Package 'dataRetrieval'

March 31, 2026

Type Package

Title Retrieval Functions for USGS and EPA Hydrology and Water Quality Data

Version 2.7.24

Description Collection of functions to help retrieve U.S. Geological Survey and U.S. Environmental Protection Agency water quality and hydrology data from web services.

License CC0

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.

Depends R (>= 4.1.0)

Imports curl (>= 7.0.0), lubridate (>= 1.5.0), stats, utils, xml2, readr (>= 1.4.0), jsonlite, httr2, whisker, sf, data.table, rlang

Suggests covr, dplyr, knitr, rmarkdown, testthat

Encoding UTF-8

BuildVignettes true

VignetteBuilder knitr

BugReports https://github.com/DOI-USGS/dataRetrieval/issues

RoxygenNote 7.3.3

NeedsCompilation no

Author Laura DeCicco [aut, cre] (ORCID: <https://orcid.org/0000-0002-3915-9487>), Robert Hirsch [aut] (ORCID: <https://orcid.org/0000-0002-4534-075X>), David Lorenz [aut], Jordan Read [ctb], Jordan Walker [ctb], Lindsay Platt [ctb], David Watkins [aut] (ORCID: <https://orcid.org/0000-0002-7544-0700>),

David Blodgett [aut] (ORCID: <https://orcid.org/0000-0001-9489-1710>),
Mike Johnson [aut] (ORCID: <https://orcid.org/0000-0002-5288-8350>),
Aliesha Krall [ctb] (ORCID: <https://orcid.org/0000-0003-2521-5043>),
Lee Stanish [ctb] (ORCID: <https://orcid.org/0000-0002-9775-6861>),
Joseph Zemmels [aut] (ORCID: <https://orcid.org/0009-0008-1463-6313>),
Elise Hinman [aut] (ORCID: <https://orcid.org/0000-0001-5396-1583>),
Michael Mahoney [ctb] (ORCID: <https://orcid.org/0000-0003-2402-304X>)

Maintainer Laura DeCicco <ldecicco@usgs.gov>

Repository CRAN

Date/Publication 2026-03-31 14:40:10 UTC

Contents

`addWaterYear` *add a water year column*

Description

Add a column to the dataRetrieval data frame with the water year. WQP queries will return a water year column for the start and end dates of the data.

Usage

```
addWaterYear(rawData)
```

Arguments

rawData	the daily- or unit-values datset retrieved from NWISweb. Must have at least one of the following columns to add the new water year columns: dateTime, Date, ActivityStartDate, or ActivityEndDate. The date column(s) can be character, POSIXct, Date. They cannot be numeric.

Value

data.frame with an additional integer column with "WY" appended to the date column name. For WQP, there will be 2 columns: ActivityStartDateWY and ActivityEndDateWY.

Examples

```
nwisData <- readNWISdv("04085427", "00060", "2022-01-01", "2022-06-30")
nwisData <- addWaterYear(nwisData)

wqpData <- readWQPqw("USGS-01594440", "01075", "", "")
wqpData <- addWaterYear(wqpData)
```

---

calcWaterYear *Extract WY from a date*

Description

Determine the correct water year based on a calendar date.

Usage

```
calcWaterYear(dateVec)
```

Arguments

dateVec	vector of dates as character ("YYYY-DD-MM"), Date, or POSIXct. Numeric does not work.

Details

This function calculates a water year based on the USGS definition that a water year starts on October 1 of the year before, and ends on September 30. For example, water year 2015 started on 2014-10-01 and ended on 2015-09-30.

Value

numeric vector indicating the water year

Examples

```
x <- seq(as.Date("2010-01-01"), as.Date("2010-12-31"), by = "month")
calcWaterYear(x)

y <- c("2010-01-01", "1994-02", "1980", "2009-11-01", NA)
calcWaterYear(y)
```

---

checkWQPdates *Date Check for Water Quality Portal*

Description

Checks date format for inputs to the Water Quality Portal. Used in `readWQPqw` and `readWQPdata`.

