Package ‘treeclim’

November 26, 2015

Type Package
Title Numerical Calibration of Proxy-Climate Relationships
Version 1.0.13
Date 2015-11-26
Depends R (>= 3.0.2), base, stats, utils, Rcpp (>= 0.10.6)
Imports abind, plyr, ggplot2, lmtest, lmodel2, np, boot
LinkingTo Rcpp, RcppArmadillo (>= 0.4.000.2)
Suggests dplR, testthat (>= 0.7.1.99)
Description Bootstrapped response and correlation functions,
seasonal correlations and evaluation of reconstruction
skills for use in dendroclimatology and dendroecology.

NeedsCompilation yes
License GPL-3
LazyData yes
RoxygenNote 5.0.1
Author Christian Zang [aut, cre, cph, trl],
Franco Biondi [ctb, cph]
Maintainer Christian Zang <christian.zang@wzw.tum.de>
Repository CRAN
Date/Publication 2015-11-26 14:25:51

R topics documented:

.range ................................................................. 2
dcc ................................................................. 3
exclude_from ...................................................... 6
g_test .............................................................. 7
muc_clim .......................................................... 8
muc_fake .......................................................... 8
muc_spruce ......................................................... 9
.range

Modifiers for climate parameter selection

Description

These modifiers are used to select specific months from specific climate parameters, and potentially transform the selections into their respective sums or means. The modifiers can be chained together using ‘+’.

Usage

.range(.months = NULL, .variables = NULL)

.mean(.months = NULL, .variables = NULL)

.sum(.months = NULL, .variables = NULL)

Arguments

.months numeric identifiers for the months (-1 for previous January until 12 for current December, with -6 for previous June, etc.)

.variables names of the variables the modifier shall be applied to
Description

This function calculates (potentially moving) response and correlation functions from tree-ring chronologies and monthly climatic data. For the moving case, the calculation is performed repeatedly for consecutive time windows. Function parameters may be bootstrapped to calculate their significance and confidence intervals.

Usage

dcc(chrono, climate, selection = -6:9, method = "response", moving = FALSE, win_size = 25, win_offset = 1, start_last = TRUE, timespan = NULL, var_names = NULL, ci = 0.05, boot = "stationary", sb = TRUE)

Arguments

chrono data.frame containing a tree-ring chronologies, e.g. as obtained by chron of package dplR.
climate either a data.frame or matrix with climatic data in monthly resolution, with year, month and climate parameters in columns (all columns except year and month will be recognized as parameters for response or correlation function), or a single data.frame or matrix in 13-column format (see below), or list of several of the latter.
selection either a numeric vector, a modifier, or a chain of modifiers specifying the parameter selection for the model (see Details).
method character string specifying the calculation method. Possible values are “response” and “correlation”. Partial strings are ok.
moving logical; should the analysis be carried out in moving windows. Defaults to FALSE.
win_size integer giving the window size for each recalculation in years.
win_offset integer giving the number of years between each window start in years.
start_last logical flag indicating whether the first window should start at the rear end (youngest part of the series) or not.
timespan integer vector of length 2 specifying the time interval (in years) to be considered for analysis. Defaults to the maximum possible interval.
var_names character vector with variable names. Defaults to corresponding column names of data.frame clim.
ci numerical value to set the test level for significance test (values 0.01, 0.05 and 0.1 are allowed); the confidence intervals are adapted accordingly.
boot character indicating which bootstrap method should be used, one of c("stationary", "std", "exact")
sb logical flag indicating whether textual status bar for moving case should be suppressed. Suppression is recommended for e.g. Sweave files.
Details

This function builds upon and extends the functionality of programme DENDROCLIM2002 (Biondi and Waikul, 2004), and will calculate bootstrapped (and non-bootstrapped) moving and static response and correlation functions in a similar fashion as described in the above mentioned paper. Important extensions include a very flexible parameter selection model (see below), the possibility to use an unlimited number of climate parameters, and the option to use exact bootstrapping.

Input chronology data can be a data.frame such as produced by function chron of package dplR. It has to be a data.frame with at least one column containing the tree-ring indices, and the corresponding years as rownames.

