Package 'hddtools'

May 25, 2020

Title Hydrological Data Discovery Tools

Version 0.9.1

Maintainer Claudia Vitolo <cvitolodev@gmail.com>

URL <http://docs.ropensci.org/hddtools>,

   <https://github.com/ropensci/hddtools>

BugReports <https://github.com/ropensci/hddtools/issues>

Description Tools to discover hydrological data, accessing catalogues and databases from various data providers.

Depends R (>= 3.2.1), rgdal

Imports zoo, sp, curl, XML, rnrfa, raster, readxl, tidyr

Suggests testthat, leaflet, rmarkdown, knitr, pkgdown, dplyr

VignetteBuilder knitr

License GPL-3

Repository CRAN

RoxygenNote 7.1.0

Encoding UTF-8

NeedsCompilation no

Author Claudia Vitolo [aut, cre] (<https://orcid.org/0000-0002-4252-1176>),
   Wouter Buytaert [ctb] (Supervisor),
   Erin Le Dell [ctb] (Erin reviewed the package for rOpenSci, see
   https://github.com/ropensci/onboarding/issues/73),
   Michael Sumner [ctb] (Michael reviewed the package for rOpenSci, see
   https://github.com/ropensci/onboarding/issues/73)

Date/Publication 2020-05-25 10:50:02 UTC

R topics documented:

Index   **13**

---

bboxSpatialPolygon   *Convert a bounding box to a SpatialPolygons object Bounding box is*
                            first created (in lat/lon) then projected if specified

---

Description

Convert a bounding box to a SpatialPolygons object Bounding box is first created (in lat/lon) then
projected if specified

Usage

```
bboxSpatialPolygon(boundingbox, proj4stringFrom = NULL, proj4stringTo = NULL)
```

Arguments

boundingbox   Bounding box: a 2x2 numerical matrix of lat/lon coordinates

proj4stringFrom

                  Projection string for the current boundingbox coordinates (defaults to lat/lon,
                  WGS84)

proj4stringTo   Projection string, or NULL to not project

Value

A SpatialPolygons object of the bounding box

References

<https://bit.ly/2NA5PSF>

Examples

```
## Not run:
  boundingbox <- raster::extent(-180, +180, -50, +50)
  bbSP <- bboxSpatialPolygon(boundingbox = boundingbox)

## End(Not run)
```

---

catalogueData60UK   *Data source: Data60UK catalogue*

Description

This function interfaces the Data60UK database catalogue listing 61 gauging stations.

Usage

```
catalogueData60UK(areaBox = NULL)
```

Arguments

areaBox   bounding box, a list made of 4 elements: minimum longitude (lonMin), mini-
                 mum latitude (latMin), maximum longitude (lonMax), maximum latitude (lat-
                 Max)

Value

This function returns a data frame containing the following columns:

id Station id number.

River String describing the river's name.

Location String describing the location.

gridReference British National Grid Reference.

Latitude

Longitude

Author(s)

Claudia Vitolo

Source

<http://nrfaapps.ceh.ac.uk/datauk60/data.html>

Examples

```
## Not run:
  # Retrieve the whole catalogue
  Data60UK_catalogue_all <- catalogueData60UK()

  # Filter the catalogue based on a bounding box
  areaBox <- raster::extent(-4, -2, +52, +53)
  Data60UK_catalogue_bbox <- catalogueData60UK(areaBox)

## End(Not run)
```

---

catalogueGRDC   *Data source: Global Runoff Data Centre catalogue*

Description

This function interfaces the Global Runoff Data Centre database which provides river discharge data for almost 1000 sites over 157 countries.

Usage

```
catalogueGRDC()
```

Value

This function returns a data frame made with the following columns:

- `grdc_no`: GRDC station number
- `wmo_reg`: WMO region
- `sub_reg`: WMO subregion
- `river`: river name
- `station`: station name
- `country`: 2-letter country code (ISO 3166)
- `lat`: latitude, decimal degree
- `long`: longitude, decimal degree
- `area`: catchment size, km2
- `altitude`: height of gauge zero, m above sea level
- `d_start`: daily data available from year
- `d_end`: daily data available until year
- `d_yrs`: length of time series, daily data
- `d_miss`: percentage of missing values (daily data)
- `m_start`: monthly data available from
- `m_end`: monthly data available until
- `m_yrs`: length of time series, monthly data
- `m_miss`: percentage of missing values (monthly data)
- `t_start`: earliest data available
- `t_end`: latest data available
- `t_yrs`: maximum length of time series, daily and monthly data
- `lta_discharge`: mean annual streamflow, m3/s
- `r_volume_yr`: mean annual volume, km3
- `r_height_yr`: mean annual runoff depth, mm

Author(s)

Claudia Vitolo

Examples

```
## Not run:
  # Retrieve the catalogue
  GRDC_catalogue_all <- catalogueGRDC()

## End(Not run)
```

---

catalogueMOPEX   *Data source: MOPEX catalogue*

Description

This function retrieves the list of the MOPEX basins.

