Package 'EGRETci'

March 15, 2019

Type Package

Title Exploration and Graphics for RivEr Trends Confidence Intervals

Version 2.0.3

Description Collection of functions to evaluate uncertainty of results from water quality analysis using the Weighted Regressions on Time Discharge and Season (WRTDS) method. This package is an add-on to the EGRET package that performs the WRTDS analysis. The WRTDS modeling method was initially introduced and discussed in Hirsch et al. (2010) <doi:10.1111/j.1752-1688.2010.00482.x>, and expanded in Hirsch and De Cicco (2015) <doi:10.3133/tm4A10>. The paper describing the uncertainty and confidence interval calculations is Hirsch et al. (2015) <doi:10.1016/j.envsoft.2015.07.017>.

License CC0

Depends R (>= 3.0)

Imports EGRET(>= 3.0.0), binom, stats, graphics, utils

Suggests knitr, testthat, doParallel, iterators, rmarkdown

LazyLoad yes

LazyData yes

VignetteBuilder knitr

BuildVignettes true

URL https://github.com/USGS-R/EGRETci

BugReports https://github.com/USGS-R/EGRETci/issues

Copyright This software is in the public domain because it contains materials that originally came from the United States Geological Survey, an agency of the United States Department of Interior. For more information, see the official USGS copyright policy at http://www.usgs.gov/visual-id/credit_usgs.html#copyright

RoxygenNote 6.1.1

NeedsCompilation no

Author Laura DeCicco [aut, cre] (<https://orcid.org/0000-0002-3915-9487>),
Robert Hirsch [aut] (<https://orcid.org/0000-0002-4534-075X>),
Jennifer Murphy [ctb]

Maintainer Laura DeCicco <ldecicco@usgs.gov>

Repository CRAN

Date/Publication 2019-03-15 05:17:35 UTC

R topics documented:

`EGRETci-package` *EGRETci package for bootstrap hypothesis tests and confidence interval analysis for WRTDS (Weighted Regressions on Time, Discharge, and Season) statistical models. This package is designed to be used in conjunction with the EGRET package, which estimates and describes WRTDS models.*

Description

Package:	EGRETci
Type:	Package
License:	Unlimited for this package, dependencies have more restrictive licensing.
Copyright:	This software is in the public domain because it contains materials that originally came from the United States G
LazyLoad:	yes

Collection of functions to evaluate uncertainty of results from water quality analysis using the Weighted Regressions on Time Discharge and Season (WRTDS) method. This package is an add-on to the EGRET package that performs the WRTDS analysis.

Author(s)

Robert M. Hirsch <rhirsch@usgs.gov>, Laura De Cicco <ldecicco@usgs.gov>

References

Hirsch, R.M., and De Cicco, L.A., 2015, User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data: U.S. Geological Survey Techniques and Methods book 4, chap. A10, 94 p., https://doi.org/10.3133/tm4A10

Hirsch, R.M., Archfield, S.A., and De Cicco, L.A., 2015, A bootstrap method for estimating uncertainty of water quality trends. Environmental Modelling & Software, 73, 148-166. https://www.sciencedirect.com/science/article/pii/S1364815215300220

`blockSample` *blockSample*

Description

Get a random subset of the Sample data frame based on the user-specified blockLength for use in bootstrap estimation process. The subset is a random subset of blocks of data from Sample dataframe. The subset is based on the random selection (with replacement) of starting dates from the full Sample data frame. The Sample selected has the same number of observations as the original Sample (some observations are are included once, some included multiple times, and some are not included).

Usage

```
blockSample(localSample, blockLength, startSeed = NA)
```

Arguments

`localSample`	Sample data frame
`blockLength`	integer size of subset expressed in days.
`startSeed`	setSeed value. Defaults to 494817. This is used to make repeatable output.

Value

newSample data frame in same format as Sample data frame

Examples

```
library(EGRET)
eList <- Choptank_eList
Sample <- eList$Sample
bsReturn <- blockSample(Sample, 200)
```

bootAnnual *Single confidence interval bootstrap run*

Description

One bootstrap run used to calculate confidence interval bands.