For climatic input data, there are three possibilities: Firstly, input climatic data can be a data.frame or matrix consisting of at least 3 rows for years, months and at least one climate parameter in the given order. Secondly, input climatic data can be a single data.frame or matrix in the style of the original DENDROCLIM2002 input data, i.e. one parameter with 12 months in one row, where the first column represents the year. Or thirdly, input climatic data can be a list of several of the latter described data.frame or matrices. As an internal format dispatcher checks the format automatically, it is absolutely necessary that in all three cases, only complete years (months 1-12) are provided. It is not possible to mix different formats in one go.

Parameters can be selected with the ‘selection’ parameter in two different ways. The default value is -6:9. This is equivalent to the standard settings in DENDROCLIM2002 and bootRes, and selects from all climate variables all months from previous year’s June (-6, previous year’s months are specified as negative integers) to current years September (9, months of the current year are specified as positive integers) as model parameters.

More complex parameter selections can be obtained by the modifiers provided in treeclim: .range, .mean, and .sum. .range corresponds the example above, where all specified months are used, while .sum and .mean will use the sums and means of the specified months. These modifiers also allow to select specific climatic variables, addressed by name. Thus, .mean(4:8,”temp”) will select the mean for climate parameter ”temp” for the months April to August. Not only ranges, but also individual vectors can be used for month specification, like e.g., .range(c(1,3,4,5)).

The modifiers can be chained together using the ‘+’ symbol, which makes it possible to create arbitrarily complex selections of climate parameters for calibration. E.g., .mean(2:5,”temp”) + .sum(2:5,”prec”) will yield the February-to-May mean for the variable ”temp” and the sum of the variable ”prec” for the same time. While there is no limitation for number of lists that can be chained together, ‘dcc’ will not check for meaningful specifications. Testing smart hypotheses is up the researcher.

For the exclusion of months, the convenience function excludefrom() (or short exfr()) is provided. E.g., .range(excludefrom(-6:10, -11:3)) will yield the monthly values of all parameters for the months previous June (-6) to current October (10), but without the months previous November (-11) to current March (3) in between. While it is also possible to supply arbitrary vectors as month specification, and not only ranges as shown in most of the examples here, this way of excluding e.g. the dormant season is far more convenient.

1000 bootstrap samples are taken from the original distributions of climate and tree-ring data, either using the stationary bootstrap (Politis and Romano 1994, boot = ”stationary”) or classical bootstrap (DENDROCLIM2002-style, boot = ”std”). The stationary bootstrap mimics the stationary properties of the original time series in the resampled time series by resampling within blocks. Within each block, the number of observations is random and has a geometric distribution. Consequently, the choice of the distribution parameter will affect the autocorrelation structure of the resampled time series. Optimal (expected) block length is chosen according to Politis and White
(2004). In the case of response function analysis, an eigen decomposition of the standardized predictor matrix is performed. Nonrelevant eigenvectors are removed using the PVP criterion (Guiot, 1990), principal component scores are then calculated from the matrices of reduced eigenvectors and standardized climatic predictors. Response coefficients are found via singular value decomposition, and tested for significance using the 95% percentile range method (Dixon, 2001). In case of correlation function analysis, the coefficients are Pearson’s correlation coefficients. The same method for significance testing is applied.

There is also the option to use exact bootstrapping like implemented in seascorr (Meko et al. 2011, `boot = "exact"`). In this case, circulant embedding is used to simulate the tree-ring data 1000 times as time series with the same frequency characteristics like the original time-series (Percival & Constantine, 2006). Empirical non-exceedence probabilities are used to test the coefficients of the response/correlation function with the original data for significance. For the exact bootstrapping case, no confidence intervals for the response/correlation coefficients can be computed.

**Value**

'dcc' returns an 'object' of class "tc_dcc".

The functions 'summary' and 'plot' are used to obtain and print a summary of the results, and to create a plot. The function 'coef' can be used to extract the coefficients.

