GIZ)" and the Indonesian weather service BMKG, funded by the German environmental ministery under the "Internationale Klimaschutz Initiative **(iki.)**". Hence the name.

Author(s)

Author: Boris Orlowsky <boris@climate-babel.org>

References

ECA&D: http://eca.knmi.nl/documents/atbd.pdf

ETCCDI climate indices: http://cccma.seos.uvic.ca/ETCCDMI/list_27_indices.shtml,

dataclim project: http://www.giz.de/en/worldwide/16743.html

Examples

```
## A typical work flow could be:

## load the package
library(iki.dataclim)

## load example data and create vector of class Date
data(potsdam)
date <- as.Date(potsdam$date)

## create a dataclim object
myDataclim <- createDataclim(date=date, tmin=potsdam$tmin, tmax=potsdam$tmax,
                          prec=potsdam$prec, basePeriod=1981:2010)

## look at the days with quality issues
slot(myDataclim, "flaggedData")

## evaluate homogeneity of temperature and precipitation
evalHomogeneity(myDataclim)

## look at summary climate statistics for the period 1980-2000
summary(myDataclim, 1980:2000)

## create a Walter-Lieth climate diagram
plotWalterLieth(myDataclim)

## convert the dataclim object to a climdexInput object (from package
## climdex.pcic) and compute a comprehensive set of ETCCDI climate
## indices
myClimdex <- createClimdex(myDataclim, basePeriod=1981:2010)
climdexIndices <- computeClimdex(myClimdex)

## plot the annual timeseries of maximum daily maximum temperature
plot(climdexIndices$annual[, "txx"])
```

`computeClimdex` *Compute a comprehensive set of climdex indices*

Description

This function computes a comprehensive set of indices using the `climdex.pcic` package.

Usage

```
computeClimdex(myClimdex)
```

Arguments

`myClimdex`	A `climdexInput` object.

Details

This function calls a number of functions for annual and monthly indices from the `climdex.pcic` package.

The following annual indices are computed, always using the default options of the `climdex.pcic` package: cdd, cwd, fd, id, r20mm, r99ptot, rx1day, sdii, tn10p, tnn, tr, tx90p, txx, csdi, dtr, gsl, prcptot, r10mm, r95ptot, rnnmm, rx5day, su, tn90p, tnx, tx10p, txn, wsdi.

The following monthly indices are computed, always using the default options of the `climdex.pcic` package: rx1day, tn10p, tnn, tx90p, txx, dtr, rx5day, tn90p, tnx, tx10p, txn.

Value

A list:

`annual.climdex` A zoo object containing a comprehensive set of annual climdex indices.

`monthly.climdex` A zoo object containing a comprehensive set of monthly climdex indices.

Author(s)

Boris Orlowsky <boris@climate-babel.org>

References

http://cccma.seos.uvic.ca/ETCCDMI/list_27_indices.shtml

See Also

`climdex.pcic-package`, `zoo-package`

Examples

```
data(potsdam)
date <- as.Date(potsdam$date)
myDataclim <- createDataclim(date=date, tmin=potsdam$tmin, tmax=potsdam$tmax,
                             prec=potsdam$prec, basePeriod=1981:2010)
myClimdex <- createClimdex(myDataclim, basePeriod=1981:2010)
climdexIndices <- computeClimdex(myClimdex)
```

---

createClimdex *Create a climdex.pcic::climdexInput from a dataclim object*

Description

This function creates a `climdex.pcic::climdexInput` object from a `dataclim` object.

Usage

```
createClimdex(myDataclim, basePeriod)
```

Arguments

myDataclim	A dataclim object.
basePeriod	A numeric vector containing the years of the base period, e.g. 1961:1990. The basePeriod is used for the computation of the quantile threshold needed for some indices in the climdex.pcic package. If NULL (the default), the basePeriod will be extracted from the respective slot of the dataclim object.

Details

This is a simple wrapper to convert a dataclim object to a climdexInput object.

Note that if basePeriod is not a sequence of consecutive years, the range min(basePeriod):max(basePeriod) will be input to the climdex.pcic computations.

Value

A climdexInput object.