Usage

```
checkWQPdates(values)
```

Arguments

values	named list with arguments to send to the Water Quality Portal

Value

values named list with corrected arguments to send to the Water Quality Portal

Examples

```
values <- list(
  startDateLo = "01-01-2002",
  characteristicName = "Phosphorous",
  endDate = as.Date("2014-01-01")
)
values <- checkWQPdates(values)
```

---

check_waterdata_sample_params
Check values from codeservice

Description

Call a service to check on values from: https://api.waterdata.usgs.gov/samples-data/codeservice/docs.

Usage

```
check_waterdata_sample_params(service = "characteristicgroup", ...)
```

Arguments

service	Options are: "characteristicgroup", "states", "counties", "countries", "sitetype", "samplemedia", "characteristics", "observedproperty"
...	Optional additional query arguments. Only "characteristics" and "observedproperty" have additional parameters options.

Value

List, structure depends on service.

Examples

```
groups <- check_waterdata_sample_params("characteristicgroup")
states <- check_waterdata_sample_params("states")
countries <- check_waterdata_sample_params("countries")
counties <- check_waterdata_sample_params("counties")
sitetypes <- check_waterdata_sample_params("sitetype")
samplemedia <- check_waterdata_sample_params("samplemedia")
characteristics <- check_waterdata_sample_params("characteristics",
                                 group = "Biological")
observedProperties <- check_waterdata_sample_params("observedproperty",
                                    text = "phosphorus")
ref_list <- check_waterdata_sample_params("reference-list")
```

---

constructNWISURL *Construct NWIS url for data retrieval*

Description

Using USGS water web services to construct urls.

Usage

```
constructNWISURL(
  siteNumbers,
  parameterCd = "00060",
  startDate = "",
  endDate = "",
  service,
  statCd = "00003",
  format = "xml",
  expanded = TRUE,
  ratingType = "base",
  statReportType = "daily",
```

```
  statType = "mean"
)
```

Arguments

siteNumbers	string or vector of strings USGS site number. This is usually an 8 digit number
parameterCd	string or vector of USGS parameter code. This is usually an 5 digit number.
startDate	character starting date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the earliest possible record.
endDate	character ending date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the latest possible record.
service	string USGS service to call. Possible values are "dv" (daily values), "uv" (unit/instantaneous values), and "rating" (rating curve), "peak", "stat" (statistics web service BETA).
statCd	string or vector USGS statistic code only used for daily value service. This is usually 5 digits. Daily mean (00003) is the default.
format	string, can be "tsv" or "xml", and is only applicable for daily and unit value requests. "tsv" returns results faster, but there is a possibility that an incomplete file is returned without warning. XML is slower, but will offer a warning if the file was incomplete (for example, if there was a momentary problem with the internet connection). It is possible to safely use the "tsv" option, but the user must carefully check the results to see if the data returns matches what is expected. The default is therefore "xml".
expanded	logical defaults to TRUE. If TRUE, retrieves additional information, only applicable for qw data.
ratingType	can be "base", "corr", or "exsa". Only applies to rating curve data.
statReportType	character Only used for statistics service requests. Time division for statistics: daily, monthly, or annual. Default is daily. Note that daily provides statistics for each calendar day over the specified range of water years, i.e. no more than 366 data points will be returned for each site/parameter. Use readNWISdata or readNWISdv for daily averages. Also note that "annual" returns statistics for the calendar year. Use readNWISdata for water years. Monthly and yearly provide statistics for each month and year within the range individually.
statType	character Only used for statistics service requests. Type(s) of statistics to output for daily values. Default is mean, which is the only option for monthly and yearly report types.

Value

url string

Examples

```
site_id <- "01594440"
startDate <- "1985-01-01"
endDate <- ""
pCode <- c("00060", "00010")
```

```
url_daily <- constructNWISURL(site_id, pCode,
  startDate, endDate, "dv",
  statCd = c("00003", "00001")
)
url_unit <- constructNWISURL(site_id, pCode, "2012-06-28", "2012-06-30", "iv")

url_daily_tsv <- constructNWISURL(site_id, pCode, startDate, endDate, "dv",
  statCd = c("00003", "00001"), format = "tsv"
)
url_rating <- constructNWISURL(site_id, service = "rating", ratingType = "base")
url_peak <- constructNWISURL(site_id, service = "peak")
```