Usage

```
bootAnnual(eList, blockLength = 200, startSeed = 494817,
  verbose = FALSE)
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes. Created from the EGRET package, after running modelEstimation.
blockLength	integer suggested value is 200
startSeed	setSeed value. Defaults to 494817. This is used to make repeatable output.
verbose	logical specifying whether or not to display progress message

Examples

```
library(EGRET)
eList <- Choptank_eList
## Not run:
annualResults <- bootAnnual(eList)

## End(Not run)
```

Choptank_eBoot *Example eBoot*

Description

Example data representing data from the Choptank River at Greensboro, MD, USGS data Data is a named list of the Daily, Sample, INFO dataframes, and xConc, and xFlux vectors.

ciBands *ciBands*

Description

Computes confidence intervals for Flow-Normalized Concentration and Flow-Normalized Flux for a WRTDS model.

Usage

```
ciBands(eList, repAnnualResults, probs = c(0.05, 0.95))
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes. Created from the EGRET package, after running modelEstimation.
repAnnualResults	named list returned from bootstrapping process
probs	vector high and low confidence interval percentages

Examples

```
library(EGRET)
eList <- Choptank_eList
nBoot <- 100
blockLength <- 200
## Not run:

repAnnualResults <- vector(mode = "list", length = nBoot)
for(n in 1:nBoot){
   annualResults <- bootAnnual(eList, blockLength, startSeed = n)
   repAnnualResults[[n]] <- annualResults
}

CIAnnualResults <- ciBands(eList, repAnnualResults)


## End(Not run)
```

ciCalculations *ciCalculations*

Description

Interactive function to calculate WRTDS confidence bands

Usage

```
ciCalculations(eList, startSeed = 494817, verbose = TRUE, ...)
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes. Created from the EGRET package, after running modelEstimation.
startSeed	setSeed value. Defaults to 494817. This is used to make repeatable output.
verbose	logical specifying whether or not to display progress message
...	optionally include nBoot, blockLength, or widthCI

Examples

```
library(EGRET)
eList <- Choptank_eList
## Not run:
CIAnnualResults <- ciCalculations(eList)

seriesOut_2 <- runSeries(eList, windowSide = 7)
CIAnnualResults <- ciCalculations(seriesOut_2,
                     nBoot = 10,
                     blockLength = 200,
                     widthCI = 90)

 plotConcHistBoot(seriesOut_2, CIAnnualResults)


## End(Not run)
```

---

plotConcHistBoot *Graph of annual concentration, flow normalized concentration, and confidence bands for flow normalized concentrations*

Description

Uses the output of modelEstimation in the EGRET package (results in the named list eList), and the data frame CIAnnualResults (produced by the function ciCalculations in the EGRETci package using scripts described in the EGRETci vignette) to produce a graph of annual concentration, flow normalized concentration, and confidence bands for flow-normalized concentrations. In addition to the arguments listed below, it will accept any additional arguments that are listed for the EGRET function plotConcHist.

Usage

```
plotConcHistBoot(eList, CIAnnualResults, yearStart = NA, yearEnd = NA,
  plotFlowNorm = TRUE, col.pred = "green", concMax = NA,
  printTitle = TRUE, cex.main = 1.1, ...)
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes. Created from the EGRET package, after running modelEstimation.
CIAnnualResults	data frame generated from ciBands (includes nBoot, probs, and blockLength attributes)
yearStart	numeric is the calendar year containing the first estimated annual value to be plotted, default is NA (which allows it to be set automatically by the data)
yearEnd	numeric is the calendar year just after the last estimated annual value to be plotted, default is NA (which allows it to be set automatically by the data)
plotFlowNorm	logical variable if TRUE flow normalized line is plotted, if FALSE not plotted
col.pred	character prediction color
concMax	number specifying the maximum value to be used on the vertical axis, default is NA (which allows it to be set automatically by the data)
printTitle	logical, default = TRUE.
cex.main	numeric title scale, default = 1.1.
...	graphical parameters

Examples

```
library(EGRET)
eList <- Choptank_eList
CIAnnualResults <- Choptank_CIAnnualResults
plotConcHistBoot(eList, CIAnnualResults)
plotConcHistBoot(eList, CIAnnualResults, yearStart=1990, yearEnd=2002)
## Not run:
CIAnnualResults <- ciCalculations(eList, nBoot = 100, blockLength = 200)
plotConcHistBoot(eList, CIAnnualResults)

## End(Not run)
```

---

plotFluxHistBoot *Graph of annual flux, flow normalized flux, and confidence bands for flow normalized flux*

Description

Uses the output of modelEstimation in the EGRET package (results in the named list eList), and the data frame CIAnnualResults (produced by EGRETci package using scripts described in the vignette) to produce a graph of annual flux, flow normalized flux, and confidence bands for flow-normalized flux. In addition to the arguments listed below, it will accept any additional arguments that are listed for the EGRET function plotFluxHist.

