Package 'waterData'

October 4, 2012

Version 1.0

Date 2012-10-01

Title An R Package for Retrieval, Analysis, and Anomaly Calculation of Daily Hydrologic Time Series Data

Author Karen R. Ryberg and Aldo V. Vecchia

Maintainer Karen R. Ryberg <kryberg@usgs.gov>

License file LICENSE

Type Package

LazyLoad yes

Description
This package imports U.S. Geological Survey (USGS) daily hydrologic data from USGS web services, plots the data, addresses some common data problems, and calculates and plots anomalies.

Depends R (>= 2.14.1), lattice, latticeExtra, XML

Suggests xtable, maps, mapdata

URL http://pubs.usgs.gov/of/2012/1168/

Collate 'waterData-package.R' 'importDVs.R' 'fillMiss.R''summaryStats.R' 'compAnom.R' 'plotAnoms.R'

R topics documented:

`waterData-package` *Hydrologic Data Retrieval and Analysis and Anomaly Calculation*

Description

An R package for retrieval, analysis, and anomaly calculation of daily hydrologic time deries data.

Details

This package imports U.S. Geological Survey (USGS) daily hydrologic data from USGS web services, plots the data, addresses some common data problems, and calculates and plots anomalies. For a description of anomalies see Vecchia (2003), and for examples of the application of streamflow anomalies in trend analysis of nutrients, pesticides and surface water, see Alexander and Smith (2006), Ryberg and Vecchia (2006), Ryberg and others (2010), Sullivan and others (2009), Vecchia (2005), and Vecchia and others (2008).

Package:	waterData
Type:	Package
Version:	1.0
Date:	2012-10-01
License:	Unlimited for this package, dependencies have more restrictive licensing.
LazyLoad:	yes

Author(s)

Karen R. Ryberg <kryberg@usgs.gov> and Aldo V. Vecchia <avecchia@usgs.gov>

References

Alexander, R.B. and Smith, R.A., 2006, Trends in the nutrient enrichment of U.S. rivers during the late 20th century and their relation to changes in probable stream trophic conditions: Limnology and Oceanography, v. 51, no. 1, Part 2: Eutrophication of Freshwater and Marine Ecosystems, p. 639–654., accessed August 1, 2012, at http://www.jstor.org/stable/4499617.

Ryberg, K.R. and Vecchia, A.V., 2006, Water-quality trend analysis and sampling design for the Devils Lake Basin, North Dakota, January 1965 through September 2003: U.S. Geological Survey Scientific Investigations Report 2006–5238, 64 p., accessed August 1, 2012, at http://pubs.usgs.gov/sir/2006/5238/.

Ryberg, K.R. and Vecchia, A.V., 2012, waterData—An R package for retrieval, analysis, and anomaly calculation of daily hydrologic time series data, version 1.0: U.S. Geological Survey Open-File Report 2012–1168, 8 p. (Also available at http://pubs.usgs.gov/of/2012/1168/.)

Ryberg, K.R., Vecchia, A.V., Martin, J.D., Gilliom, R.J., 2010, Trends in pesticide concentrations in urban streams in the United States, 1992–2008: U.S. Geological Survey Scientific Investigations Report 2010-5139, 101 p., accessed August 1, 2012, at http://pubs.usgs.gov/sir/2010/5139/.

Sullivan, D.J., Vecchia, A.V., Lorenz, D.L., Gilliom, R.J., Martin, J.D., 2009, Trends in pesticide concentrations in corn-belt streams, 1996–2006: U.S. Geological Survey Scientific Investigations Report 2009-5132, 75 p., accessed August 1, 2012, at http://pubs.usgs.gov/sir/2009/5132/.

Vecchia, A.V., 2003, Relation between climate variability and stream water quality in the continental United States, Hydrological Science and Technology, v. 19 no. 1, 77–98.

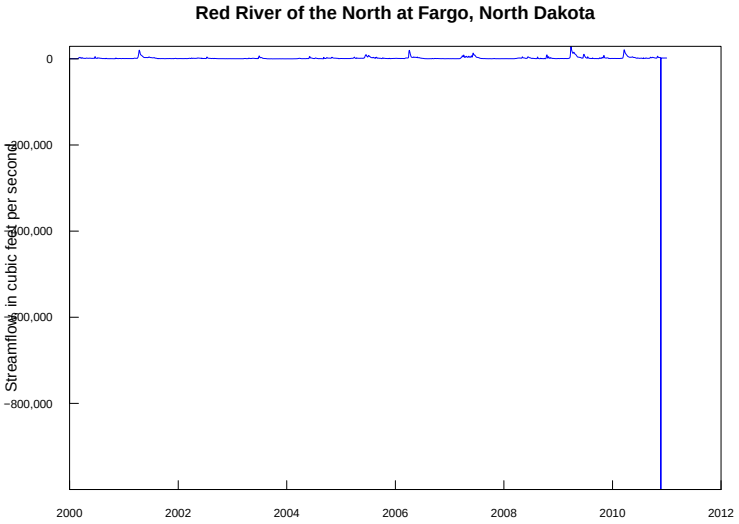
Vecchia, A.V., 2003, Water-quality trend analysis and sampling design for streams in North Dakota, 1971–2000: U.S. Geological Survey Scientific Investigations Report 2003–4094, 73 p., accessed August 1, 2012, at http://nd.water.usgs.gov/pubs/wri/wri034094/index.html.