`constructWQPURL` *Construct WQP url for data retrieval*

Description

Construct WQP url for data retrieval. This function gets the data from here: https://www.waterqualitydata.us

Usage

```
constructWQPURL(siteNumbers, parameterCd, startDate, endDate, legacy = TRUE)
```

Arguments

`siteNumbers`	string or vector of strings USGS site number.
`parameterCd`	string or vector of USGS parameter code. This is usually an 5 digit number.
`startDate`	character starting date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the earliest possible record.
`endDate`	character ending date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the latest possible record.
`legacy`	Logical. If TRUE, uses legacy WQP services. Default is TRUE. Setting legacy = FALSE uses WQX3.0 WQP services, which are in-development, use with caution.

Value

url string

Examples

```
site_ids <- c("USGS-02292010", "USGS-02276877")
startDate <- "2020-01-01"
endDate <- ""
pCode <- c("80154", "00613")
url_wqp <- constructWQPURL(
  site_ids,
```

```
  pCode,
  startDate, endDate
)
url_wqp
charNames <- c(
  "Temperature",
  "Temperature, sample",
  "Temperature, water",
  "Temperature, water, deg F"
)
obs_url_orig <- constructWQPURL(
  siteNumbers = c(
    "IIDFG-41WSSPAHS",
    "USGS-02352560"
  ),
  parameterCd = charNames,
  startDate, ""
)
obs_url_orig
```

---

countyCd *US County Code Lookup Table*

Description

Classic lookup table for counties. Has been replaced in functions with `check_waterdata_sample_params("counties")`.

Value

countyCd data frame.

Name	Type	Description
STUSAB	character	State abbreviation
STATE	character	two-digit ANSI code
COUNTY	character	three-digit county code
COUNTY_NAME	character	County full name
COUNTY_ID	character	County id

Examples

```
head(countyCd)
```

countyCdLookup *US county code look up*

Description

Function to simplify finding county and county code definitions. Used in `readNWISdata` and `readNWISuse`. Currently only has US counties.

Usage

```
countyCdLookup(state, county, outputType = "fips")
```

Arguments

`state`	could be character (full name, abbreviation, id), or numeric (id)
`county`	could be character (name, with or without "County") or numeric (id)
`outputType`	character can be "fullName", "tableIndex", "id", or "fullEntry".

Examples

```
fips <- countyCdLookup(state = "WI", county = "Dane")
id <- countyCdLookup(state = "WI", county = "Dane", outputType = "id")
name <- countyCdLookup(state = "OH", county = 13, output = "fullName")
entry <- countyCdLookup(state = "Pennsylvania", county = "ALLEGHENY COUNTY", output = "fullEntry")
fromIDs <- countyCdLookup(state = 13, county = 5, output = "fullName")
```

create_NWIS_bib *Create NWIS data citation*

Description

Uses attributes from the NWIS functions to create data citations.

Usage

```
create_NWIS_bib(x)
```

Arguments

`x`	Any data returned from an NWIS function, must include "queryTime" and "url" attributes, which should come with the data by default.

Details

See ?bibentry for more information.

Value

bibentry object to use for citing the data.

Examples

```
nwisData <- read_waterdata_daily(monitoring_location_id = "USGS-04085427",
                                 parameter_code = "00060",
                                 time = c("2012-01-01", "2012-06-30"))
nwis_citation <- create_NWIS_bib(nwisData)
nwis_citation

print(nwis_citation, style = "Bibtex")
print(nwis_citation, style = "citation")
```

---

create_WQP_bib *Create WQP data citation*

Description

Uses attributes from the WQP functions to create data citations.

Usage

```
create_WQP_bib(x)
```

Arguments

x	Any data returned from an NWIS function, must include "queryTime" and "url" attributes, which should come with the data by default.

Details

See ?bibentry for more information.

Value

bibentry object to use for citing the data.

Examples

```
WQPData <- readWQPqw("USGS-05288705",
                     parameterCd = "00300")
wqp_citation <- create_WQP_bib(WQPData)
wqp_citation

print(wqp_citation, style = "Bibtex")
print(wqp_citation, style = "citation")
```

---

findNLDI *R Client for the Network Linked Data Index*

Description

Provides a formal client to the USGS Network Linked Data Index.