An object of class "tc_dcc" is a list containing at least the following components:

- **call**: the call made to function 'dcc'
- **coef**: the coefficients, themselves being an object of class 'tc_coef' for the static case, and of class 'tc_mcoef' for the moving case. Objects of class 'tc_coef' are single data.frames, while objects of class 'tc_mcoef' are lists of separate data.frames for the coefficients ($coef), upper and lower confidence interval ($ci_upper and $ci_lower), and significance flags ($significant)
- **design**: the design matrix on which this call to 'dcc' operates
- **truncated**: the input data truncated to the common timespan or the specified timespan
- **original**: the original input data, with the climate data being recast into a single data.frame

**Author(s)**

Christian Zang; the original MATLAB code for exact bootstrapping was written by Dave Meko

**References**


Politis, D.N. and H. White (2004), Automatic block-length selection for the dependent bootstrap, Econometric Reviews 23(1), 53-70.

Examples

```
## Not run:
de_resp <- dcc(muc_spruce, muc_clim)

## End(Not run)
```

---

**exclude_from**

Exclude months from analysis

Description

Single months or ranges of months can be excluded from analysis. This is helpful for e.g. excluding winter months without cambial activity.

Usage

```
exclude_from(month, exclude = NULL)
exfr(month, exclude = NULL)
```

Arguments

- `month`: range of numeric month ids
- `exclude`: range or set of months to exclude

Value

A reduced set of numeric month ids

See Also

- link{.range}, link{.mean}, link{.sum}

Examples

```
exfr(-5:10, -10:3)
```
Gershunov test for spurious low-frequency modulations

Description

This function provides a test to decide whether low-frequency modulations in the relationship between climate and tree-growth are significantly stronger or weaker than could be expected by chance.

Usage

```r
g_test(x, boot = FALSE, ci = 0.05, sb = TRUE)
```

Arguments

- `x` an object of class "tc_dcc" as returned from a call to `dcc` with moving correlations enabled
- `boot` logical shall the individual correlation be bootstrapped? (see details)
- `ci` Significance level for testing, one of c(0.01, 0.05, 0.1)
- `sb` logical shall a status bar be drawn?

Details

This function is a multivariate extension of the test for spurious low-frequency modulations for moving correlations of time series as proposed by Gershunov et al. (2001). In short, 1000 simulations of random data sets are generated, where the climate data is simulated as Gaussian noise, and the tree-data as linear combinations of the climate parameters using the original coefficients of the correlation function, and an error component with a variance equal to the variance unexplained by the individual parameters.

For each iteration, a moving correlation function is calculated with exactly the same settings as the original model. The standard deviation over the individual windows for each parameter is then compared to the bootstrapped distribution of the standard deviation of the simulated data to test for significantly higher or lower low-frequency modulations.

Value

A `data.frame` with columns for variable identification and logical flags for whether the low-frequency modulation of the correlation of this variable with tree-growth is higher than expected by chance, lower than expected by chance, or can be considered as noise.

References

Examples

```r
## Not run:
dc_cor <- dcc(muc_spruce, muc_clim, 3:9, method = "cor", moving = TRUE)
g_test(dc_cor)
## End(Not run)
```

data(muc_clim)

Description

This dataset gives the monthly mean temperature and total precipitation at Forstenrieder Park, Munich, Bavaria, Germany.

Usage

data(muc_clim)

Format

A data.frame containing four columns with year, month, temperature and precipitation.

Source


data(muc_fake)

Description

This dataset gives the modelled tree-ring widths for Picea abies at Forstenrieder Park, Munich, Bavaria, Germany. Tree growth was modeled as a response of low temperatures in previous and current July and August, high temperatures in current February and March, and high precipitation amounts in current July and August.

Usage

data(muc_fake)

Format

A data.frame containing tree-ring indices and replication depth with respective years as rownames.
muc_spruce

Source


---

Description

This dataset gives the tree-ring indices for *Picea abies* at Forstenrieder Park, Munich, Bavaria, Germany. The chronology represents 20 cores from 10 trees. The series were read in using `read.rwl` from package dp1R, detrended using a 60a spline with 50% frequency cutoff (function `detrend`), and averaged to a chronology using a robust mean `chron`.

Usage

data(muc_spruce)

Format

A data frame containing tree-ring indices and replication depth with respective years as rownames.