Vecchia, A.V., 2005, Water-quality trend analysis and sampling design for streams in the Red River of the North Basin, Minnesota, North Dakota, and South Dakota, 1970–2001: U.S. Geological Survey Scientific Investigations Report 2005–5224, 54 p., accessed August 1, 2012, at http://pubs.usgs.gov/sir/2005/5224/.

`badDataSet` *Hydrologic dataset with provisional and negative data.*

Description

Daily mean streamflow for the Red River of the North at Fargo, North Dakota, streamgage 0504000, from January 1, 2000, to December 31, 2010. At the time these data were downloaded, some of the data were provisional and subject to revision, including a negative value.



Usage

```
badDataSet
```

Format

A dataframe containing 4,018 observations of daily mean streamflow. There are four variables shown in the following table.

Name	Type	Description
staid	factor	USGS Station identification number
val	numberic	Daily mean streamflow
dates	Date	Date of daily mean
qualcode	factor	Qualification code

Details

Streamflow data for U.S. Geological Survey streamgage 05054000, Red River of the North at Fargo, North Dakota, from January 1, 2000, to December 31, 2010, were downloaded, then 2,000 of the 4,018 daily values were randomly changed to missing values. At the time these data were downloaded, some of the data were provisional and subject to revision, including a negative value. For example purposes, the three smallest nonnegative values were changed to zero. Data for this site and specific dates have since been revised and approved. Definitions of the daily data qualification codes can be found at http://help.waterdata.usgs.gov/codes-and-parameters/daily-value-qualification-code-dv_rmk_cd (U.S. Geological Survey, 2011a).

Source

Imported from the USGS Daily Values Site Web Service http://waterservices.usgs.gov/rest/DV-Service.html (U.S. Geological Survey, 2011b) using the function importDVs, modified by replacing the three smallest values with zero. Data for this site and specific dates have since been revised and approved. The user would not be able to retrieve the same dataset from the USGS Daily Values Web Site Service.

References

U.S. Geological Survey, 2011a, Daily value qualification code (dv_rmk_cd): National Water Information System: Help System, accessed September 7, 2012, at http://help.waterdata.usgs.gov/codes-and-parameters/daily-value-qualification-code-dv_rmk_cd.

U.S. Geological Survey, 2011b, National Water Information System (NWISWeb): U.S. Geological Survey database, accessed November 2, 2011, at http://waterservices.usgs.gov/.

Examples

```
data(exampleWaterData)
par(las=1, tck=0.02, mar=c(5,5,4,2), mex=1.2, cex.axis=0.75)
plot(badDataSet$dates, badDataSet$val, type="l",
     ylab="Streamflow, in cubic feet per second",
     xlab="", yaxs='i', xaxs='i',
     xlim=c(as.Date("2000-01-01"), as.Date("2012-01-01")),
     yaxt="n", col="blue")
axis(2, axTicks(2), format(axTicks(2), big.mark=",", scientific=FALSE))
title("Red River of the North at Fargo, North Dakota")
summary(badDataSet$val)
```

---

cleanUp *Cleans up hydrologic time series data*

Description

Function to identify and fix common problems with hydrologic data

Usage

```
cleanUp(dataset, task = "view", replace = 0.1)
```

Arguments

dataset	is a data frame in format of the data frame returned by importDVs.
task	is either "view" or "fix." View will return a list containing rows with negative values and rows with missing values. Fix will replace negative values with NA and replace zeroes with the value specified by the replace argument.
replace	is the value used to replace 0 values. The default is 0.1. For streamflow in small streams, one might want to use 0.01. For daily data other than streamflow, such as turbidity, users may not want to replace 0 values with a nonzero value. In those cases, replace can be set to 0.

Value

A list showing days with negative values and days with values of 0 when task is "view." When task is "fix" the fixed dataset is returned. When a negative value is replaced with NA, an "N" is added to the qualcode field to indicate that there had been a negative number. When a zero value is replaced, an "R" is added to the qualcode field to indicate that a zero value was replaced.

Note

If calculating anomalies (see compAnom), the user may need to replace isolated missing values with with a value; however, if there are larger periods with missing values, streamflow anomalies may not be an appropriate use of the data. The substitution of some missing data with values may be done using the function fillMiss that is part of this package. However, care needs to be taken when filling in missing data.

See Also

fillMiss

Examples

```
data(exampleWaterData)
cleanUp(badDataSet, task="view")
q05054000Fix <- cleanUp(badDataSet, task="fix")
summary(badDataSet)
summary(q05054000Fix)
```

---

compAnom *Calculates anomalies*

Description

Function to calculate short-, medium-, and long-term hydrologic anomalies

Usage

```
compAnom(dataset, which = 1)
```

Arguments

dataset	is the daily hydrologic data returned from importDVs or data otherwise obtained and in the same format as that produced by importDVs.
which	indicates which set of anomalies; 1 calculates the 1-year, 30-day, and 1-day anomalies; 2 calculates the 100-day, 10-day, and 1-day anomalies; 3 calculates the 30-day and 1-day anomalies; and 4 calculates the 10-year, 5-year, 1-year, one-quarter-year (seasonal), and 1-day anomalies.

