Package ‘HS’

September 10, 2019

Type  Package
Title  Homogenous Segmentation for Spatial Lines Data
Date  2019-08-31
Version  1.1
Maintainer  Yongze Song <yongze.song@postgrad.curtin.edu.au>
Description  Methods of homogenous segmentation for spatial lines data, such as pavement performance indicators and traffic volumes. Three methods are available for homogenous segmentation, including cumulative difference approach, minimization coefficient of variation, and spatial heterogeneity based method.
Imports  zoo, ggpubr, data.table, tidyr, reshape2, ggplot2, partitions, utils
Depends  R (>= 3.4.0)
License  GPL-2
Encoding  UTF-8
LazyData  true
RoxygenNote  6.0.1
NeedsCompilation  no
Author  Yongze Song [aut, cre] (<https://orcid.org/0000-0003-3420-9622>)
Repository  CRAN
Date/Publication  2019-09-10 04:40:03 UTC

R topics documented:

cda ........................................................................................................ 2
deflection ............................................................................................ 3
findbreak ................................................................................................ 3
hs ............................................................................................................. 4
hsctg ...................................................................................................... 5
mcv ......................................................................................................... 5
preprocessing ......................................................................................... 6
segbycategory ....................................................................................... 7
Cummulative difference approach (CDA) for homogeneous segmentation of spatial lines data.

Description

Function for homogeneous segmentation of spatial lines data using a cumulative difference approach (CDA).

Usage

cda(var = "deflection", length = "length", data, range = NULL)

Arguments

- var: A character or a character vector of variable names, such as a road pavement performance indicator.
- length: A character of road length name in data.
- data: A data frame of a dataset.
- range: A vector of length threshold.

Examples

testdata <- tsdwa[1:100,]
testdata$Length <- testdata$SLK.end - testdata$SLK.start
testdata <- cda(var = "Deflection", length = "Length", testdata)
**deflection**

*Road deflection dataset.*

**Description**

The "deflection" dataset is a sample of the road deflection data monitored and collected by Main Roads Western Australia.

**Usage**

`deflection`

**Format**

`deflection`: A data frame with 1000 rows and 4 variables.

- `id`: Number of observation.
- `SLK.start`: Spatial start location of data. SLK is short for the straight line kilometer.
- `SLK.end`: Spatial end location of data.
- `Deflection`: The monitored road deflection value.

**Author(s)**

Yongze Song <yongze.song@postgrad.curtin.edu.au>

**findbreak**

*Find spatial breaking locations and add a column of breaks.*

**Description**

Find spatial breaking locations and add a column of breaks.

**Usage**

`findbreak(start = "SLK.start", end = "SLK.end", data, dist.allow = 0.05, line.no = NULL)`

**Arguments**

- `start`: A character of start location name of a spatial line.
- `end`: A character of end location name of a spatial line.
- `data`: A data frame of a dataset.
- `dist.allow`: A number of the maximum allowed breaks within a line segment.
- `line.no`: A character of spatial line name.
hs

Homogeneous segmentation function with continuous variables.

Description

Homogeneous segmentation function with continuous variables.

Usage

hs(start = "SLK.start", end = "SLK.end", var = "deflection", data, method = "shs", range = NULL)

Arguments

- **start**: A character of start location name of a spatial line.
- **end**: A character of end location name of a spatial line.
- **var**: A character or a character vector of variable names, such as a road pavement performance indicator.
- **data**: A data frame of a dataset.
- **method**: A character of homogeneous segmentation method. Available methods include "shs", "cda" and "mcv".
- **range**: A vector of segment length threshold.

Examples

testdata <- tsdwa[1:100,]
testdata <- findbreak(start = "SLK.start", end = "SLK.end",
                      data = testdata, dist.allow = 0.05)

hs1 <- hs(start = "SLK.start", end = "SLK.end", var = c("Curvature", "Deflection", "BLI"),
          testdata, method = "shs", range = c(0.1, 0.5))
hsctg

Homogeneous segmentation function with both categorical and continuous variables.

Description
Homogeneous segmentation function with both categorical and continuous variables.

Usage
hsctg(start = "SLK.start", end = "SLK.end", var = "deflection", data, method = "shs", range = NULL, by.ctg = NULL)

Arguments
- start: A character of start location name of a spatial line.
- end: A character of end location name of a spatial line.
- var: A character or a character vector of variable names, such as a road pavement performance indicator.
- data: A data frame of a dataset.
- method: A character of homogeneous segmentation method. Available methods include "shs", "cda" and "mcv".
- range: A vector of segment length threshold.
- by.ctg: A vector of categorical variable names.

Examples
testdata <- tsdwa[1:300,]
hc1 <- hsctg(start = "SLK.start", end = "SLK.end", var = c("Curvature", "Deflection", "BLI"),
data, method = "shs", range = c(0.1, 0.5), by.ctg = c("SurfType", "PvtType"))

mcv

Minimization coefficient of variation (MCV) for homogeneous segmentation of spatial lines data.