Source


norw01U

Tree-Ring Chronology of a Pine Population at Visdalen, Norway

Description

This dataset gives the tree-ring indices for *Pinus sylvestris* at a site near Visdalen, Norway. The chronology represents 34 dated series and spans the period between 1600 and 1983. The raw measurements were obtained from the ITRDB (see source), and read in, detrended with cubic splines (frequency cutoff of 50 percent at two thirds of the series length) and averaged using `read.rwl`, `detrend`, and `chron` from package dp1R.

Usage

data(norw01U)

Format

A data frame containing tree-ring indices and replication depth with respective years as rownames.
### Source


Accessed 2014/07/03.

---

<table>
<thead>
<tr>
<th>norway_prec</th>
<th>Monthly Precipitation Sums for Norway</th>
</tr>
</thead>
</table>

### Description

This data set gives the monthly precipitation sums at country level for Norway in a decadal (DENDROCLIM2002-style) format from the TYN CY 1.1 data set.

#### Usage

```r
data(norway_prec)
```

#### Format

A `data.frame` containing thirteen columns with year and twelve months of temperature data in degree Celsius.

### Source


---

<table>
<thead>
<tr>
<th>norway_temp</th>
<th>Monthly Temperature Means for Norway</th>
</tr>
</thead>
</table>

### Description

This data set gives the monthly temperature means at country level for Norway in a decadal (DENDROCLIM2002-style) format from the TYN CY 1.1 data set.

#### Usage

```r
data(norway_temp)
```

#### Format

A `data.frame` containing thirteen columns with year and twelve months of temperature data in degree Celsius.
**rt_prec**

**Source**

**Description**
This dataset gives the monthly precipitation sum at Rothenburg ob der Tauber, Bavaria, Germany in a decadal (DENDROCLIM2002-style) format.

**Usage**
data(rt_prec)

**Format**
A data.frame containing thirteen columns with year and twelve months of precipitation data in mm rainfall.

**Source**

**rt_spruce**

**Description**
This dataset gives the tree-ring indices for *Picea abies* at Rothenburg ob der Tauber, Bavaria, Germany. The chronology represents 20 cores from 10 trees. The series were read in using `read.rwl` from package `dplR`, detrended using a 60a spline with 50% frequency cutoff (function `detrend`), and averaged to a chronology using a robust mean `chron`.

**Usage**
data(rt_spruce)

**Format**
A data.frame containing tree-ring indices and replication depth with respective years as rownames.
**Source**


---

**rt_temp**  
*Monthly Temperature Means for Rothenburg ob der Tauber, Germany*

**Description**

This dataset gives the monthly temperature means at Rothenburg ob der Tauber, Bavaria, Germany in a decadal (DENDROCLIM2002-style) format.

**Usage**

```r
data(rt_temp)
```

**Format**

A `data.frame` containing thirteen columns with year and twelve months of temperature data in degree Celsius.

**Source**


---

**seascorr**  
*Seasonal (partial) correlation analysis*

**Description**

Calculate seasonal correlation with primary and secondary climate variables and tree-ring data, similar to the seascorr function for MATLAB.

**Usage**

```r
seascorr(chrono, climate, var_names = NULL, timespan = NULL, complete = 9, season_lengths = c(1, 3, 6), primary = 1, secondary = 2, ci = 0.95)
```
Arguments

- **chrono**: data.frame containing a tree-ring chronologies, e.g. as obtained by `chron` of package `dplR`.
- **climate**: either a data.frame or matrix with climatic data in monthly resolution, with year, month and climate parameters in columns (all columns except year and month will be recognized as parameters for response or correlation function), or a single data.frame or matrix in 13-column format (see below), or a list of several of the latter.
- **var_names**: character vector with variable names. Defaults to corresponding column names of data.frame clim.
- **timespan**: integer vector of length 2 specifying the time interval (in years) to be considered for analysis. Defaults to the maximum possible interval.
- **complete**: integer scalar, month when tree-ring growth is expected to have finished.
- **season_lengths**: numeric vector giving the lengths of the seasons for variable grouping
- **primary** and **secondary**: position numeric or name character of primary and secondary climate variable
- **ci**: numeric scalar to set the test level for significance test (values 0.01, 0.05 and 0.1 are allowed); the confidence intervals are adapted accordingly.