Usage

```
findNLDI(
  comid = NULL,
  nwis = NULL,
  wqp = NULL,
  huc12 = NULL,
  location = NULL,
  origin = NULL,
  nav = NULL,
  find = c("flowlines"),
  distance_km = 100,
  no_sf = FALSE,
  warn = TRUE
)
```

Arguments

comid	numeric or character. An NHDPlusV2 COMID
nwis	numeric or character. A USGS NWIS surface water siteID
wqp	numeric or character. A water quality point ID
huc12	numeric or character. A WBD HUC12 unit ID
location	numeric vector. Coordinate pair in WGS84 SRS ordered lng/lat (X,Y)
origin	named list. Specifying a feature type and ID (e.g. list("comid" = 101))
nav	character vector. where to navigate from the starting point. Options include along the upper mainsteam (UM), upstream tributary (UT), downstream mainstem (DM) and downstream divergences (DD). You may select one or more of the abbreviations ("UM", "UT", DM", "DD").

find	character vector. Define what resources to find along the navigation path(s) (see get_nldi_sources()$source). Can also include 'basin' or 'flowline', which will return the upstream basin of the starting feature or flowlines along the navigation respectively. The default is "flowlines". If you provide any other resource, AND want flowlines, then flowlines must be explicitly requested.
distance_km	numeric. Define how far to look along the navigation path in kilometers (default = 100)
no_sf	if available, should sf be used for parsing, defaults to TRUE if sf is locally installed
warn	(default TRUE) should warnings be printed

Details

The function is useful for topology and location based feature discovery. A user must specify an origin feature, optional navigation direction(s) along the network, as well as features to identify along the navigated paths. Valid starting options can be given by one of the following arguments: comid, nwis, huc12, wqp, location, and start.

Value

a list of data.frames if sf is not installed, a list of sf objects if it is

Examples

```
# Find Features / Define origin features

## Find feature by COMID
findNLDI(comid = 101)

## Find feature by NWIS ID
findNLDI(nwis = "11120000")


## Find feature by LOCATION
findNLDI(location = c(-115, 40))

## GENERAL ORIGIN: COMID
findNLDI(origin = list("comid" = 101))

# Navigation (flowlines will be returned if find is unspecified)
# UPPER MAINSTEM of USGS-11120000
findNLDI(nwis = "11120000", nav = "UM")

# MULTI-REQUEST
# UPPER MAINSTEM and TRIBUTARY of USGS-11120000
findNLDI(nwis = "11120000", nav = c("UT", "UM"))

# Discover Features(flowlines will not be returned unless included in find)
```

```
## Find feature(s) on the upper tributary of USGS-11120000
findNLDI(nwis = "11120000", nav = "UT", find = c("nwis"))

## Find upstream basin boundary and  of USGS-11120000
findNLDI(nwis = "11120000", find = "basin")

# Control Distance
## Limit search to 50 km
findNLDI(comid = 101, nav = "DM", find = c("nwis", "flowlines"), distance_km = 50)
```

`getQuerySummary` *getting header information from a WQP query*

Description

getting header information from a WQP query

Usage

```
getQuerySummary(url)
```

Arguments

`url`	the query url

`getWebServiceData` *Function to return data from web services*

Description

This function accepts a url parameter, and returns the raw data.

Usage

```
getWebServiceData(obs_url, ...)
```

Arguments

`obs_url`	character containing the url for the retrieval
`...`	information to pass to header request

Details

To add a custom user agent, create an environmental variable: CUSTOM_DR_UA

Value

Returns xml, json, or text depending on the requested data.

Examples

```
siteNumber <- "02177000"
startDate <- "2012-09-01"
endDate <- "2012-10-01"
offering <- "00003"
property <- "00060"
obs_url <- constructNWISURL(siteNumber, property, startDate, endDate, "dv")

rawData <- getWebServiceData(obs_url)
```

---

get_nldi_sources *Get current NLDI offerings*

Description

Used to query the current resources available through the NLDI

Usage

```
get_nldi_sources(url = pkg.env$nldi_base)
```

Arguments

url	URL for NLDI sources. Default is supplied by package environment.

