

Package ‘HEDA’

July 20, 2021

Title 'Hydropeaking Events Detection Algorithm'

Type Package

Version 0.1.5

Description

This tool identifies hydropeaking events from raw time-series flow record, a rapid flow variation induced by the hourly-adjusted electricity market. The novelty of 'HEDA' is to use vector angle instead of the first-order derivative to detect change points which not only largely improves the computing efficiency but also accounts for the rate of change of the flow variation. More details <[doi:10.1016/j.jhydrol.2021.126392](https://doi.org/10.1016/j.jhydrol.2021.126392)>.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Depends R (>= 4.0), dplyr (>= 1.0.2), ggplot2 (>= 3.1.0), zoo (>= 1.8.7), lubridate (>= 1.7.4), rlang (>= 0.4.11)

RoxygenNote 7.1.1

NeedsCompilation no

Author Tingyu Li [aut, cre],
Xiaotian Zou [aut],
Gregory Pasternack [aut]

Maintainer Tingyu Li <styli@ucdavis.edu>

Repository CRAN

Date/Publication 2021-07-20 19:00:02 UTC

R topics documented:

Clean_conectD	2
Clean_position	3
Clean_Spt	3
HEDA_Tidy	4
HPK_frq_mgt	5
HPK_plot	6

HPK_rt_dur	7
HPK_SampleData	8
ReversalCount	8

Index	10
--------------	-----------

Clean_conectD	<i>Clean_conectD</i>
---------------	----------------------

Description

Evaluate difference between peaking and off-peaking discharge

Usage

```
Clean_conectD(df, alpha3 = 0.7, alpha4 = 0.5)
```

Arguments

df	Dataframe to be processed
alpha3	Default value: 0.7
alpha4	Default value: 0.5

Details

Evaluate whether the difference in discharge between peaking and off-peaking points is qualified to be identified as hydropeaking events.

Value

The output will be a dataframe in the same form with the input.

Examples

```
## Not run:
# before running the function
HPK_SampleData$dateTime <- parse_date_time(HPK_SampleData$dateTime, "mdy HM")

hpk_flow_cln <- HEDA_Tidy(HPK_SampleData, season = c(6,7,8,9))

hpk_flow_cg <- ReversalCount(hpk_flow_cln)
hpk_flow_cg <- Clean_position(hpk_flow_cg)
hpk_flow_cg <- Clean_Spt(hpk_flow_cg)

# clean the unqualified peaking and off-peaking discharge
hpk_flow_cg <- Clean_conectD(hpk_flow_cg)

# or change the default values
hpk_flow_cg <- Clean_conectD(hpk_flow_cg, alpha3 = 0.7, alpha4 = 0.5)

## End(Not run)
```

Clean_position	<i>Clean_position</i>
----------------	-----------------------

Description

Change points are excluded if they are in the wrong position. For example, both point 3 and the peak pair represent the peaking discharge whose value (position) should be close to the daily maximum discharge. If the peaking discharge is close to the daily minimum discharge, change points will be removed since they are in the wrong positions

Usage

```
Clean_position(df, alpha2 = 0.3)
```

Arguments

df	dataframe to be processed
alpha2	Default value : 0.3

Value

output will be a dataframe in the same form as the input

Examples

```
# before running the function
HPK_SampleData$dateTime <- parse_date_time(HPK_SampleData$dateTime, "mdy HM")

hpk_flow_cln <- HEDA_Tidy(HPK_SampleData, season = c(6,7,8,9))

hpk_flow_cg <- ReversalCount(hpk_flow_cln)

# running the function to clean unqualified change points
hpk_flow_cg <- Clean_position(hpk_flow_cg)

# default value is open to change
hpk_flow_cg <- Clean_position(hpk_flow_cg, alpha2 = 0.3)
```

Clean_Spt	<i>Clean_Spt</i>
-----------	------------------

Description

Clean repeated points

Usage

```
Clean_Spt(df, alpha3 = 0.7, alpha4 = 0.5)
```

Arguments

df	Dataframe to be processed
alpha3	Default value: 0.7
alpha4	Default value: 0.5

Details

Evaluate whether the difference in discharge between peaking and off-peaking points is qualified to be identified as hydropeaking events

Value

Output file will be a dataframe in the same form as the input

Examples

```
# before running the function
HPK_SampleData$dateTime <- parse_date_time(HPK_SampleData$dateTime,"mdy HM")

hpk_flow_cln <- HEDA_Tidy(HPK_SampleData, season = c(6,7,8,9))

hpk_flow_cg <- ReversalCount(hpk_flow_cln)
hpk_flow_cg <- Clean_position(hpk_flow_cg)

# clean continuous points of the same type
hpk_flow_cg <- Clean_Spt(hpk_flow_cg)

# or change the default values
hpk_flow_cg <- Clean_Spt(hpk_flow_cg, alpha3 = 0.7, alpha4 = 0.5)
```

HEDA_Tidy

HEDA_Tidy

Description

Format flow record into hourly record; Split record by designated season; Interpolate and smooth the record.