Usage

```
plotFluxHistBoot(eList, CIAnnualResults, yearStart = NA, yearEnd = NA,
  plotFlowNorm = TRUE, fluxUnit = 9, fluxMax = NA,
  col.pred = "green", printTitle = TRUE, cex.main = 1.1, ...)
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes. Created from the EGRET package, after running modelEstimation.
CIAnnualResults	data frame from ciBands (needs nBoot, probs, and blockLength attributes)
yearStart	numeric is the calendar year containing the first estimated annual value to be plotted, default is NA (which allows it to be set automatically by the data)
yearEnd	numeric is the calendar year just after the last estimated annual value to be plotted, default is NA (which allows it to be set automatically by the data)
plotFlowNorm	logical variable if TRUE flow normalized line is plotted, if FALSE not plotted
fluxUnit	number representing entry in pre-defined fluxUnit class array. printFluxUnitCheatSheet
fluxMax	number specifying the maximum value to be used on the vertical axis, default is NA (which allows it to be set automatically by the data)
col.pred	character prediction color
printTitle	logical
cex.main	numeric title scale
...	graphical parameters

Examples

```
library(EGRET)
eList <- Choptank_eList
CIAnnualResults <- Choptank_CIAnnualResults
plotFluxHistBoot(eList, CIAnnualResults, fluxUnit=5)

## Not run:
CIAnnualResults <- ciCalculations(eList, nBoot = 100, blockLength = 200)
plotFluxHistBoot(eList, CIAnnualResults, fluxUnit=5)

## End(Not run)
```

---

plotHistogramTrend *plotHistogramTrend*

Description

Histogram of trend results from bootstrap process. The histogram shows the trend results expressed as percentage change between the first year (or first period) and the second year (or second period). It shows the zero line (no trend) and also shows the WRTDS estimate of the trend in percent.

Usage

```
plotHistogramTrend(eList, eBoot, caseSetUp, flux = TRUE, xMin = NA,
  xMax = NA, xStep = NA, printTitle = TRUE, cex.main = 1.1,
  cex.axis = 1.1, cex.lab = 1.1, col.fill = "grey", ...)
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes. Created from the EGRET package, after running modelEstimation.
eBoot	named list. Returned from wBT.
caseSetUp	data frame. Returned from trendSetUp.
flux	logical if TRUE, plots flux results, if FALSE plots concentration
xMin	minimum bin value, it is good to have the xMin and xMax arguments straddle zero.
xMax	maximum bin value
xStep	step size, should probably be multiples of 10 or 20
printTitle	logical if TRUE, includes title
cex.main	numeric title font size
cex.axis	numeric axis font size
cex.lab	numeric label font size
col.fill	character fill color
...	base R graphical parameters that can be passed to the hist function

Examples

```
library(EGRET)
eList <- Choptank_eList
eBoot <- Choptank_eBoot
caseSetUp <- Choptank_caseSetUp
plotHistogramTrend(eList, eBoot, caseSetUp, flux=FALSE)

## Not run:
caseSetUp <- trendSetUp(eList)
eBoot <- wBT(eList,caseSetUp)
plotHistogramTrend(eList, eBoot, caseSetUp,
                   flux=FALSE, xMin = -20, xMax = 60, xStep = 5)
plotHistogramTrend(eList, eBoot, caseSetUp,
                   flux=TRUE, xMin = -20, xMax = 60, xStep = 5)

# Using runPairs:
year1 <- 1985
year2 <- 2009
pairOut_2 <- runPairs(eList, year1, year2, windowSide = 7)
boot_pair_out <- runPairsBoot(eList, pairOut_2, nBoot = 10)

plotHistogramTrend(eList, boot_pair_out,caseSetUp=NA,
```

```
                    flux=TRUE, xMin = -20, xMax = 60, xStep = 5)

## End(Not run)
```

---

pVal *pVal*

Description

Computes the two-sided p value for the null hypothesis, where null hypothesis is that the slope is zero, based on binomial distribution. Should be noted that the result does not depend on the magnitude of the s values only depends on the number of plus values and number of negative values.