Value

data.frame

Examples

```
get_nldi_sources()
```

`get_ogc_params` *Get parameter descriptions*

Description

This function returns a list of properties available from an endpoint. When available, it will also contain a description.

Usage

```
get_ogc_params(service)
```

Arguments

`service`	Character, can be any of the USGS Waterdata API endpoints or collections.

Value

list

Examples

```
ml <- dataRetrieval::get_ogc_params("monitoring-locations")
ml$national_aquifer_code
```

`importNGWMN` *Function to return data from the National Ground Water Monitoring Network waterML2 format*

Description

This function accepts a url parameter for a WaterML2 getObservation. This function is still under development, but the general functionality is correct.

Usage

```
importNGWMN(input, asDateTime = FALSE, tz = "UTC")
```

Arguments

input	character or raw, containing the url for the retrieval or a path to the data file, or raw XML.
asDateTime	logical, if TRUE returns date and time as POSIXct, if FALSE, character
tz	character to set timezone attribute of dateTime. Default is "UTC", and converts the date times to UTC, properly accounting for daylight savings times based on the data's provided time zone offset. Possible values to provide are "America/New_York", "America/Chicago", "America/Denver", "America/Los_Angeles", "America/Anchorage", as well as the following which do not use daylight savings time: "America/Honolulu", "America/Jamaica", "America/Managua", "America/Phoenix", and "America/Metlakatla". See also OlsonNames() for more information on time zones.

Value

mergedDF a data frame source, time, value, uom, uomTitle, comment, gmlID

Examples

```
params <- list(request = "GetObservation",
               service = "SOS",
               version = "2.0.0",
               observedProperty = "urn:ogc:def:property:OGC:GroundWaterLevel",
               responseFormat = "text/xml",
               featureOfInterest = "VW_GWDP_GEOSERVER.USGS.403836085374401")

obs_url <- httr2::request("https://cida.usgs.gov") |>
 httr2::req_url_path_append("ngwmn_cache") |>
 httr2::req_url_path_append("sos") |>
 httr2::req_url_query(!!!params)

#data_returned <- importNGWMN(obs_url)
```

---

importRDB1 *Function to return data from the NWIS RDB 1.0 format*

Description

This function accepts a url parameter that already contains the desired NWIS site, parameter code, statistic, startdate and enddate. It is not recommended to use the RDB format for importing multisite data.

Usage

```
importRDB1(obs_url, asDateTime = TRUE, convertType = TRUE, tz = "UTC")
```

Arguments

obs_url	character containing the url for the retrieval or a file path to the data file.
asDateTime	logical, if TRUE returns date and time as POSIXct, if FALSE, Date
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates, datetimes, numerics based on a standard algorithm. If false, everything is returned as a character
tz	character to set timezone attribute of datetime. Default converts the datetimes to UTC (properly accounting for daylight savings times based on the data's provided tz_cd column). Recommended US values include "UTC", "America/New_York", "America/Chicago", "America/Denver", "America/Los_Angeles", "America/Anchorage", "America/Honolulu", "America/Jamaica", "America/Managua", "America/Phoenix", and "America/Metlakatla". For a complete list, see https://en.wikipedia.org/wiki/List_of_tz_database_time_zones

Value

A data frame with the following columns:

Name	Type	Description
agency_cd	character	The NWIS code for the agency reporting the data
site_no	character	The USGS site number
datetime	POSIXct	The date and time of the value converted to UTC (if asDateTime = TRUE
	character	or raw character string (if asDateTime = FALSE)
tz_cd	character	The time zone code for datetime
code	character	Any codes that qualify the corresponding value
value	numeric	The numeric value for the parameter
tz_cd_reported	The originally reported time zone	

Note that code and value are repeated for the parameters requested. The names are of the form XD_P_S, where X is literal, D is an option description of the parameter, P is the parameter code, and S is the statistic code (if applicable). If a date/time (dt) column contained incomplete date and times, a new column of dates and time was inserted. This could happen when older data was reported as dates, and newer data was reported as a date/time.