Author(s)

Boris Orlowsky <boris@climate-babel.org>

See Also

climdex.pcic-package

Examples

```
data(potsdam)
date <- as.Date(potsdam$date)
myDataclim <- createDataclim(date=date, tmin=potsdam$tmin, tmax=potsdam$tmax,
                             prec=potsdam$prec, basePeriod=1981:2010)
myClimdex <- createClimdex(myDataclim, 1981:2010)
```

---

createDataclim *Create a dataclim object*

Description

Create a dataclim object from daily temperature and precipitation (and other meteorological) data

Usage

```
createDataclim(date = NULL, tmin = NULL, tmax = NULL, prec = NULL,
  basePeriod=1961:1990, convertFlaggedToNA=TRUE, ...)
```

Arguments

date	A vector of class Date.
tmin	A numeric vector of the same length as date.
tmax	A numeric vector of the same length as date.
prec	A numeric vector of the same length as date.
basePeriod	A numeric vector of years. The annual cycle is computed with respect to to these years during the consistency checks. Defaults to the standard climatological period of 1961:1990.
convertFlaggedToNA	Logical, whether detected inconsistent data shall be set to NA. Defaults to TRUE
...	Further vectors of the same length as date. This allows to include observations other than temperature and precipitation.

Details

This generates a dataclim object. It calls the dataclim constructor which performs consistency checks (according to the Algorithm Theoretical Basis Document (ATBD) of the European Climate Assessment & Dataset project (ECA&D), currently only for temperature and precipitation) and aggregates the data to monthly and annual timescales. NAs in the variables are allowed.

Value

An object of class dataclim.

Author(s)

Boris Orlowsky <boris@climate-babel.org>

References

The ATBD: http://www.ecad.eu/documents/atbd.pdf

See Also

See also dataclim.

Examples

```
data(potsdam)
date <- as.Date(potsdam$date)
myDataclim <- createDataclim(date=date, tmin=potsdam$tmin, tmax=potsdam$tmax,
                          prec=potsdam$prec, basePeriod=1981:2010)
```

dataclim-class *Class* "dataclim"

Description

The class provides a standardized container for climate data with in-built Quality Control (QC). Its objects ensure a correct input for the tests and functions of this package.

Objects from the Class

Objects can be created by calls of the form

`new("dataclim", date, tmin, tmax, prec, basePeriod, convertFlaggedToNA, ...)`,

although `createDataclim` is more handy.

Dataclim objects contain monthly and annual aggregates of daily input data after after Qualitiy Control (QC) and padded with NA to ensure complete calendar year coverage. Suspect entries are flagged.

The QC follows closely the Algorithm Theoretical Basis Document (ATBD) of the European Climate Assessment & Dataset project (ECA&D), hosted by the Dutch meteorological service KNMI. Currently, only QCs for temperature and precipitation are implemented.

Slots

`basePeriod`**:** Object of class `"numeric"`. A vector of years. Temperature QC evaluates the deviation from the mean annual cycle computed over these years.

`flagged`**:** Object of class `"list"`. A list with components of each variable, containing the (row-) indices of suspect days according to the QC of the respective variable.

`flaggedData`**:** Object of class `"zoo"` from package `zoo`. A zoo containing all days where QC detected something.

`data`**:** Object of class `"zoo"`. Daily data, where suspect entries are set to NA (if `convertFlaggedToNA` is TRUE).

`monthlyAvg`**,** `monthlyMin`**,** `monthlyMax`**:** Objects of class `"zoo"`. Monthly aggregates of the daily data for months with not more than three NA days. Precipitation in `monthlyAvg` is given as sums.

`annualAvg`**,** `annualMin`**,** `annualMax`**:** Objects of class `"zoo"`. Annual aggregates of the above monthly data. Only years with non-missing monthly data are considered here.

`convertFlaggedToNA`**:** logical, are entries with quality issues set to NA?

Methods

initialize `signature(.Object = "dataclim")`: Constructor of a dataclim object. Performs QC and fills the slots.

show `signature(object = "dataclim")`: Nice display.

summary `signature(object = "dataclim")`: Computes summary statistics from a dataclim objects: means, standard deviations, annual cycles of monthly averages, minimum and maximum and annual and monthly trends (slope and p-value). These statistics are computed over the years given in the optional basePeriod vector (defaults to the `basePeriod`-slot of the `dataclim` object).

Note

Slots can be accessed with slot(object, name). See ?slot for help.

Author(s)

Boris Orlowsky <boris@climate-babel.org>

References

The ATBD: http://www.ecad.eu/documents/atbd.pdf

Examples

```
showClass("dataclim")
```

`ECA&D homogeneity tests of climatological time series`

4 homogeneity tests as employed in the ECA&D project

Description

These functions implement the 4 homogeneity tests of climatological time series described in the Algorithm Theoretical Basis Document (ATBD) of the European Climate Assessment & Dataset project (ECA&D). These are the Standard Normal Homogeneity test (SNHtest), the BuisHand Range test (BHRtest), the PETtitt test (PETtest) and the VON Neumann ratio (VONtest).