Usage

```
pVal(s)
```

Arguments

s	numeric vector of slope values from the bootstrap (already flipped)

Value

pVal numeric value

Examples

```
s <- c(-1.0, 0, 0.5, 0.55, 3.0)
pValue <- pVal(s)
```

---

runGroupsBoot *The bootstrap uncertainty analysis for runGroups results*

Description

This function that does the uncertainty analysis for determining the change between two groups of years. The process is virtually identical to what is used for runPairsBoot.

Usage

```
runGroupsBoot(eList, groupResults, nBoot = 100, startSeed = 494817,
  blockLength = 200)
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes
groupResults	data frame returned from runGroups
nBoot	the maximum number of bootstrap replicates to be used, typically 100
startSeed	setSeed value. Defaults to 494817. This is used to make repeatable output.
blockLength	days, typically 200 is a good choice

Value

eBoot, a named list with bootOut,wordsOut,xConc,xFlux values. bootOut is a data frame with the results of the bootstrapping tests. wordsOut is a character vector describing the results. xConc, xFlux are vectors of length iBoot, of the change in flow normalized concentration or flux computed by each bootstrap replicate (mg/L). pConc and pFlux are vectors of length iBoot, of the change in flow normalized concentration or flux computed from each bootstrap replicate expressed as

See Also

runPairsBoot, runGroups

Examples

```
library(EGRET)
eList <- Choptank_eList

## Not run:
groupResults <- runGroups(eList,
                          group1firstYear = 1995,
                          group1lastYear = 2004,
                          group2firstYear = 2005,
                          group2lastYear = 2014,
                          windowSide = 7, wall = TRUE,
                          sample1EndDate = "2004-10-30",
                          paStart = 4, paLong = 2,
                          verbose = FALSE)

boot_group_out <- runGroupsBoot(eList, groupResults)

plotHistogramTrend(eList, boot_group_out, caseSetUp=NA)

## End(Not run)
```

---

runPairsBoot *The bootstrap uncertainty analysis for runPairs results*

Description

The function that does the uncertainty analysis for determining the change between any pair of years. It is very similar to the wBT function that runs the WRTDS bootstrap test. It differs from wBT in that it runs a specific number of bootstrap replicates, unlike the wBT approach that will stop running replicates based on the status of the test statistics along the way.

Usage

```
runPairsBoot(eList, pairResults, nBoot = 100, startSeed = 494817,
  blockLength = 200)
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes
pairResults	data frame returned from runPairs
nBoot	the maximum number of bootstrap replicates to be used, typically 100
startSeed	setSeed value. Defaults to 494817. This is used to make repeatable output.
blockLength	days, typically 200 is a good choice

Value

eBoot, a named list with bootOut,wordsOut,xConc,xFlux values. bootOut is a data frame with the results of the bootstrapping tests. wordsOut is a character vector describing the results. xConc, xFlux are vectors of length iBoot, of the change in flow normalized concentration or flux computed by each bootstrap replicate (mg/L). pConc and pFlux are vectors of length iBoot, of the change in flow normalized concentration or flux computed from each bootstrap replicate expressed as

See Also

runGroupsBoot, runPairs

Examples

```
library(EGRET)
eList <- Choptank_eList
year1 <- 1985
year2 <- 2009

## Not run:
pairOut_2 <- runPairs(eList, year1, year2, windowSide = 7)

boot_pair_out <- runPairsBoot(eList, pairOut_2)

plotHistogramTrend(eList, boot_pair_out, caseSetUp=NA)

## End(Not run)
```

saveEGRETci *Save EGRETci workspace after running wBT (the WRTDS bootstrap test)*

Description

Saves critical information in a EGRETci workflow when analyzing trends over a set of two years.