Details

This function was written with streamflow data in mind because streamflow is the most commonly used exogenous variable for trend models for water quality; however, the function is generic so that users may experiment with anomalies from other daily hydrologic data. Examples of the inclusion of streamflow anomalies in trend analysis of nutrients, pesticides and surface water can be found in Alexander and Smith (2006), Ryberg and Vecchia (2006), Ryberg and others (2010), Sullivan and others (2009), Vecchia (2003), Vecchia (2005), and Vecchia and others (2008).

Value

A list. In the cases of "which" equal to 1 or 2, the first element of the list is a data frame containing the station identification number, dates, streamflow, and long-term, mid-term, and short-term anomalies. The next three elements of the list are the length in days of the long-term, mid-term, and short-term streamflow anomalies. In the case of "which" equal to 3, the first element of the list is a data frame containing the station identification number, dates, streamflow, and mid-term and short-term anomalies. The next two elements of the list are the length in days of the mid-term and short-term streamflow anomalies. In the case of "which" equal to 4, the first element of the list is a data frame containing the station identification number, dates, streamflow, and 10-year, 5-year, annual, seasonal, and daily streamflow anomalies. The next five elements of the list are the length in days of the 10-year, 5-year, annual, seasonal, and daily streamflow anomalies.

References

Alexander, R.B. and Smith, R.A., 2006, Trends in the nutrient enrichment of U.S. rivers during the late 20th century and their relation to changes in probable stream trophic conditions: Limnology and Oceanography, v. 51, no. 1, Part 2: Eutrophication of Freshwater and Marine Ecosystems, p. 639–654, accessed August 1, 2012, at http://www.jstor.org/stable/4499617.

Ryberg, K.R. and Vecchia, A.V., 2006, Water-quality trend analysis and sampling design for the Devils Lake Basin, North Dakota, January 1965 through September 2003: U.S. Geological Survey Scientific Investigations Report 2006–5238, 64 p., accessed August 1, 2012, at http://pubs.usgs.gov/sir/2006/5238/.

Ryberg, K.R. and Vecchia, A.V., 2012, waterData—An R packge for retrieval, analysis, and anomaly calculation of daily hydrologic time series data, version 1.0: U.S. Geological Survey Open-File Report 2012–1168, 8 p. (Also available at http://pubs.usgs.gov/of/2012/1168/.)

Ryberg, K.R., Vecchia, A.V., Martin, J.D., Gilliom, R.J., 2010, Trends in pesticide concentrations in urban streams in the United States, 1992-2008: U.S. Geological Survey Scientific Investigations Report 2010–5139, 101 p. (Also available at http://pubs.usgs.gov/sir/2010/5139/.)

Sullivan, D.J., Vecchia, A.V., Lorenz, D.L., Gilliom, R.J., Martin, J.D., 2009, Trends in pesticide concentrations in corn-belt streams, 1996–2006: U.S. Geological Survey Scientific Investigations Report 2009–5132, 75 p., accessed Ausugst , 2012, at http://pubs.usgs.gov/sir/2009/5132/.

Vecchia, A.V., 2003, Relation between climate variability and stream water quality in the continental United States, Hydrological Science and Technology, v. 19, no. 1, 77–98.

Vecchia, A.V., 2003, Water-quality trend analysis and sampling design for streams in North Dakota, 1971–2000: U.S. Geological Survey Scientific Investigations Report 2003–4094, 73 p., accessed August 1, 2012, at http://nd.water.usgs.gov/pubs/wri/wri034094/index.html.

Vecchia, A.V., 2005, Water-quality trend analysis and sampling design for streams in the Red River of the North Basin, Minnesota, North Dakota, and South Dakota, 1970–2001: U.S. Geological Survey Scientific Investigations Report 2005–5224, 54 p. accessed August 1, 2012, at http://pubs.usgs.gov/sir/2005/5224/.

Examples

```
## Not run:
q05054000.85 <- importDVs("05054000", sdate="1985-10-01", edate="2010-09-30")
anoms05054000.1 <- compAnom(q05054000.85, which=1)
anoms05054000.2 <- compAnom(q05054000.85, which=2)
anoms05054000.3 <- compAnom(q05054000.85, which=3)
anoms05054000.4 <- compAnom(q05054000.85, which=4)

## End(Not run)
```

---

fillMiss *Fill-in missing hydrological values*

Description

Function to fill in missing time series data.

Usage

```
fillMiss(dataset, block = 30, pmiss = 40,
  model = "trend", smooth = TRUE, ...)
```

Arguments

dataset	is a data frame in the format of the data frame returned by importDVs, with missing values indicated by NA.
block	is the size of the largest block of missing data that the function will fill-in.
pmiss	is the maximum amount of the missing data that can be missing in the dataset for fill-in procedure to be performed.
model	is the type of structural time series model, see StructTS. The default value is trend. If level is used, the results of fillMiss, which by default applies a fixed-interval smoothing to the time series, tsSmooth, will be very close to linear interpolation.
smooth	a logical that indicates whether or not to apply tsSmooth to the structured time series.
...	further arguments to be passed to plotting method (see par).