Usage

```
SNHtest(X)
BHRtest(X)
PETtest(X)
VONtest(X)
```

Arguments

`X`	A numeric vector of equidistant observations. NAs are not allowed. At least 20 observations are required.

Details

The ATBD describes the test statistics for these 4 tests and provides critical values for significance levels of 1% and 5% for different sample sizes between 20 and 100. Significance is assessed conservatively, i.e. if the actual sample size lies between two of the provided sample sizes in the ATBD, we compare to the critical value of the smaller size. The Von Neumann ratio test (VONtest) only indicates whether the series has a break while the other tests also indicate the approximate location of the break.

Value

`statistic`	Test statistic
`breakpoint`	Index where the break occurs approximately.
`significance`	One of p5, p1 or NS for significant at 5%, at 1% or Not Significant, respectively.

Author(s)

Boris Orlowsky <boris@climate-babel.org>

References

The ATBD: http://www.ecad.eu/documents/atbd.pdf

See Also

`evalHomogeneity`

Examples

```
v <- c(rep(1, 10), rep(2, 10))
SNHtest(v)
```

`evalHomogeneity` *Quality labels for temperature and precipitation observation series*

Description

This function classifies time series as "useful", "doubtful" or "suspect", based on 4 homogeneity tests (see below). It follows the Algorithm Theoretical Basis Document (ATBD) of the European Climate Assessment & Dataset project (ECA&D), hosted by the Dutch meteorological service KNMI.

Usage

```
evalHomogeneity(X)
```

Arguments

`X`	A `dataclim` object or a `data.frame`.

Details

If X is an object of dataclim, the function calls the 4 homogeneity tests Standard Normal Homogeneity test (SNHtest), the BuisHand Range test (BHRtest), the PETtitt test (PETtest) and the VON Neumann ratio (VONtest) and, for temperature, applies them to annual means of daily temperature ranges (Tmax-Tmin) and the day-to-day difference of the daily temperature ranges. For precipitation, the annual counts of days with more than 1mm of rain are evaluated.

If no more than one test indicates a break at the 1% level of significance, the temperature or precipitation observations are considered as "useful". If two tests indicate a break at the 1% level, the respective observations are considered as "doubtful". With three or more tests indicating breaks at the 1% level, the respective observations are considered as "suspect". The two derived temperature characteristics are aggregated by using the maximum number of 1%-significant breaks.

Note that the four tests cannot handle NAs. If X is a dataclim object and the derived annual means have missing values, the function checks whether there are at least 20 valid years and at least 70% of the years are valid for all variables. In this case, the missing values are linearly interpolated.

If X is a data.frame, the function assumes the data.frame to contain equidistant non-NA observations. The series are tested directly, without any processing of derived quantities as in the case of the dataclim object.

Value

A list of

tests	A data.frame with the significance levels (Not Significant, NS - 5% significance level, p5 - 1% significance level, p1) for each test for the (derived annual) variables.
breaks	A data.frame with the approximate locations of the break for each test for the (derived annual) variables (except for the VONtest). For the dataclim method, the years are returned. For the data.frame method, a row-index is returned.
classes	A named vector containing the resulting classification of temperature and precipitation as "useful", "doubtful" or "suspect".

Author(s)

Boris Orlowsky <boris@climate-babel.org>

References

The ATBD: http://www.ecad.eu/documents/atbd.pdf

See Also

SNHtest, BHRtest, PETtest, VONtest

Examples

```
data(potsdam)
date <- as.Date(potsdam$date)
myDataclim <- createDataclim(date=date, tmin=potsdam$tmin, tmax=potsdam$tmax,
```

```
                          prec=potsdam$prec, basePeriod=1981:2010)
evalHomogeneity(myDataclim)
```

`evalHomogeneity-methods`
Methods for Function `evalHomogeneity`

Description

Homogeneity and quality labels for temperature and precipitation observation series

Methods

`signature(X = "dataclim")` This function classifies annual temperature and precipitation series as "useful", "doubtful" or "suspect" based on 4 homogeneity tests. It follows the Algorithm Theoretical Basis Document (ATBD) of the European Climate Assessment & Dataset project (ECA&D), hosted by the Dutch meteorological service KNMI.

`signature(X = "data.frame")` As above, but for a `data.frame` input. Column(s) are supposed to contain equidistant observations. NAs are not allowed. nrow(X) >= 20 is required.

`plotWalterLieth` *Plot a Walter Lieth climate diagram*

Description

This produces a Walter Lieth climate diagram from annual cycles of temperature and precipitation.