Usage

```
saveEGRETci(eList, eBoot, caseSetUp, fileName = "")
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes. Created from the EGRET package, after running modelEstimation.
eBoot	named list. Returned from wBT.
caseSetUp	data frame. Returned from trendSetUp.
fileName	character. If left blank (empty quotes), the function will interactively ask for a name to save.

See Also

wBT, trendSetUp, modelEstimation

Examples

```
library(EGRET)
eList <- Choptank_eList
## Not run:
caseSetUp <- trendSetUp(eList)
eBoot <- wBT(eList,caseSetUp)
saveEGRETci(eList, eBoot, caseSetUp)

## End(Not run)
```

setForBoot *Allows user to set window parameters for the WRTDS model prior to running the bootstrap procedure*

Description

Adds window parameters to INFO file in eList.

Usage

```
setForBoot(eList, caseSetUp, windowY = 7, windowQ = 2, windowS = 0.5,
  edgeAdjust = TRUE)
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes. Created from the EGRET package, after running modelEstimation.
caseSetUp	data frame returned from trendSetUp
windowY	numeric specifying the half-window width in the time dimension, in units of years, default is 7
windowQ	numeric specifying the half-window width in the discharge dimension, units are natural log units, default is 2
windowS	numeric specifying the half-window with in the seasonal dimension, in units of years, default is 0.5
edgeAdjust	logical specifying whether to use the modified method for calculating the windows at the edge of the record.

Value

eList list with Daily,Sample, INFO data frames and surface matrix.

Examples

```
library(EGRET)
eList <- Choptank_eList
## Not run:
caseSetUp <- trendSetUp(eList)
bootSetUp <- setForBoot(eList,caseSetUp)

## End(Not run)
```

---

trendSetUp *Interactive setup for running wBT, the WRTDS Bootstrap Test*

Description

Walks user through the set-up for the WRTDS Bootstrap Test. Establishes start and end year for the test period. Sets the minimum number of bootstrap replicates to be run, the maximum number of bootstrap replicates to be run, and the block length (in days) for the block bootstrapping.

Usage

```
trendSetUp(eList, ...)
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes. Created from the EGRET package, after running modelEstimation.
...	additional arguments to bring in to reduce interactive options (year1, year2, nBoot, bootBreak, blockLength)

Value

caseSetUp data frame with columns year1, yearData1, year2, yearData2, numSamples, nBoot, bootBreak, blockLength, confStop. These correspond to:

Column Name	Manuscript Variable
year1	y_s
year2	y_e
nBoot	M_max
bootBreak	M_min
blockLength	B

See Also

setForBoot, wBT

Examples

```
library(EGRET)
eList <- Choptank_eList
## Not run:
# Completely interactive:
caseSetUp <- trendSetUp(eList)
# Semi-interactive:
caseSetUp <- trendSetUp(eList, nBoot=100, blockLength=200)

## End(Not run)
```

---

wBT *Run the WBT (WRTDS Bootstrap Test)*

Description

Runs the WBT for a given data set to evaluate the significance level and confidence intervals for the trends between two specified years. The trends evaluated are trends in flow normalized concentration and flow normalized flux. Function produces text outputs and a named list (eBoot) that contains all of the relevant outputs. Check out runPairsBoot and runGroupsBoot for more bootstrapping options.

Usage

```
wBT(eList, caseSetUp, saveOutput = TRUE, fileName = "temp.txt",
  startSeed = 494817)
```

Arguments

eList	named list with at least the Daily, Sample, and INFO dataframes. Created from the EGRET package, after running modelEstimation.
caseSetUp	data frame. Returned from trendSetUp.
saveOutput	logical. If TRUE, a text file will be saved in the working directory.
fileName	character. Name to save the output file if saveOutput=TRUE.
startSeed	setSeed value. Defaults to 494817. This is used to make repeatable output.

Value

eBoot, a named list with bootOut,wordsOut,xConc,xFlux values. bootOut is a data frame with the results of the bootstrapping tests. wordsOut is a character vector describing the results. xConc, xFlux are vectors of length iBoot, of the change in flow normalized concentration or flux computed by each bootstrap replicate (mg/L). pConc and pFlux are vectors of length iBoot, of the change in flow normalized concentration or flux computed from each bootstrap replicate expressed as

See Also

trendSetUp, setForBoot, runGroupsBoot, runPairsBoot

Examples

```
library(EGRET)
eList <- Choptank_eList
## Not run:
  caseSetUp <- trendSetUp(eList,
  year1=1985,
  year2=2005,
  nBoot = 50,
  bootBreak = 39,
  blockLength = 200)
eBoot <- wBT(eList,caseSetUp)

## End(Not run)
```

Index