Format

The returned data frame has the following columns:

Name	Type	Description
staid	factor	USGS station identification number
val	numeric	The value of the hydrologic variable
dates	Date	Date of daily value
qualcode	factor	Qualification code

Details

This function will check the percent of missing values and the size of the largest missing block of data. By default, if less than 40 percent of the data are missing and the largest block is less than 30 days, the data will be filled-in by using a structural time series, StructTS from the base stats package in R (R Development Core Team, 2012). The fitted structural time series is then smoothed via a state-space model, tsSmooth from the base stats package in R.

Value

a data frame with NAs in the "val" column replaced by estimated values and a plot showing observed and estimated data. If there are too many missing values, based on default or user defined limits, the unaltered dataset is returned as well as a message, such as "Error in fillMiss(misQ05054000) : Too much missing data. Cannot fill in missing values."

Note

Many methods have been suggested for estimating missing hydrological data. However, experiments showed that the functions in the base stats package worked very well if the blocks of missing data were not long. Users with larger blocks of missing data may want to explore other methods including using nearby gages to estimate missing values at a streamgage. Additional methods for filling in missing hydrological data are summarized in Beauchamp and others (1989) and Elshorbagy and others (2000).

References

Beauchamp, J.J., 1989, Comparison of regression and time-series methods for synthesizing missing streamflow records: Water Resources Bulletin, v. 25, no. 5, p. 961–975.

Elshorbagy, A.A., Panu, U.S., Simonovic, S.P., 2000, Group-based estimation of missing hydrological data—I. Approach and general methodology: Hydrological Sciences Journal, v. 45, no. 6, p. 849–866.

R Development Core Team, 2012, R—A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing, [ISBN 3-900051-07-0]. (Also available at http://www.R-project.org/.)

See Also

StructTS, tsSmooth, cleanUp

Examples

```
data(exampleWaterData)
my.newdata <- fillMiss(misQ05054000, block=30, pmiss=50, log="y")
 ## Not run:
  # ph example
  pH05082500<-importDVs("05082500", code="00400", stat="00008",
  sdate="2000-01-01", edate="2011-12-31")
```

```
  plotParam(pH05082500)
  pHfilled<-fillMiss(pH05082500, block=45, ylim=c(7.5,9), yaxs="i")

## End(Not run)
```

`importDVs` *Imports daily USGS hydrologic times series data*

Description

Function to import daily hydrologic time series data given a USGS streamgage identification number.

Usage

```
importDVs(staid, code = "00060", stat = "00003",
  sdate = "1851-01-01",
  edate = as.Date(Sys.Date(), format = "%Y-%m-%d"))
```

Arguments

`staid`	is the USGS site identification number, usually eight digits long, but can be longer. Users may search for surface-water sites and obtain station identification numbers using the USGS Site Web Service, http://waterservices.usgs.gov/rest/Site-Service.html (USGS, 2012e); using the National Water Information System: Mapper, http://wdr.water.usgs.gov/nwisgmap/ (U.S. Geological Survey, 2012a); or using the National Water Information System: Web Interface to daily surface-water data, http://waterdata.usgs.gov/nwis/dv/?referred_module=sw (U.S. Geological Survey, 2012f). The site identification number needs to be entered as a character, that is in quotes, because many USGS streamgage numbers begin with zero and the leading zero is necessary.
`code`	is the USGS parameter code, a 5-digit number used in the USGS computerized data system, National Water Information System (NWIS), to uniquely identify a specific hydrologic property or constituent. A list of paramater codes is available at http://nwis.waterdata.usgs.gov/usa/nwis/pmcodes (U.S. Geological Survey, 2012b).
`stat`	is the USGS statistics code, a 5-digit number used in the USGS computerized data system, National Water Information System (NWIS), to uniquely identify specific statistics, such as daily mean, daily maximum, and daily minimum. The default, 00003, is the mean daily value. A list of statistics codes is available at http://nwis.waterdata.usgs.gov/nwis/help/?read_file=stat&format=table (U.S. Geological Survey, 2012c). Not all statistics are available at every gage.
`sdate`	is the start date of the time series, in the format yyyy-mm-dd, optional.
`edate`	is the end date of the time series, in the format yyyy-mm-dd, optional.

Format

The returned data frame has the following columns

Name	Type	Description
staid	factor	USGS station identification number
val	numeric	The value of the hydrologic variable
dates	Date	Date of daily value
qualcode	factor	Qualification code

Details

This function will import data from a WaterML1 service (current USGS hydrological data standard). It will retrieve daily streamflow and continuous water-quality data from the USGS Daily Values Site Web Service http://waterservices.usgs.gov/rest/DV-Service.html (U.S. Geological Survey, 2012d).