Description
Minimization coefficient of variation (MCV) for homogeneous segmentation of spatial lines data.

Usage
mcv(var = "deflection", length = "length", data, range = NULL)
preprocessing

Arguments

<table>
<thead>
<tr>
<th>var</th>
<th>A character or a character vector of variable names, such as a road pavement performance indicator.</th>
</tr>
</thead>
<tbody>
<tr>
<td>length</td>
<td>A character of road length name in data.</td>
</tr>
<tr>
<td>data</td>
<td>A data frame of a dataset.</td>
</tr>
<tr>
<td>range</td>
<td>A vector of segment length threshold.</td>
</tr>
</tbody>
</table>

Examples

testdata <- tsdwa[1:100,]
testdata$length <- testdata$SLK.end - testdata$SLK.start
testdata <- mcv(var = "Deflection", length = "length", testdata, range = c(0.1, 0.5))

preprocessing Preprocessing for field monitoring data.

Description

The preprocessing includes two steps: removing missing data and ordering data by spatial locations.

Usage

preprocessing(var = "deflection", location = "SLK", data = data)

Arguments

<table>
<thead>
<tr>
<th>var</th>
<th>A character of the name of a variable in a dataset, such as a road pavement performance indicator.</th>
</tr>
</thead>
<tbody>
<tr>
<td>location</td>
<td>A character of the name of spatial locations in a dataset.</td>
</tr>
<tr>
<td>data</td>
<td>A data frame of monitoring data.</td>
</tr>
</tbody>
</table>

Examples

testdata <- tsdwa[1:100,]
testdata <- preprocessing(var = "Deflection", location = "SLK.start", data = testdata)
segbycategory

Segmentation with categorical variables.

Description

Segmentation with categorical variables.

Usage

segbycategory(data, by = NULL)

Arguments

data A data frame of a dataset.

by A character or a vector of categorical variable names.

Examples

testdata <- tsdw[1:100,]
testdata <- segbycategory(testdata, by = c("SurfType", "PvtType"))

segcompare

Segments comparison of different homogeneous segmentations methods.

Description

Segments comparison of different homogeneous segmentations methods.

Usage

segcompare(start = "SLK.start", end = "SLK.end", var = "deflection",
           data, segid.matrix, methods = NULL)

## S3 method for class 'segcompare'
print(x, ...)
## S3 method for class 'segcompare'
plot(x, ...)
Arguments

- **start**: A character of start location name of a spatial line.
- **end**: A character of end location name of a spatial line.
- **var**: A character or a character vector of variable names, such as a road pavement performance indicator.
- **data**: A list of segmentation result.
- **segid.matrix**: A matrix of segmentations.
- **methods**: A vector of segmentation method names, default NULL.
- **x**: A list of segments comparison result.
- **...**: Ignore

Examples

testdata <- tsdwa[1:300, ]
testdata$length <- testdata$SLK.end - testdata$SLK.start
variable <- c("Curvature", "Deflection", "BLI")

seg1 <- hs(start = "SLK.start", end = "SLK.end", var = variable, testdata, method = "shs", range = c(0.1, 0.5)) # 0.3 s
seg2 <- hs(start = "SLK.start", end = "SLK.end", var = variable, testdata, method = "cda", range = c(0.1, 0.5)) # 0.7 s
seg3 <- hs(start = "SLK.start", end = "SLK.end", var = variable, testdata, method = "mcv", range = c(0.1, 0.5)) # 0.6 s
segid.matrix <- cbind(seg1$seg.id, seg2$seg.id, seg3$seg.id)
data(segid.matrix)

cp <- segcompare(start = "SLK.start", end = "SLK.end", var = variable, testdata, segid.matrix, methods = c("SHS", "CDA", "MCV")) # 4.8 s
cp
plot(cp)

Description

Segmentation results of CDA, MCV and SHS methods for data "tsdwa".

Usage

segid.matrix

Format

segid.matrix: A matrix with 300 rows and 3 columns, representing segmentations of three methods, CDA, MCV and SHS.
Author(s)

Yongze Song <yongze.song@postgrad.curtin.edu.au>

segplot

Visualization of homogeneous segments.

Description

Visualization of homogeneous segments.

Usage

segplot(start = "SLK.start", var = "deflection",
        seg.id = "seg.id", data, plot.range = NULL)

Arguments

- **start**: A character of start location name of a spatial line.
- **var**: A character or a character vector of variable names, such as a road pavement performance indicator.
- **seg.id**: A character of the name of new segment number.
- **data**: A data frame of a dataset.
- **plot.range**: A vector of plot range.

Examples

testdata <- tsdwa[1:300,]
testdata$length <- testdata$SLK.end - testdata$SLK.start
testdata <- shs(var = c("Curvature", "Deflection"), length = "length",
                testdata, range = c(0.1, 0.5))
segplot(start = "SLK.start", var = c("Curvature", "Deflection"),
        seg.id = "seg.id", testdata, plot.range = 1:300)

segsmooth

Smoothing data using the moving average method for the homogeneous segmentation.

