Package 'boxfilter'

May 24, 2023

Type Package

Title Filter Noisy Data

Version 0.2

Date 2023-05-24

Maintainer Thomas Ruf <Thomas.P.Ruf@me.com>

Description Noise filter based on determining the proportion of neighboring points. A false point will be rejected if it has only few neighbors, but accepted if the proportion of neighbors in a rectangular frame is high. The size of the rectangular frame as well as the cut-off value, i.e. of a minimum proportion of neighbor-points, may be supplied or can be calculated automatically. Originally designed for the cleaning of heart rates, but suitable for filtering any slowly-changing physiological variable.For more information see Signer (2010)<doi:10.1111/j.2041-210X.2009.00010.x>.

License GPL (>= 3)

Encoding UTF-8

LazyData true

Imports ggplot2, gridExtra

Depends R (>= 3.50)

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Author Thomas Ruf [aut, cre] (<https://orcid.org/0000-0002-9235-7079>)

Repository CRAN

Date/Publication 2023-05-24 11:20:11 UTC

R topics documented:

---

boxclip   *Boxfilter Main Function*

Description

Filters noise from data (e.g. heartrates) using x for x-axis data and y for y-axis, based on the proportion of neighbors of each point in a surrounding box of width and height (these may be determined automatically). It discards all data points that have less than a proportion of clipit neighbors.

Usage

```
boxclip (x=NULL,y,clipit=NULL,QI=NULL,width=NULL, height=NULL,
miny=min(y,na.rm=TRUE), maxy=max(y,na.rm=TRUE), plotit=TRUE, histo=FALSE)
```

Arguments

x	The x-axis of data, a datetime for example. Optional. If x=NULL x<-1:length(y) will be generated.
y	The y-axis of data, probably noisy. Required.
clipit	Optional. Y-values with less than a proportion of clipit neighbors will be discarded. If clipit is omitted it is set equal to the first trough in the neighbor proportion histogram.
ci	
QI	Optional. An integer quality index for each data point.
width	Optional. The width of the box. If width is omitted it will be generated from floor(length(x)*0.01).
height	Optional. The height of the box. If height is omitted it will be generated from floor(mean(y,na.rm=T)/4).
miny	The minimum y-value expected. Anything below miny is discarded.
maxy	The maximum y-value expected. Anything above maxy is discarded.
plotit	Optional.If TRUE show a graph of the original and filtered data.
histo	Optional. If TRUE also show a histogram of the neighboring points.

Details

Boxfilter mimics the human criterion of self-similarity. Data points with many neighbors are more trustworthy.

Value

x	Original x-axis data
y	Original y-axis data
filtered	Filtered data. Discarded data points are set to NA.
neighbors	Proportion of neighbors of each point.

Note

To store only filtered data, use e.g.:

bc=boxclip (x,y) data=data.frame(x=bc$x, hrf=bc$filtered) data=na.omit(data) write.csv(data,file="myheartrates.csv")

Author(s)

Thomas Ruf (thomas.p.ruf@me.com)

See Also

clipview

Examples

```
data("wb_month")
data("ibex_hr")

x=wb_month$x
y=wb_month$hr


myclip=boxclip(x,y, histo=TRUE)
summary(myclip)

r=seq(1,28400,by=4)
myclip=boxclip(ibex_hr$Time[r],ibex_hr$Heartrate[r],0.65)
summary(myclip)
#store(myclip)


data("sleepduration")
Date=as.POSIXct(sleepduration$Date)
Duration=as.numeric(sleepduration$Bedtime)
boxclip(Date,Duration,miny=0)
```

---

clipview   *Get View Of Different clipit Values*

Description

Gives a view of the original data, a histogram of the neighbors, and results of four values of clipit. The cutoff-values clipit are determined by a sequence of length .-

Usage

```
clipview (x=NULL,y,clipit=NULL,width=NULL, height=NULL,
miny=min(y,na.rm=TRUE), maxy=max(y,na.rm=TRUE),plotit=FALSE)
```

Arguments

x	The x-axis of data, a datetime for example. Optional. If x=NULL x<-1:length(y) will be generated.
y	The y-axis of data, probably noisy. Required.
clipit	This function requires a sequence of four items <1, e.g. clipit=seq(0.1, 0.4, by=0.1)
width	Optional. The width of the box. If width is omitted it will be generated from floor(length(x)*0.01).
height	Optional. The height of the box. If height is omitted it will be generated from floor(mean(y,na.rm=T)/4).
miny	The minimum y-value expected. Anything below miny is discarded.
maxy	The minimum y-value expected. Anything above maxy is discarded.
plotit	Optional.If TRUE show a graph of the original and filtered data.