Value

a data frame containing daily streamflow or other hydrologic data for the site specified during the dates specified (inclusive). The USGS parameter code, code, and the statistics code, stat, are attributes of the data frame.

References

U.S. Geological Survey, 2012a, National Water Information System: Mapper, accessed September 7, 2012, at http://wdr.water.usgs.gov/nwisgmap/.

U.S. Geological Survey, 2012b, Parameter code definition, National Water Information System: Web Interface, accessed September 7, 2012, at http://nwis.waterdata.usgs.gov/usa/nwis/pmcodes.

U.S. Geological Survey, 2012c, Stat codes (stat_cd), National Water Information System: Web Interface, accessed September 7, 2012, at http://nwis.waterdata.usgs.gov/nwis/help/?read_file=stat&format=table.

U.S. Geological Survey, 2012d, USGS daily values site web service: REST Web Services, accessed September 7, 2012, at http://waterservices.usgs.gov/rest/DV-Service.html.

U.S. Geological Survey, 2012e, USGS site web service: REST Web Services, accessed September 7, 2012, at http://waterservices.usgs.gov/rest/Site-Service.html.

U.S. Geological Survey, 2012f, USGS surface-water daily data for the Nation: National Water Information System: Web Interface, accessed September 7, 2012, at http://waterdata.usgs.gov/nwis/dv/?referred_module=sw.

Examples

```
## Not run:
# import mean daily streamflow for Red River of the North at Fargo, ND
q05054000 <- importDVs("05054000", sdate="2000-01-01", edate="2010-12-31")
head(q05054000)
# additional examples of how to this function follow
# import mean daily gage height for Red River of the North at Grand Forks, ND
gh05082500 <- importDVs("05082500", code="00065", sdate="2000-01-01", edate="2010-12-31")
# import mean daily specific conductance for Red River of the North at Grand Forks, ND
sc05082500<- importDVs("05082500", code="00095", sdate="2000-01-01", edate="2010-12-31")
# import mean daily water temperature for Red River of the North at Fargo, ND
temp05054000<- importDVs("05054000", code="00010", sdate="2000-01-01", edate="2010-12-31")
# import median daily pH for Red River of the North at Fargo, ND
pH05054000<- importDVs("05054000", code="00400", stat="00008", sdate="2000-01-01", edate="2010-12-31")
# examine the attributes of the data frame to show that the parameter code
```

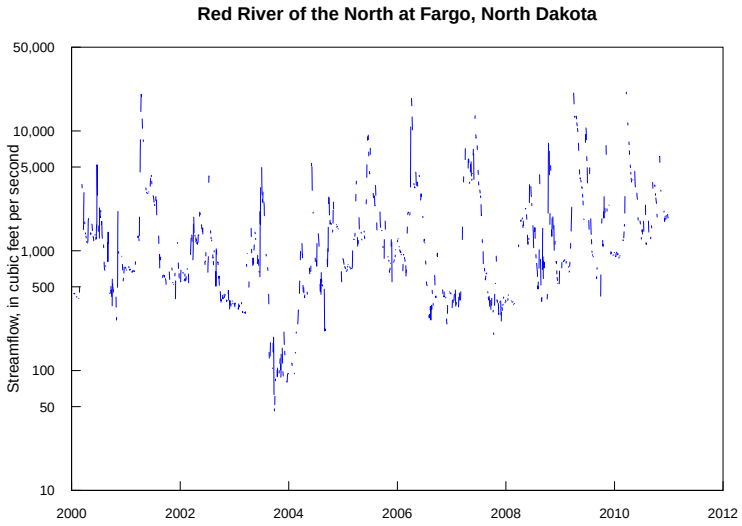
```
# and statistics code are saved with the data frame
attributes(pH05054000)[c("code","stat")]
# import mean daily oxygen for Red River of the North at Fargo, ND
do05054000 <- importDVs("05054000", code="00300", sdate="2000-01-01", edate="2010-12-31")
# import mean daily turbidity for Red River of the North at Fargo, ND
turb05054000 <- importDVs("05054000", code="63680", sdate="2000-01-01", edate="2010-12-31")

## End(Not run)
```

---

misQ05054000 *Daily mean streamflow with missing values*

Description

Daily mean streamflow data with randomly generated missing values.



Usage

```
misQ05054000
```

Format

A dataframe containing 4,018 observations of daily mean streamflow. There are four variables, staid, the USGS streamgage station number; val, the daily mean streamflow value in cubic feet per second; dates, the dates of the observations; and qualcode, the USGS data qualification code.

Name	Type	Description
staid	factor	USGS station identification number
val	numberic	Daily mean streamflow
dates	Date	Date of daily mean
qualcode	factor	Qualification code

Details

Streamflow data for USGS streamgage 05054000, Red River of the North at Fargo, North Dakota, from January 1, 2000, to December 31, 2010, were downloaded and then 2,000 of the 4,018 daily values were randomly replaced with NAs. This provides a dataset with slightly less than 50 percent of the values missing; however, the blocks of missing values are less than 30 days long. Definitions of the daily data qualification codes can be found at http://help.waterdata.usgs.gov/codes-and-parameters/daily-value-qualification-code-dv_rmk_cd (U.S. Geological Survey, 2011).