Usage

```
catalogueMOPEX(MAP = TRUE)
```

Arguments

MAP   Boolean, TRUE by default. If FALSE it returns a list of the USGS station ID's
               and the gage locations of all 1861 potential MOPEX basins. If TRUE, it return
               a list of the USGS station ID's and the gage locations of the 438 MOPEX basins
               with MAP estimates.

Value

This function returns a data frame containing the following columns:

USGS_ID Station id number

Longitude Decimal degrees East

Latitude Decimal degrees North

Drainage_Area Square Miles

R_gauges Required number of precipitation gages to meet MAP accuracy criteria

N_gauges Number of gages in total gage window used to estimate MAP

A_gauges Avaliable number of gages in the basin

Ratio_AR Ratio of Available to Required number of gages in the basin

Date_start Date when recordings start

Date_end Date when recordings end

State State of the basin

Name Name of the basin

Columns Date_start, Date_end, State, Name are taken from: https://hydrology.nws.noaa.gov/pub/gcip/mopex/US_Data/Basin
Date_start and Date_end are conventionally set to the first of the month here, however actual recordings my differ. Always refer to the recording date obtained as output of `tsMOPEX()`.

Author(s)

Claudia Vitolo

Source

https://hydrology.nws.noaa.gov/pub/gcip/mopex/US_Data/Documentation/

Examples

```
## Not run:
  # Retrieve the MOPEX catalogue
  catalogue <- catalogueMOPEX()

## End(Not run)
```

---

catalogueSEPA   *Data source: SEPA catalogue*

Description

This function provides the official SEPA database catalogue of river level data (from https://www2.sepa.org.uk/waterlevels/CS
containing info for hundreds of stations. Some are NRFA stations. The function has no input arguments.

Usage

```
catalogueSEPA()
```

Value

This function returns a data frame containing the following columns:

SEPA_HYDROLOGY_OFFICE

STATION_NAME

LOCATION_CODE Station id number.

NATIONAL_GRID_REFERENCE

CATCHMENT_NAME

RIVER_NAME

GAUGE_DATUM

```
CATCHMENT_AREA in Km2
START_DATE
END_DATE
SYSTEM_ID
LOWEST_VALUE
LOW
MAX_VALUE
HIGH
MAX_DISPLAY
MEAN
UNITS
WEB_MESSAGE
NRFA_LINK
```

Author(s)

Claudia Vitolo

Examples

```
## Not run:
  # Retrieve the whole catalogue
  SEPA_catalogue_all <- catalogueSEPA()

## End(Not run)
```

---

grdcLTMMD   *Data set: The grdcLTMMD look-up table*

Description

The grdcLTMMD look-up table

Usage

```
data("grdcLTMMD")
```

Format

A data frame with 6 rows and 4 columns.

WMO_Region an integer between 1 and 6

Coverage

Number_of_stations

Archive url to spreadsheet

Source

<http://www.bafg.de/GRDC>

---

hddtools   *hddtools: Hydrological Data Discovery Tools*

Description

Many governmental bodies and institutions are currently committed to publish open data as the result of a trend of increasing transparency, based on which a wide variety of information produced at public expense is now becoming open and freely available to improve public involvement in the process of decision and policy making. Discovery, access and retrieval of information is, however, not always a simple task. Especially when access to data APIs is not allowed, downloading a metadata catalogue, selecting the information needed, requesting datasets, de-compression, conversion, manual filtering and parsing can become rather tedious. The R package hddtools is an open source project, designed to make all the above operations more efficient by means of reusable functions.