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
queryTime	POSIXct	The time the data was returned
comment	character	Header comments from the RDB file

Examples

```
site_id <- "02177000"
startDate <- "2012-09-01"
endDate <- "2012-10-01"
offering <- "00003"
property <- "00060"

obs_url <- constructNWISURL(site_id, property,
  startDate, endDate, "dv",
  format = "tsv"
)

data <- importRDB1(obs_url)


urlMultiPcodes <- constructNWISURL("04085427", c("00060", "00010"),
  startDate, endDate, "dv",
  statCd = c("00003", "00001"), "tsv"
)

multiData <- importRDB1(urlMultiPcodes)

unitDataURL <- constructNWISURL(site_id, property,
  "2020-10-30", "2020-11-01", "uv",
  format = "tsv"
) # includes timezone switch

unitData <- importRDB1(unitDataURL, asDateTime = TRUE)

iceSite <- "04024000"
start <- "2015-11-09"
end <- "2015-11-24"
urlIce <- constructNWISURL(iceSite, "00060", start, end, "uv", format = "tsv")

ice <- importRDB1(urlIce, asDateTime = TRUE)
iceNoConvert <- importRDB1(urlIce, convertType = FALSE)

# User file:
filePath <- system.file("extdata", package = "dataRetrieval")
fileName <- "RDB1Example.txt"
fullPath <- file.path(filePath, fileName)
importUserRDB <- importRDB1(fullPath)
```

`importWaterML1` *Function to return data from the NWISWeb WaterML1.1 service*

Description

This function accepts a url parameter that already contains the desired NWIS site, parameter code, statistic, startdate and enddate.

Usage

```
importWaterML1(obs_url, asDateTime = FALSE, tz = "UTC")
```

Arguments

obs_url	character or raw, containing the url for the retrieval or a file path to the data file, or raw XML.
asDateTime	logical, if TRUE returns date and time as POSIXct, if FALSE, Date
tz	character to set timezone attribute of datetime. Default converts the datetimes to UTC (properly accounting for daylight savings times based on the data's provided tz_cd column). Recommended US values include "UTC", "America/New_York", "America/Chicago", "America/Denver", "America/Los_Angeles", "America/Anchorage", "America/Honolulu", "America/Jamaica", "America/Managua", "America/Phoenix", and "America/Metlakatla". For a complete list, see https://en.wikipedia.org/wiki/List_of_tz_database_time_zones

Value

A data frame with the following columns:

Name	Type	Description
agency_cd	character	The NWIS code for the agency reporting the data
site_no	character	The USGS site number
	POSIXct	The date and time of the value converted to UTC (if asDateTime = TRUE),
	character	or raw character string (if asDateTime = FALSE)
tz_cd	character	The time zone code for
code	character	Any codes that qualify the corresponding value
value	numeric	The numeric value for the parameter

Note that code and value are repeated for the parameters requested. The names are of the form X_D_P_S, where X is literal, D is an option description of the parameter, P is the parameter code, and S is the statistic code (if applicable).

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
siteInfo	data.frame	A data frame containing information on the requested sites
variableInfo	data.frame	A data frame containing information on the requested parameters
statisticInfo	data.frame	A data frame containing information on the requested statistics on the data
queryTime	POSIXct	The time the data was returned

See Also

renameNWISColumns()

Examples

```
site_id <- "02177000"
startDate <- "2012-09-01"
endDate <- "2012-10-01"
offering <- "00003"
property <- "00060"
obs_url <- constructNWISURL(site_id, property, startDate, endDate, "dv")

data <- importWaterML1(obs_url, asDateTime = TRUE)


unitDataURL <- constructNWISURL(
  site_id, property,
  "2013-11-03", "2013-11-03", "uv"
)
unitData <- importWaterML1(unitDataURL, TRUE)

# Two sites, two pcodes, one site has two data descriptors:
siteNumber <- c("01480015", "04085427")
obs_url <- constructNWISURL(
  siteNumber, c("00060", "00010"),
  startDate, endDate, "dv"
)
data <- importWaterML1(obs_url)
data$dateTime <- as.Date(data$dateTime)
data <- renameNWISColumns(data)
names(attributes(data))
attr(data, "url")
attr(data, "disclaimer")

inactiveSite <- "05212700"
inactiveSite <- constructNWISURL(inactiveSite, "00060",
                                 "2014-01-01", "2014-01-10", "dv")
inactiveSite <- importWaterML1(inactiveSite)

inactiveAndAcitive <- c("07334200", "05212700")
inactiveAndAcitive <- constructNWISURL(inactiveAndAcitive,
                         "00060", "2014-01-01", "2014-01-10", "dv")
inactiveAndAcitive <- importWaterML1(inactiveAndAcitive)

# Timezone change with specified local timezone:
tzURL <- constructNWISURL("04027000", c("00300", "63680"),
                          "2011-11-05", "2011-11-07", "uv")
tzIssue <- importWaterML1(tzURL,
  asDateTime = TRUE, tz = "America/Chicago"
)


filePath <- system.file("extdata", package = "dataRetrieval")
fileName <- "WaterML1Example.xml"
fullPath <- file.path(filePath, fileName)
importFile <- importWaterML1(fullPath, TRUE)
```