Source

Imported from the USGS Daily Values Site Web Service http://waterservices.usgs.gov/rest/DV-Service.html (U.S. Geological Survey, 2012) using the function importDVs, with random replacements of data with NAs.

References

U.S. Geological Survey, 2011, Daily value qualification code (dv_rmk_cd): National Water Information System: Help System, accessed September 7, 2012, at http://help.waterdata.usgs.gov/codes-and-parameters/daily-value-qualification-code-dv_rmk_cd.

U.S. Geological Survey, 2012, National Water Information System (NWISWeb): U.S. Geological Survey database, accessed May 2, 2012, at http://waterservices.usgs.gov/.

Examples

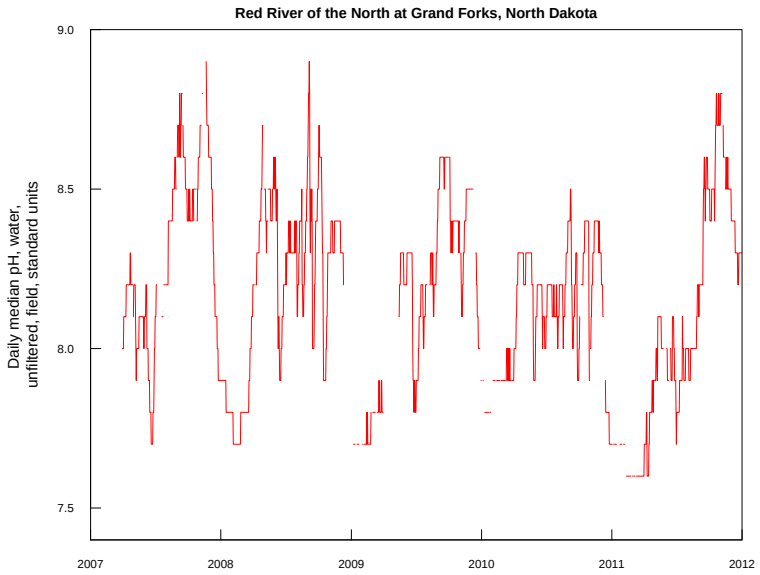
```
data(exampleWaterData)
par(las=1, tck=0.02, mar=c(5,5,4,2), mex=1.2, cex.axis=.9)
plot(misQ05054000$dates, misQ05054000$val, type="l", log="y",
     ylab="Streamflow, in cubic feet per second",
     xlab="", yaxs='i', xaxs='i',
     xlim=c(as.Date("2000-01-01"), as.Date("2012-01-01")), ylim=c(10,50000),
     yaxt="n")
axis(2, axTicks(2), format(axTicks(2), big.mark=","))
title("Red River of the North at Fargo, North Dakota")
summary(misQ05054000$val)
```

---

pH05082500 *Hydrologic dataset with provisional and negative data.*

Description

Daily median pH for the Red River of the North at Grand Forks, North Dakota, streamgage 05082500, from March 31, 2001, to December 31, 2011. At the time these data were downloaded, some of it was provisional data subject to revision.



Usage

```
pH05082500
```

Format

A dataframe containing 1,737 rows. There are four variables, staid, the USGS streamgage station number; val, the daily meadian pH value in standard unites; dates, the dates of the observations; and qualcode, the USGS data qualification code.

Name	Type	Description
staid	factor	USGS station identification number
val	numeric	Daily median pH
dates	Date	Date of daily value
qualcode	factor	Qualification code

Details

pH data for USGS streamgage 05082500, Red River of the North at Grand Forks, North Dakota, from March 31, 2001, to December 31, 2011. At the time these data were downloaded, some of the data were provisional and subject to revision, indicated by "P" in the qualcode column. Definitions of the daily data qualification codes can be found at http://help.waterdata.usgs.gov/codes-and-parameters/daily-value-qualification-code-dv_rmk_cd (U.S. Geological Survey, 2011).

Source

Imported from the USGS Daily Values Site Web Service http://waterservices.usgs.gov/rest/DV-Service.html (U.S. Geological Survey, 2012) using the function importDVs, July 23, 2012.

References

U.S. Geological Survey, 2011, Daily value qualification code (dv_rmk_cd): National Water Information System: Help System, accessed September 7, 2012, at http://help.waterdata.usgs.gov/codes-and-parameters/daily-value-qualification-code-dv_rmk_cd.

U.S. Geological Survey, 2012, National Water Information System (NWISWeb): U.S. Geological Survey database, accessed July 23, 2012, at http://waterservices.usgs.gov/.

Examples

```
data(exampleWaterData)
par(las=1, tck=0.02, mar=c(3,5,2,2), cex.axis=0.75)
plot(pH05082500$dates, pH05082500$val, type="l",
     ylab="Daily median pH, water, \nunfiltered, field, standard units",
     xlab="", yaxs='i', xaxs='i', ylim=c(7.4, 9),
     xlim=c(as.Date("2007-01-01"), as.Date("2012-01-01")), col="red")
title("Red River of the North at Grand Forks, North Dakota",cex.main=0.95)
summary(pH05082500$val)
```

`plotAnoms` *Plots streamflow anomalies*

Description

Function to plot hydrologic time series and anomalies

Usage

```
plotAnoms(data, ...)
```

Arguments

`data`	is the anomaly list from the function compAnom.
`...`	further arguments to be passed to plotting method (see par).