Usage

```
plotWalterLieth(X, station="", alt=NA, per, margin=c(4, 4, 5, 4),
pcol="#005ac8", tcol="#e81800", pfcol="#79e6e8", sfcol =
  "#09a0d1", shem=FALSE, ...)
```

Arguments

`X`	A `dataclim` object or a `matrix`.
`station`	A string describing the station. Defaults to the empty string "".
`alt`	The altitude of the station. Defaults to NA.
`per`	A string for the period over which the monthly averages are computed. Defaults to the `basePeriod` slot if X is of class `dataclim` or to the empty string "" if X is a `matrix`.
`margin`	Vector with four numeric elements describing the margins of the plot. Defaults to c(4, 4, 5,4). See `par()`.
`pcol`	Color for (p)recipitation. Defaults to "#005ac8".

tcol	Color for (t)emperature. Defaults to "#e81800".
pfcol	Color (p)robable (f)rost. Defaults to "#79e6e8".
sfcol	Color (s)ure (f)rost. Defaults to "#09a0d1".
shem	Logical. Is the station on the Southern hemisphere? Defaults to FALSE.
...	Further arguments passed to plot().

Details

If X is an object of dataclim, the extracts the monthly averages from a call to summary(X) and produces the Walter Lieth climate diagram.

If X is a matrix, X is expected to be a 4x12 matrix with no NAs, with the first row: monthly precipitation (mm), second row: monthly average maximum daily temperature (degC), third row: monthly average minimum daily temperature (degC), forth row: monthly absolute minimum daily temperature.

Value

Nothing. A plot.

Note

The code is based on code from the former CRAN package climatol by Jose A. Guijarro <jguijarrop@aemet.es>.

Author(s)

Boris Orlowsky <boris@climate-babel.org>

References

Walter, H. & Lieth, H (1960): Klimadiagramm Weltatlas. G. Fischer, Jena.

See Also

dataclim

Examples

```
data(potsdam)
date <- as.Date(potsdam$date)
myDataclim <- createDataclim(date=date, tmin=potsdam$tmin, tmax=potsdam$tmax,
                             prec=potsdam$prec, basePeriod=1981:2010)
plotWalterLieth(myDataclim)
plotWalterLieth(myDataclim, station="Potsdam", alt=80)
```

plotWalterLieth-methods

Methods for Function `plotWalterLieth`

Description

Plot a Walter Lieth climate diagram

Methods

`signature(X = "dataclim")` This produces a Walter Lieth climate diagram from annual cycles of temperature and precipitation.

`signature(X = "matrix")` As above, but for a `matrix` input. This must be a 4x12 matrix, with first row: monthly precipitation (mm), second row: monthly average maximum daily temperature (degC), third row: monthly average minimum daily temperature (degC), forth row: monthly absolute minimum daily temperature. NAs are not allowed.

Note

The code is based on code from the former CRAN package `climatol` by Jose A. Guijarro <jguijarrop@aemet.es>.

References

Walter, H. & Lieth, H (1960): Klimadiagramm Weltatlas. G. Fischer, Jena.

potsdam

Example data from the Potsdam weather station

Description

The `potsdam` data frame contains daily data of maximum, average and minimum air temperature, precipitation and air pressure (1979-2013) from the weather station at Potsdam, Germany

Usage

```
data(potsdam)
```

Format

This data frame contains the following columns:

date a character vector; calendar date (yyyy-mm-dd)

tmax a numeric vector; daily maximum air temperature in deg. C

tavg a numeric vector; daily average air temperature in deg. C

tmin a numeric vector; daily minimum air temperature in deg. C

prec a numeric vector; daily precipitation in mm

press a numeric vector; daily air pressure in hPa

Details

The original data comes from ECA&D (see link below) and was assembled into a readable csv file. Note that this data must not be used for commercial purposes, see http://eca.knmi.nl/documents/ECAD_datapolicy.pdf. Quote from the document: "These data, which include many GCOS-defined Essential Climate Variables (ECVs) for the atmosphere near the surface, are strictly for use in non-commercial research and education projects only."

Source

http://eca.knmi.nl/dailydata/customquery.php

Examples

```
data(potsdam)
names(potsdam)
head(potsdam)
str(potsdam)
```

---

zoo-class *Class* "zoo"

Description

This is the class zoo from package zoo, made available for use in S4 classes.

Objects from the Class

A virtual Class: No objects may be created from it.

Slots

.S3Class: Object of class "character" ~~

Extends

Class "oldClass", directly.

Methods

No methods defined with class "zoo" in the signature.

Author(s)

Boris Orlowsky <boris@climate-babel.org>

Examples

```
showClass("zoo")
```

Index