importWaterML2 *Parse the WaterML2 timeseries portion of a waterML2 file*

Description

Returns data frame columns of all information with each time series measurement; Anything defined as a default, is returned as an attribute of that data frame.

Usage

```
importWaterML2(input, asDateTime = FALSE, tz = "UTC")
```

Arguments

input	XML with only the wml2:MeasurementTimeseries node and children
asDateTime	logical, if TRUE returns date and time as POSIXct, if FALSE, character
tz	character to set timezone attribute of datetime. Default is an empty quote, which converts the datetimes to UTC (properly accounting for daylight savings times based on the data's provided time zone offset). Possible values are "America/New_York", "America/Chicago", "America/Denver", "America/Los_Angeles", "America/Anchorage", "America/Honolulu", "America/Jamaica", "America/Managua", "America/Phoenix", and "America/Metlakatla"

Examples

```
baseURL <- httr2::request("https://waterservices.usgs.gov/nwis/dv")
baseURL <- httr2::req_url_query(baseURL,
                                format = "waterml,2.0",
                                sites = "01646500",
                                startDT = "2014-09-01",
                                endDT = "2014-09-08",
                                statCd = "00003",
                                parameterCd = "00060" )

timeseries <- importWaterML2(baseURL, asDateTime = TRUE, tz = "UTC")
```

importWQP *Basic Water Quality Portal Data parser*

Description

Imports data from the Water Quality Portal based on a specified url.

Usage

```
importWQP(obs_url, tz = "UTC", csv = TRUE, convertType = TRUE)
```

Arguments

obs_url	character URL to Water Quality Portal#' @keywords data import USGS web service
tz	character to set timezone attribute of datetime. Default is UTC (properly accounting for daylight savings times based on the data's provided tz_cd column). Possible values include "America/New_York","America/Chicago", "America/Denver","America/Los_An "America/Anchorage","America/Honolulu","America/Jamaica","America/Managua", "America/Phoenix", and "America/Metlakatla"
csv	logical. Is the data coming back with a csv or tsv format. Default is FALSE. Currently, the summary service does not support tsv, for other services tsv is the safer choice.
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates, datetimes, numerics based on a standard algorithm. If false, everything is returned as a character.

Value

retval dataframe raw data returned from the Water Quality Portal. Additionally, a POSIXct dateTime column is supplied for start and end times, and converted to UTC. See https://www.waterqualitydata.us/portal_userguide/ for more information.

See Also

readWQPdata(), readWQPqw(), whatWQPsites()

Examples

```
# These examples require an internet connection to run

## Examples take longer than 5 seconds:

rawSampleURL <- constructWQPURL("USGS-01594440", "01075", "", "")

rawSample <- importWQP(rawSampleURL)
```

```
STORETex <- constructWQPURL("WIDNR_WQX-10032762", "Specific conductance", "", "")

STORETdata <- importWQP(STORETex)

STORETdata_char <- importWQP(STORETex, convertType = FALSE)
```

---

is_dataRetrieval_user *Is this a dataRetrieval user*

Description

Reveals if this is a user or not

Usage

```
is_dataRetrieval_user()
```

Examples

```
is_dataRetrieval_user()
```

---

parameterCdFile *List of USGS parameter codes*

Description

Complete list of USGS parameter codes as of Oct. 24, 2024.