Value

a plot.

Examples

```
## Not run:
 q05054000.85 <- importDVs("05054000", sdate="1985-01-01", edate="2010-09-30")
 anoms05054000 <- compAnom(q05054000.85, which=1)
 plotAnoms(anoms05054000)

## End(Not run)
```

plotParam *Plot Streamflow and Continous Water-Quality Data*

Description

Function to plot hydrologic times series. Will plot more than one site at a time.

Usage

```
plotParam(data, logscale = FALSE, metric = FALSE,
  ylabel = NULL, ...)
```

Arguments

data	is the data frame in the foramt of that returned by importDVs.
metric	USGS streamflow data are usually in cubic feet per second; however it may be converted to cubic meters per second for publication. Likewise, gage height is usually in feet, but could be converted to meters. The metric argument only has an effect on streamflow and gage height.
logscale	is a logical indicating whether or not the y-scale should be log 10. Streamflow generally is plotted with a log scale and this only has an effect on the plotting of streamflow data.
ylabel	optionally allows user to pass a y-axis label.
...	further arguments to be passed to plotting method (see par). (see xyplot).

Value

a lattice plot

Examples

```
data(exampleWaterData)
plotParam(misQ05054000, code="00060", stat="00003", logscale=TRUE)
plotParam(misQ05054000, code=attributes(misQ05054000)$code,
stat=attributes(misQ05054000)$stat, logscale=TRUE)
```

siteInfo *Retrieve site information*

Description

Function to retrieve information about a USGS streamgage site

Usage

```
siteInfo(staid)
```

Arguments

staid	is a character vector containing USGS site identification number(s). USGS site numbers are usually eight digits long, but can be longer. Users may search for surface-water sites and obtain station identification numbers using the USGS Site Web Service, http://waterservices.usgs.gov/rest/Site-Service.html (U.S. Geological Survey, 2012b); using the National Water Information System: Mapper, http://wdr.water.usgs.gov/nwisgmap/ (U.S. Geological Survey, 2012a); or using the National Water Information System: Web Interface to daily surface-water data, http://waterdata.usgs.gov/nwis/dv/?referred_module=sw (U.S. Geological Survey, 2012c). The site identification number needs to be entered as a character, that is in quotes, because many USGS streamgage numbers begin with zero and the leading zero is necessary.

Format

a data frame with the following columns:

Name	Type	Description
staid	factor	USGS station identification number
staname	character	USGS station name
lat	numeric	Decimal latitude
lng	numeric	Decimal longitude

Details

This provides some limited metadata about the USGS streamgage site.

Value

a data frame containing the station identification number(s), the USGS streamgage name(s), the decimal latitude(s), and decimal longitude(s).

Note

Information retrieved using this function can be used to create a map of multiple streamgage sites—see package vignette.

References

U.S. Geological Survey, 2012a, National Water Information System: Mapper, accessed September 7, 2012, at http://wdr.water.usgs.gov/nwisgmap/.

U.S. Geological Survey, 2012b, USGS site web service: REST Web Services, accessed September 7, 2012, at http://waterservices.usgs.gov/rest/Site-Service.html.

U.S. Geological Survey, 2012c, USGS surface-water daily data for the Nation: National Water Information System: Web Interface, accessed September 7, 2012, at http://waterdata.usgs.gov/nwis/dv/?referred_module=sw.

Examples

```
## Not run:
 staInfo <- siteInfo("05054000")
 staInfo
 staInfo <- siteInfo(c("05054000", "05082500", "06342500"))
```

```
  staInfo
  # a list with an invalid station identification number
  staInfo <- siteInfo(c("05054000", "05082500", "06342501"))

## End(Not run)
```

`summaryStats` *Calculate summary statistics*

Description

Function to calculate summary statistics for daily hydrologic time series.

Usage

```
summaryStats(dataset, staid = 1)
```

Arguments

`dataset`	is the data frame containing hydrologic data
`staid`	is used to label the output

Format

The returned matrix has the following columns, which are formatted for putting in a report or table.

Name	Type	Description
Begin	character	The beginning date of the time series
End	character	The ending date of the time series
n	character	Number of rows
NA	character	Number of missing values
Neg	character	Number of negative values
Min	character	The minimum value
Q1	character	The first quartile, 25th percentile
Med	character	The median
Mean	character	The mean
Q3	character	The third quartile, 75th percentile
Max	character	The maximum value
StdDev	character	The standard deviation
IQR	character	The interquartile range

Details

The summary statistics returned are useful for exploratory data analysis and for describing the date set.

Value

a data frame containing a number of summary statistics of the daily hydrologic data series

Note

Hydrologic data are often skewed (Helsel and Hirsch, 2002). Summary statistics help describe the degree of skewness and help to determine the degree of applicability of hypothesis tests. Some data, in particular streamflow, may need to be transformed to produce approximately normal data.