Description

A center aligned moving window is used for the moving average method.

Usage

segsmooth(var = "deflection", range = 11, data)
segsummary

Statistical summary of homogeneous segments.

Description
Statistical summary of homogeneous segments.

Usage
segsummary(start = "SLK.start", end = "SLK.end", var = "deflection", 
seg.id = "seg.id", data, by.ctg = NULL)
## S3 method for class 'segsummary'
print(x, ...)

Arguments

start A character of start location name of a spatial line.
end A character of end location name of a spatial line.
var A character or a character vector of variable names, such as a road pavement 
performance indicator.
seg.id A character of the name of new segment number.
data A data frame of a dataset.
by.ctg A character of categorical variable names.
x A list of segmentation result.
... Ignore

Vari
A character of the name of a variable in a dataset, such as a road pavement 
performance indicator.

range A number of the size of moving window. An odd number is required.
data A data frame of monitoring data.

Examples

# preprocessing
testdata <- tsdwa[1:500,]
testdata <- preprocessing(var = "Deflection", location = "SLK.start", data = testdata)
# smoothing
testdata <- segsmooth(var = "Deflection", range = 11, data = testdata)
# plot
plot(testdata$SLK.start, testdata$Deflection, type = "l", 
col = "lightblue", xlab = "location", ylab = "deflection")
lines(testdata$SLK.start, testdata$smooth.Deflection)
Examples

```r
testdata <- tsdw[1:100,]
testdata$length <- testdata$SLK.end - testdata$SLK.start
testdata <- shs(var = c("Curvature", "Deflection"), length = "length",
               testdata, range = c(0.1, 0.5))
s1 <- segsummary(start = "SLK.start", end = "SLK.end", var = c("Curvature", "Deflection"),
                seg.id = "seg.id", testdata)
s1
```

shs

Spatial heterogeneity-based segmentation (SHS) for homogeneous segmentation of spatial lines data.

Description

Spatial heterogeneity-based segmentation (SHS) for homogeneous segmentation of spatial lines data.

Usage

```r
shs(var = "deflection", length = "length", data, range = NULL)
```

Arguments

- `var`: A character or a character vector of variable names, such as a road pavement performance indicator.
- `length`: A character of road length name in data.
- `data`: A data frame of a dataset.
- `range`: A vector of segment length threshold.

Examples

```r
testdata <- tsdw[1:100,]
testdata$length <- testdata$SLK.end - testdata$SLK.start
testdata <- shs(var = "Deflection", length = "length", testdata, range = c(0.1, 0.5))
```
### splitlong

*Split long segments to segments within length threshold.*

**Description**

Split long segments to segments within length threshold.

**Usage**

```r
splitlong(var = "deflection", length = "length", 
           seg.id = "seg.id", data, range = NULL)
```

**Arguments**

- **var**: A character or a character vector of variable names, such as a road pavement performance indicator.
- **length**: A character of length name.
- **seg.id**: A character of the name of new segment number.
- **data**: A data frame of a dataset.
- **range**: A vector of segment length threshold.

**Examples**

```r
testdata <- tsdwa[1:1000,]
testdata$length <- round(testdata$SLK.end - testdata$SLK.start, digits = 10)
testdata <- cda(var = "Deflection", length = "length", testdata, range = c(0.1, 0.5))
testdata <- splitlong(var = "Deflection", length = "length", 
                      seg.id = "seg.id", testdata, range = c(0.1, 0.5))
seglength.summary <- testdata[, .(sum(length)), by = .(seg.id)]
```

---

**tsdwa**

*Traffic speed deflectometer (TSD) data of pavement deteriorations.*

**Description**

The "tsdwa" dataset is a sample of the pavement deterioration data monitored and collected by Main Roads Western Australia.

**Usage**

```r
tsdwa
```
Format

tsdwa: A data frame with 5000 rows and 8 variables.

- id. Number of observation.
- SLK.start. Spatial start location of data. SLK is short for the straight line kilometer.
- SLK.end. Spatial end location of data.
- SurfType. Surfacing type.
- PvtType. Pavement type.
- Curvature.
- Deflection.
- BLI. Base layer index.

Author(s)

Yongze Song <yongze.song@postgrad.curtin.edu.au>
Index

*Topic **TSD**
  tsdwa, 12

*Topic **dataset**
  deflection, 3
  segid.matrix, 8
  tsdwa, 12

*Topic **deflection**
  deflection, 3

*Topic **segmentations**
  segid.matrix, 8

cda, 2

deflection, 3

findbreak, 3

hs, 4
hsctg, 5

mcv, 5

plot.segcompare(segcompare), 7
preprocessing, 6
print.segcompare(segcompare), 7
print.segsummary(segsummary), 10

segbycategory, 7
segcompare, 7
segid.matrix, 8
segplot, 9
segsMOOTH, 9
segsummary, 10
shs, 11
splitlong, 12

tsdwa, 12