The package facilitate access to various online data sources such as:

- **KGClimateClass** (<http://koeppen-geiger.vu-wien.ac.at/>): The Koppen Climate Classification map is used for classifying the world's climates based on the annual and monthly averages of temperature and precipitation

- **GRDC** (<http://www.bafg.de/GRDC/EN/Home/homepage_node.html>): The Global Runoff Data Centre (GRDC) provides datasets for all the major rivers in the world

- **Data60UK** (<http://tdwg.catchment.org/datasets.html>): The Data60UK initiative collated datasets of areal precipitation and streamflow discharge across 61 gauging sites in England and Wales (UK).

- **MOPEX** (<https://www.nws.noaa.gov/ohd/mopex/mo_datasets.htm>): This dataset contains historical hydrometeorological data and river basin characteristics for hundreds of river basins in the US.

- **SEPA** (<http://apps.sepa.org.uk/waterlevels/>): The Scottish Environment Protection Agency (SEPA) provides river level data for hundreds of gauging stations in the UK.

This package complements R's growing functionality in environmental web technologies by bridging the gap between data providers and data consumers. It is designed to be an initial building block of scientific workflows for linking data and models in a seamless fashion.

References

Vitolo C, Buytaert W, 2014, HDDTOOLS: an R package serving Hydrological Data Discovery Tools, AGU Fall Meeting, 15-19 December 2014, San Francisco, USA.

KGClimateClass | *Function to identify the updated Koppen-Greiger climate zone (on a 0.1 x 0.1 degrees resolution map).*

Description

Given a bounding box, the function identifies the overlapping climate zones.

Usage

```
KGClimateClass(areaBox = NULL, updatedBy = "Peel", verbose = FALSE)
```

Arguments

areaBox   bounding box, a list made of 4 elements: minimum longitude (lonMin), mini-
              mum latitude (latMin), maximum longitude (lonMax), maximum latitude (lat-
              Max)

updatedBy   this can either be "Kottek" or "Peel"

verbose   if TRUE more info are printed on the screen

Value

List of overlapping climate zones.

Author(s)

Claudia Vitolo

References

Kottek et al. (2006): <http://koeppen-geiger.vu-wien.ac.at/>. Peel et al. (2007): http://people.eng.unimelb.edu.au/mpeel/koppen.html.

Examples

```
## Not run:
  # Define a bounding box
  areaBox <- raster::extent(-3.82, -3.63, 52.41, 52.52)
  # Get climate classes
  KGClimateClass(areaBox = areaBox)

## End(Not run)
```

---

tsData60UK   *Interface for the Data60UK database of Daily Time Series*

Description

This function extract the dataset containing daily rainfall and streamflow discharge at one of the Data60UK locations.

Usage

```
tsData60UK(id)
```

Arguments

id   String which identifies the station ID number

Value

The function returns a data frame containing 2 time series (as zoo objects): "P" (precipitation) and "Q" (discharge).

Author(s)

Claudia Vitolo

Examples

```
## Not run:
  Morwick <- tsData60UK(id = "22001")

## End(Not run)
```

---

tsMOPEX   *Interface for the MOPEX database of Daily Time Series*

Description

This function extract the dataset containing daily rainfall and streamflow discharge at one of the MOPEX locations.

Usage

```
tsMOPEX(id, MAP = TRUE)
```

Arguments

`id`	String for the station ID number (USGS_ID)
`MAP`	Boolean, TRUE by default. If FALSE it looks for data through all the 1861 potential MOPEX basins. If TRUE, it looks for data through the 438 MOPEX basins with MAP estimates.

Value

If MAP = FALSE, this function returns a time series of daily streamflow discharge (Q, in mm). If MAP = TRUE, this function returns a data frame containing the following columns (as zoo object):

`Date` Format is "yyyymmdd"

`P` Mean areal precipitation (mm)

`E` Climatic potential evaporation (mm, based NOAA Freewater Evaporation Atlas)

`Q` Daily streamflow discharge (mm)

`T_max` Daily maximum air temperature (Celsius)

`T_min` Daily minimum air temperature (Celsius)

Author(s)

Claudia Vitolo

Examples

```
## Not run:
  BroadRiver <- tsMOPEX(id = "01048000")

## End(Not run)
```

---

tsSEPA	*Interface for the MOPEX database of Daily Time Series*

Description

This function extract the dataset containing daily rainfall and streamflow discharge at one of the MOPEX locations.

Usage

```
tsSEPA(id)
```

Arguments

`id`	hydrometric reference number (string)

Value

The function returns river level data in metres, as a zoo object.

Author(s)

Claudia Vitolo

Examples

```
## Not run:
  sampleTS <- tsSEPA(id = "10048")

## End(Not run)
```

Index