References

Helsel, D.R. and Hirsch, R. M., 2002, Statistical methods in water resources: U.S. Geolgical Survey Techniques of Water Resources Investigations, book 4, chap. A3, 522 p. (Also available at http://pubs.usgs.gov/twri/twri4a3/).

Examples

```
data(exampleWaterData)
summaryStats(pH05082500, staid="05082500")
```

---

tellMeSiteURL *USGS Site Information Service URL*

Description

Function that returns USGS Site Information Service URL for troubleshooting or building a URL for other purposes.

Usage

```
tellMeSiteURL(staid)
```

Arguments

staid	is the USGS site identification number, which is usually eight digits long, but can be longer. Users may search for surface-water sites and obtain station identification numbers using the USGS Site Web Service, http://waterservices.usgs.gov/rest/Site-Service.html (U.S. Geological Survey, 2012b); using the National Water Information System: Mapper, http://wdr.water.usgs.gov/nwisgmap/ (U.S. Geological Survey, 2012a); or using the National Water Information System: Web Interface to daily surface-water data, http://waterdata.usgs.gov/nwis/dv/?referred_module=sw (U.S. Geological Survey, 2012c). The site identification number needs to be entered as a character, that is in quotes, because many USGS streamgage numbers begin with zero and the leading zero is necessary.

Value

URL for USGS site information

References

U.S. Geological Survey, 2012a, National Water Information System: Mapper, accessed September 7, 2012, at http://wdr.water.usgs.gov/nwisgmap/.

U.S. Geological Survey, 2012b, USGS site web service: REST Web Services, accessed September 7, 2012, at http://waterservices.usgs.gov/rest/Site-Service.html.

U.S. Geological Survey, 2012c, USGS surface-water daily data for the Nation: National Water Information System: Web Interface, accessed September 7, 2012, at http://waterdata.usgs.gov/nwis/dv/?referred_module=sw.

Examples

```
tellMeSiteURL("05054000")
```

`tellMeURL` *USGS Daily Values Site Service URL*

Description

Function that returns USGS Daily Values Site Service URL for troubleshooting or building a URL for other purposes.

Usage

```
tellMeURL(staid, code = "00060", stat = "00003",
  sdate = "1851-01-01",
  edate = as.Date(Sys.Date(), format = "%Y-%m-%d"))
```

Arguments

staid	is the USGS site identification number, which is usually eight digits long, but can be longer. Users may search for surface-water sites and obtain station identification numbers using the USGS Site Web Service, http://waterservices.usgs.gov/rest/Site-Service.html (U.S. Geological Survey, 2012d); using the National Water Information System: Mapper, http://wdr.water.usgs.gov/nwisgmap/ (U.S. Geological Survey, 2012a); or using the National Water Information System: Web Interface to daily surface-water data, http://waterdata.usgs.gov/nwis/dv/?referred_module=sw (U.S. Geological Survey, 2012e). The site identification number needs to be entered as a character, that is in quotes, because many USGS streamgage numbers begin with zero and the leading zero is necessary.
code	is the USGS parameter code, a 5-digit number used in the USGS computerized data system, National Water Information System (NWIS), to uniquely identify a specific hydrologic property or constituent. A list of paramater codes is available at http://nwis.waterdata.usgs.gov/usa/nwis/pmcodes (U.S. Geological Survey, 2012b).
stat	is the USGS statistics code, a 5-digit number used in the USGS computerized data system, NWIS, to uniquely identify specific statistics, such as daily mean, daily maximum, and daily minimum. The default, 00003, is the mean daily value. A list of statistics codes is available at http://nwis.waterdata.usgs.gov/nwis/help/?read_file=stat&format=table (U.S. Geological Survey, 2012c). Not all statistics are available at every gage.

sdate	is the start date of the time series, in the format yyyy-mm-dd.
edate	is the end date of the time series, in the format yyyy-mm-dd.

Value

URL for USGS data

References

U.S. Geological Survey, 2012a, National Water Information System: Mapper, accessed September 7, 2012, at http://wdr.water.usgs.gov/nwisgmap/.

U.S. Geological Survey, 2012b, Parameter code definition, National Water Information System: Web Interface, accessed September 7, 2012, at http://nwis.waterdata.usgs.gov/usa/nwis/pmcodes.

U.S. Geological Survey, 2012c, Stat codes (stat_cd), National Water Information System: Web Interface, accessed September 7, 2012, at http://nwis.waterdata.usgs.gov/nwis/help/?read_file=stat&format=table.

U.S. Geological Survey, 2012d, USGS site web service: REST Web Services, accessed September 7, 2012, at http://waterservices.usgs.gov/rest/Site-Service.html.

U.S. Geological Survey, 2012e, USGS surface-water daily data for the Nation: National Water Information System: Web Interface, accessed September 7, 2012, at http://waterdata.usgs.gov/nwis/dv/?referred_module=sw.

Examples

```
tellMeURL("05054000", code="00060", stat="00003", sdate="2000-01-01",
 edate=as.Date(Sys.Date(), format="%Y-%m-%d"))
```

Index