Package 'GNSSseg'

June 2, 2020

Type Package

Title Homogenization of GNSS Series

Version 6.0

Date 2020-05-12

Author Annarosa Quarello [aut, cre], Emilie Lebarbier [aut], Olivier Bock [aut]

Maintainer Annarosa Quarello <quarello.annarosa@gmail.com>

Description Homogenize GNSS (Global Navigation Satellite System) time-series. The general model is a segmentation in the mean model including a periodic function and considering monthly variances, see Quarello (2020) <arXiv:2005.04683>.

LazyData true

License GPL-3

Encoding UTF-8

Imports robustbase, capushe

RoxygenNote 7.0.2

Suggests knitr, rmarkdown

NeedsCompilation no

Repository CRAN

Date/Publication 2020-06-02 15:20:16 UTC

R topics documented:

Data   *Example of data*

Description

A data frame [n x 2] containing a simulated Gaussian series for the two years 1995 and 1996, with size n=731. 3 changes are considered at positions 100, 150 and 500 (or at the dates 1995-04-10, 1995-05-30 and 1996-05-14). The means of the segments alternates between 0 and 1 (beginning by 0). The functional part is 0.4*cos(2*pi*time/lyear) where lyear is 365.25 and time is centred according to the first date and expressed in seconds: time=(date-date[1])/86400. The standard deviation of the noise of the 12 months are drawn from an uniform distribution between 0.1 and 0.8. The date is expressed as yyyy-mm-dd in the "calendar time" (class POSIXct).

Usage

```
data(Data)
```

Format

A data frame with 731 observations on the following 2 variables.

signal a numeric vector

date a date vector expressed as yyyy-mm-dd in the "calendar time" (class POSIXct)

Details

signal: the values of the observed signal; date: the dates in calendar time

Examples

```
library(GNSSseg)
data(Data)
class(Data$date)
plot(Data$date,Data$signal,type="l")
```

GNSSseg   *Homogeneization of GNSS series*

Description

fit a segmentation in the mean model by taken into account for a functional part and a heterogeneous variance (default is monthly)

Usage

```
GNSSseg(
  Data,
  lyear = 365.25,
  lmin = 1,
  Kmax = 30,
  selection.K = "BM_BJ",
  S = 0.75,
  f = TRUE,
  selection.f = FALSE,
  threshold = 0.001,
  tol = 1e-04
)
```

Arguments

Data	a data frame, with size [n x 2], containing the signal (e.g. the daily GPS-ERAI series for GNSS) and the dates (in format yyyy-mm-dd of type "calendar time" (class POSIXct))
lyear	the length of the year in the signal. Default is 365.25
lmin	the minimum length of the segments. Default is 1
Kmax	the maximal number of segments (must be lower than n). Default is 30
selection.K	a name indicating the model selection criterion to select the number of segments K (mBIC, Lav, BM_BJ or BM_slope). "none" indicates that no selection is claimed and the procedure considers Kmax segments or Kmax-1 changes. If selection.K="All", the results for the four possible criteria are given. Default is "BM_BJ"
S	the threshold used in the Lav's criterion. Default is 0.75
f	a boolean indicating if the functional part is taking into account in the model. Default is TRUE and note that if f=FALSE, only a segmentation is performed
selection.f	a boolean indicating if a selection on the functions of the Fourier decomposition of order 4 is performed. Default is FALSE
threshold	a numeric value lower than 1 used for the selection of the functions of the Fourier decomposition of order 4. Default is 0.001
tol	the stopping rule for the iterative procedure. Default is 1e-4

Details

The function performs homogeneization of GNSS series. The considered model is such that: (1) the average is composed of a piecewise function (changes in the mean) with a functional part and (2) the variance is heterogeneous on fixed intervals. By default the latter intervals are the months. The inference procedure consists in two steps. First, the number of segments is fixed to Kmax and the parameters are estimated using the maximum likelihood procedure using the following procedure: first the variances are robustly estimated and then the segmentation and the functional parts are iteratively estimated. Then the number of segments is chosen using model selection criteria. The possible criteria are mBIC the modified BIC criterion, Lav the criterion proposed by Lavielle, BM_BJ

and `BM_slope` the criteriain which the penalty constant is calibrated using the Biggest Jump and the slope.

- The data is a data frame with 2 columns: $signal is the signal to be homogeneized (a daily series) and $date is the date. The date will be in format yyyy-mm-dd of type "calendar time" (class POSIXct).

- The function part is estimated using a Fourier decomposition of order 4 with `selection.f=FALSE`. `selection.f=TRUE` consists in selecting the significative functions of the Fourier decomposition of order 4 (for which p.values are lower than `threshold`)

- If `selection.K="none"`, the procedure is performed with Kmax segments.

- Missing data in the signal are accepted.

Value

A file containing

- `K` that corresponds to the selected number of segments or K-1 corresponds to the number of changes. If `selection.K="none"`, the number of segments is Kmax.

- `seg` that corresponds to the estimation of the segmentation parameters (the begin and the end positions of each segment with the estimated mean).

- `funct` that corresponds to the estimation of the functional part. If f==FALSE, `funct` is FALSE

- `coeff` that corresponds to the estimation of the coefficients of the Fourier decomposition. The vector contains 8 coefficients if `selection.f=FALSE` or as many coefficients as the number of selected functions if `selection.f=TRUE`. If f==FALSE, `coeff` is FALSE

- `variances` that corresponds to the estimated variances of each fixed interval

- `SSR` that corresponds to the Residuals Sum of Squares for k=1,...,Kmax. If `selection.K="none"`, it contains only the SSR for Kmax segments

- `Tot` is a list. Each component contains all the results k segments (k=1,...,Kmax). If `selection.K="none"`, `Tot` is NA

If `selection.K="All"`, the outputs `K`, `seg`, `funct` and `coeff` are each a list containing the corresponding results obtained for the four model selection criteria

Examples

```
data(Data)
lyear=365.25
Kmax=4
lmin=1
result=GNSSseg(Data,lyear,Kmax=Kmax,selection.K="none")
plot_GNSS(Data,result$seg,result$funct)
```

---

plot_GNSS   *Graph of the result*

Description

plot the signal with the estimated average

Usage

```
plot_GNSS(Data, segmentation, functional)
```

Arguments

Data   a data frame, with size [n x 2], containing the signal (e.g. the daily GPS-ERAI
                  series for GNSS) and the dates (in format yyyy-mm-dd of type "calendar time"
                  (class POSIXct))

segmentation   the estimated segmentation (result of the GNSSseg function)

functional   the estimated functional (result of the GNSSseg function)

Details

The function gives the plot of the results with the signal

Value

a plot of the results with the signal

Examples

```
data(Data)
lyear=365.25
Kmax=4
lmin=1
result=GNSSseg(Data,lyear,selection.K="none",Kmax=Kmax)
plot_GNSS(Data,result$seg,result$funct)
```

Index