Details

This function mimics the behaviour of the MATLAB function `seascorr` (Meko et al. 2011), which calculates partial correlations of tree-ring data with a primary and a secondary climatic variable for seasons of different lengths.

Input chronology data can be a data.frame such as produced by function `chron` of package `dplR`. It has to be a data.frame with at least one column containing the tree-ring indices, and the corresponding years as rownames.

For climatic input data, there are three possibilities: Firstly, input climatic data can be a data.frame or matrix consisting of at least 3 rows for years, months and at least one climate parameter in the given order. Secondly, input climatic data can be a single data.frame or matrix in the style of the original DENDROCLIM2002 input data, i.e. one parameter with 12 months in one row, where the first column represents the year. Or thirdly, input climatic data can be a list of several of the latter described data.frame or matrices. As an internal format dispatcher checks the format automatically, it is absolutely necessary that in all three cases, only complete years (months 1-12) are provided. It is not possible to mix different formats in one go.

The 'complete' parameter specifies the months of the current year in which tree-growth is assumed to finish. This month marks the last month of the first season, and starting from here, 14 different seasons are computed for each specified season length in one-month steps. E.g., for a starting value of 9 (current September) and season length of 3 months, the first season comprises current July to current September, the second season comprises current June to current August, and the last season comprises previous June to previous August. This results in 14 seasons for a given season length. An arbitrary number of season lengths can be specified.

The choice for primary vs. secondary variable can be made either via numeric selection (the integer value 1 stands for the first variable in the supplied climate data set), or by name ("temp", when one
of the variables is named "temp"). The correlation of the primary variable with tree-growth is computed as the simple (Pearson) correlation coefficient, while the influence of the secondary variable on tree-growth is computed with the influence of the primary variable on tree-growth removed.

Like in the original seascorr program, the significance of each (partial) correlation is evaluated using exact bootstrapping by circulant embedding of the tree-ring data (Percival & Constantine, 2006).

Value

'seascorr' returns an 'object' of class ""tc_seascorr"".

The 'plot' function is used to obtain a plot of the results.

An object of class ""tc_seascorr"" is a list containing at least the following components:

call the call made to 'seascorr'
seasons a list of length n, where n is the number of season lengths provided; each list element consists of a data.frame with end month, correlation coefficient and significance flag
truncated the input data truncated to the common timespan or the specified timespan
original the original input data, with the climate data being recast into a single data.frame

Author(s)

Christian Zang; the procedure incl. exact bootstrapping was implemented first by Dave Meko in MA

References


Examples

sc <- seascorr(muc_fakeL, muc_clim)
sc
plot(sc)
Evaluate reconstruction skills using split-calibration

Description

This function allows to evaluate the reconstruction skills for a given proxy time series in split-calibration approach.

Usage

skills(object, formula, model = "ols", calibration = "50%",
       timespan = NULL)

Arguments

- **object**: an object of class "tc_dcc" or "tc_seascorr"
- **formula**: a formula specifying the regressand, optionally omitting the intercept, see details
- **model**: one of "ols" or "rma"
- **calibration**: which part of the data shall be used as calibration subset? Given as either a range of years, an integer corresponding to the first or last number of observations, or a percentage as character string corresponding to the part of the data set to be used as calibration subset.
- **timespan**: timespan to be used to truncate the data

Details

The result of a call to **dcc** or **seascorr** can be used as object for the function. The required data is then taken from this object and no further processing of the tree and climate variables has to be done for by the user. This reflects the flow of analysis, where first general climate/growth relations are explored, and then the strongest ones are deployed for reconstruction purposes.

**formula** is a one-sided formula of the form ~ selection, where ‘selection’ is an aggregation modifier (one of .mean, .sum, and .range). The user should be aware of the fact that in case the aggregation modifier evaluates to more than one variable (e.g., summer means for both temperature and precipitation), a warning message is issued, and only the first variable is taken into consideration for evaluating the reconstruction skills.

The type of regression model (ordinary least squares or errors-in-variables via reduced major axis regression) can be selected.