Details

Note that a sequence is required for clipit here, while a scalar is required in boxclip().

Value

There is no return value.Six graphs are generated, original, histogram, and four for each value of clipit.

Author(s)

Thomas Ruf (thomas.p.ruf@me.com)

See Also

boxclip()

Examples

```
data("wb_year")


r=seq(1,54179,by=5)

x=wb_year$x[r]
y=wb_year$y[r]

clipview(x,y,clipit=seq(0.2,0.5,0.1), miny=10)
```

---

ibex_hr   *One year of heart rates of a capricorn free-living in the alps.*

Description

Heart rates were obtained using acoustic loggers in the rumen.

Usage

```
data("ibex_hr")
```

Format

A data frame with 28454 observations on the following 2 variables.

Time datetime

Heartrate in bpm

References

Signer, C., Ruf, T., & Arnold, W. (2011). Hypometabolism and basking: The strategies of Alpine ibex to endure harsh over-wintering conditions. Functional Ecology, 25(3), 537-547.

Examples

```
data(ibex_hr)
```

---

showdata   *Boxfilter function*

Description

Shows a graph of the data and its change over time.Called by boxclip()

Usage

```
showdata(x,y)
```

Arguments

x   The x-axis of data, a datetime for example.

y   The y-axis of data, probably noisy.

Details

Asks the user whether to continue or not.Continue only if you want points to be deleted

Author(s)

Thomas Ruf (thomas.p.ruf@me.com)

Examples

```
set.seed(1234)
y=runif(1000,20,30)
ix=sample(1:1000,50)
y[ix]=runif(50,60,70)
showdata(1:1000,y)
```

---

sleepduration   *data centering around sleeop in a humam*

Description

data on duration in bed

Usage

```
data("sleepduration")
```

Format

A data frame with 881 observations on the following 2 variables.

`Date` a character vector

`Bedtime` a character vector, convert to decimal number

Source

https://www.kaggle.com/datasets/danagerous/sleep-data

Examples

```
data(sleepduration)
## maybe str(sleepduration) ; plot(sleepduration) ...
```

---

store	*Store data*

Description

Stores original and filtered data.

Usage

```
store(object)
```

Arguments

object   Must be of class "boxclip", resulting from boxclip.

Value

returns nothing

Author(s)

Thomas Ruf (thomas.p.ruf@me.com)

Examples

```
data ("wb_month.RData")

x=wb_month$x
y=wb_month$hr

myclip=boxclip(x,y,clipit=0.25,width=15)
summary(myclip)
store(myclip)
```

---

summary.boxclip   *Summarize boxclip results*

Description

Summary method for class boxclip.

Usage

```
## S3 method for class 'boxclip'
summary(object,...)
```

Arguments

object	an object of class boxclip
...	currently, no other arguments are required.

Value

summary.boxclip prints the following items:

clipit	Cut-off value. Only points with a proportion >= clipit will be retained.
width	Rectangle width in x-units. Filter criterion is the proportion of data points inside the rectangle.
height	Rectangle height in y-units. Filter criterion is the proportion of data points inside the rectangle.
full	Number of original data.
remaining	Number of data remaining.
rest	Percentage remaining.

Author(s)

Thomas Ruf <thomas.p.ruf@me.com>

References

Signer (2010) <doi: 10.1111/j.2041-210X.2009.00010.x>

See Also

boxclip

Examples

```
data ("wb_month.RData")

x=wb_month$x
y=wb_month$hr

myclip=boxclip(x,y,clipit=0.25,width=15)
summary(myclip)
```

---

wb_month	*One month of heart rates and their quality in a wild boar. Quality was assessed by Star-Oddi, Island.*

Description

Heart rates were obtained from DST centi- HRT, Star-Oddi, Gardabaer, Iceland.

Usage

```
data("wb_month")
```

Format

A data frame with 3720 observations on the following 3 variables.

x x-axis, datetime

hr y-axis, heart rate

QI quality index (0-3) of the signal

Examples

```
data(wb_month)
```

---

wb_year	*One year of heart rates of a wild boar female.*

Description

Heart rates were obtained from DST centi- HRT, Star-Oddi, Gardabaer, Iceland.

Usage

```
data("wb_year")
```

Format

A data frame with 3720 observations on the following 3 variables.

x x-axis, datetime

y y-axis, heart rate

References

Ruf, T., Vetter, S. G., Painer, J., Stalder, G., & Bieber, C. (2021). Atypical for northern ungulates, energy metabolism is lowest during summer in female wild boars (Sus scrofa). Scientific Reports, 11(1), 1-12.

Examples

```
data(wb_year)
```

Index