Usage

```
HEDA_Tidy(df, season)
```

Arguments

df Dataframe to be processed
 season Subset data by season

Value

Output dataframe has 4 columns containing location id, datetime, parameter value, ann_thre. ann_thre is originally the mean annual discharge. If flow record is subset by season, ann_thre will be the mean discharge during the subset period

Examples

```
## Not run:
# format the time column to time column
HPK_Sample_data$dateTime <- parse_date_time(HPK_Sample_data$dateTime,"mdy HM")

# clean subset data by season
hpk_flow_cln <- HEDA_Tidy(HPK_Sample_data, season = c(6,7,8,9))

## End(Not run)
```

HPK_frq_mgt	<i>HPK_frq_mgt</i>
-------------	--------------------

Description

Extract frequency and magnitude metrics

Usage

```
HPK_frq_mgt(df)
```

Arguments

df Dataframe to be processed

Details

Extract Qpeak, offQpeak, pk_no and pkraio from the identified hydropeaking events

Value

Output dataframe includes time-series values of four metrics. Qpeak is the hydropeaking discharge, offQpeak is base flow, pk_no is the daily number of rise and fall process, pkratio is the ratio of days with hydropeaking to the total number of days analyzed

Examples

```
## Not run:  
HpkFrqMgt <- HPK_frq_mgt(hpk_flow_cg)  
  
## End(Not run)
```

HPK_plot

HPK_plot

Description

Plot hydrograph

Usage

```
HPK_plot(df)
```

Arguments

df Dataframe to be processed

Details

Plot the hydrograph of the processed data with change points marked by different colors

Value

The plot will be presented under the “Plots” tab in RStudio. Users can determine whether to save out the diagram by themselves. The plot shows flow in m^3/s

Examples

```
## Not run:  
## subset the data  
tt = hpk_flow_cg[13100:13400,]  
  
## plot the hydrograph of the subset data  
HPK_plot(tt)  
  
## End(Not run)
```

HPK_rt_dur

HPK_rt_dur

Description

Extract rate of change and duration related metrics

Usage

```
HPK_rt_dur(df)
```

Arguments

df Dataframe to be processed

Details

Extract rate of change and duration related metrics: pk_rtn is the retention of peaking process. offpk_rtn is the retention of base flow. D_rampup/D_ramp is the duration of rise/fall process. RB_Indx_up/RB_Index_dw is the flashness of rise/fall process. Ramp_up/Ramp_dw is the rate of change of rise/fall process. Strange_up/Strange_dw is the standardized rise/fall amplitude

Value

Output dataframe includes time-series values of four metrics. Qpeak is the hydropeaking discharge, offQpeak is base flow, pk_no is the daily number of rise and fall process, pkratio is the ratio of days with hydropeaking to the total number of days analyzed

Examples

```
## Not run:
HpkRtDur <- HPK_rt_dur(hpk_flow_cg)

## to extract the time-series of one metric without missing value
pk_rtn <- HpkRtDur[na.omit(HpkRtDur$pk_rtn),c("location_id", "datetime", "pk_rtn")]

## End(Not run)
```

HPK_SampleData	<i>HPK_SampleData</i>
----------------	-----------------------

Description

Flow record of Cherry creek below Dion R Holm Powerhouse

Usage

```
data(HPK_SampleData)
```

Format

site_no identification number of gauge station
 dateTime Date and time flow when discharge recorded
 dateTime Date and time discharge recorded

Details

A data frame with 3 columns and 30000 flow records

Source

https://waterdata.usgs.gov/nwis/uv?site_no=11278400

Examples

```
data(HPK_SampleData)
```

ReversalCount	<i>ReversalCount</i>
---------------	----------------------

Description

Detect hydropeaking events

Usage

```
ReversalCount(df, alpha1 = 0.03, theta = 60, gamma = 1.1)
```

Arguments

df	Dataframe to be processed
alpha1	Default value : 0.03
theta	Default value : 60 degree for m ³ /s, 85 degree for cfs.
gamma	gamma default value: 1.1 m/s or 40 cfs.

Value

Output dataframe has six columns containing location id, datatime, parameter value, ann_thre, vt_degree and dtag. ann_thre is originally the mean annual discharge. If flow record is subset by season, ann_thre will be the mean discharge of that period. vt_degree is the vector angle between two flow vectors. dtag is the catogrized change points.

References

Li, T. and Pasternack, G.B., 2021. Revealing the diversity of hydropeaking flow regimes. *Journal of Hydrology*, 598, p.126392.

Examples

```
## Not run:  
hpk_flow_cg <- ReversalCount(hpk_flow_cln)  
  
hpk_flow_cg <- ReversalCount(hpk_flow_cln, alpha1 = 0.03, theta = 85, gamma = 40)  
  
## End(Not run)
```

Index

Clean_conectD, [2](#)
Clean_position, [3](#)
Clean_Spt, [3](#)

HEDA_Tidy, [4](#)
HPK_frq_mgt, [5](#)
HPK_plot, [6](#)
HPK_rt_dur, [7](#)
HPK_SampleData, [8](#)

ReversalCount, [8](#)