The part of the data to be used as a calibration subset can be specified in three different ways: 1) as a range of years, these are then taken as calibration period; 2) as a single integer, if positive, this number of observations at the recent end of the data set is taken as calibration set, if negative, this number of oldest observations is taken; and 3) as a character string giving a percentage of values, e.g., “-40%” would select the 40% oldest observations, while “55%” would select the 55% most recent ones.
The relationship between climate and tree-ring data is evaluated for the calibration period and the complete data set. Frequently used verification statistics are computed: reduction of error (RE), coefficient of efficiency (CE), and the Durban-Watson statistic (DW) (Cook et al. 1994, Durbin and Watson, 1951).

Value

'skills' returns an 'object' of class "tc_skills".

An object of class "tc_skills" is a list containing at least the following components:

- `call`: the call made to function 'skills'
- `r.cal`: the coefficient of correlation for the calibration timespan
- `r.full`: the coefficient of correlation for the complete data set
- `coef.cal`: regression coefficients for the calibration model
- `coef.full`: regression coefficients for the full model
- `p.cal`: significance for the calibration model
- `p.full`: significance for the full model
- `RE`: reduction of error statistic
- `CE`: coefficient of efficiency statistic
- `DW`: Durbin-Watson statistic
- `cal.model`: the complete calibration model (an object of class 'lmodel2')
- `full.model`: the complete full model (an object of class 'lmodel2')

References


Examples

```r
## Not run:
dc <- dcc(muc_fakeL muc_climL Nmean(6:9L "temp") + Nsum(6:9L "prec"))
sk <- skills(dc)
sk
plot(sk)
## End(Not run)
```
**spai020**

*Tree-Ring Chronology of a Pine Population at Penota, Spain*

**Description**

This dataset gives the tree-ring indices for *Pinus sylvestris* at a site near Penota, Spain. The chronology represents 18 dated series and spans the period between 1763 and 1991. The chronology was obtained from the ITRDB (see source), and read in using *read.crn* from package dp1R.

**Usage**

```r
data(spai020)
```

**Format**

A data frame containing tree-ring indices and replication depth with respective years as rownames.

**Source**


---

**spain_prec**

*Monthly Precipitation Sums for Spain*

**Description**

This data set gives the monthly precipitation sums at country level for Spain in a decadal (DENDROCLIM2002-style) format from the TYN CY 1.1 data set.

**Usage**

```r
data(spain_prec)
```

**Format**

A data frame containing thirteen columns with year and twelve months of temperature data in degree Celsius.

**Source**

**Description**

This data set gives the monthly temperature means at country level for Spain in a decadal (DENDROCLIM2002-style) format from the TYN CY 1.1 data set.

**Usage**

data(spain_temp)

**Format**

A data.frame containing thirteen columns with year and twelve months of temperature data in degree Celsius.

**Source**


---

**tc_design**

Deparse list structure from month specification and return parameter set

**Description**

Deparse list structure from month specification into single calls, call an aggregating function to collect the data, and return a parameter set for calibration. In general, the months must be specified as a list or nested list structure. Possible cases are: list() list(...) list(list(...)) list(list(...), list(...), ...)

**Usage**

tc_design(selection, climate)

**Arguments**

- **selection**: the list structure to specify the month selection
- **climate**: the climate data as returned by make_pmat

**Value**

a data.frame
# Index

**Topic datasets**
- muc_clim, 8
- muc_fake, 8
- muc_spruce, 9
- norw015, 9
- norway_prec, 10
- norway_temp, 10
- rt_prec, 11
- rt_spruce, 11
- rt_temp, 12
- spai020, 17
- spain_prec, 17
- spain_temp, 18

**Topic manip**
- exclude_from, 6

**Topic test**
- g_test, 7
  - .mean, 15
  - .mean(.range), 2
  - range, 2, 15
  - .sum, 15
  - .sum(.range), 2

- chron, 9, 11
- dcc, 3, 7, 15
- detrend, 9, 11
- seascorr, 12, 15
- skills, 15
- spai020, 17
- spain_prec, 17
- spain_temp, 18
- tc_design, 18

- read.crn, 17
- read.rwl, 9, 11
- rt_prec, 11
- rt_spruce, 11
- rt_temp, 12