Value

parameterData data frame with information about USGS parameters.

Name	Type	Description
parameter_cd	character	5-digit USGS parameter code
parameter_group_nm	character	USGS parameter group name
parameter_nm	character	USGS parameter name
casrn	character	Chemical Abstracts Service (CAS) Registry Number
srsname	character	Substance Registry Services Name
parameter_units	character	Parameter units

Examples

```
head(parameterCdFile[, 1:2])
```

parse_WQP *Convert WQP columns to correct types*

Description

Takes the character results and converts to numeric and dates.

Usage

```
parse_WQP(retval, tz = "UTC")
```

Arguments

retval	Data frame from WQP
tz	character to set timezone attribute of datetime. Default is UTC (properly accounting for daylight savings times based on the associated "TimeZone" column). Possible values include "America/New_York","America/Chicago", "America/Denver","America/Los_Angeles", "America/Anchorage","America/Honolulu","America/Jamaica","A "America/Phoenix", and "America/Metlakatla"

Value

data frame retval with converted columns

Examples

```
# These examples require an internet connection to run
rawSampleURL <- constructWQPURL("USGS-01594440", "01075", "", "")

## Examples take longer than 5 seconds:


rawSample <- importWQP(rawSampleURL, convertType = FALSE)
convertedSample <- parse_WQP(rawSample)
```

pcode_to_name *Parameter code to characteristic name*

Description

This function is useful to fine what characteristic name, result sample fraction, unit code, and other parameters are mapped with USGS parameter codes. This information is useful for converting workflows from a more traditional NWIS water quality retrieval to a Water Quality Portal retrieval.

Usage

```
pcode_to_name(parameterCd = "all")
```

Arguments

parameterCd	character that contains the code for a character vector of 5-digit parameter codes. Default is "all" which will return a complete list of parameter codes that have been mapped to a characteristic name.

Value

a data frame with columns "parm_cd", "description", "characteristicname", "measureunitcode", "resultsamplefraction", "resulttemperaturebasis", "resultstatisticalbasis", "resulttimebasis", "resultweightbasis", "resultparticlesizebasis", "last_rev_dt"

Examples

```
pcodes <- c("00070", "00075", "00430", "52642")


all <- pcode_to_name()
some <- pcode_to_name(pcodes)
```

readNGWMNdata *Import data from the National Groundwater Monitoring Network.*

Description

Only water level data and site locations and names are currently available through the web service.

Usage

```
readNGWMNdata(service, ..., asDateTime = TRUE, tz = "UTC")
```

Arguments

service	char Service for the request - "observation" and "featureOfInterest" are implemented.
...	Other parameters to supply, namely siteNumbers or bbox
asDateTime	logical if TRUE, will convert times to POSIXct format. Currently defaults to FALSE since time zone information is not included.
tz	character to set timezone attribute of dateTime. Default is "UTC", and converts the date times to UTC, properly accounting for daylight savings times based on the data's provided time zone offset. Possible values to provide are "America/New_York", "America/Chicago", "America/Denver", "America/Los_Angeles", "America/Anchorage", as well as the following which do not use daylight savings time: "America/Honolulu", "America/Jamaica", "America/Managua", "America/Phoenix", and "America/Metlakatla". See also OlsonNames() for more information on time zones.

Examples

```
# one site
site <- "USGS.430427089284901"
#oneSite <- readNGWMNdata(siteNumbers = site, service = "observation")

# multiple sites
sites <- c("USGS.272838082142201", "USGS.404159100494601", "USGS.401216080362703")
# Very slow:
# multiSiteData <- readNGWMNdata(siteNumbers = sites, service = "observation")
# attributes(multiSiteData)

# non-USGS site
# accepts colon or period between agency and ID
site <- "MBMG:702934"
# data <- readNGWMNdata(siteNumbers = site, service = "featureOfInterest")

# bounding box
# bboxSites <- readNGWMNdata(service = "featureOfInterest", bbox = c(30, -102, 31, 99))
# retrieve  sites.  Set asDateTime to false since one site has an invalid date
# Very slow:
# bboxData <- readNGWMNdata(service = "observation", siteNumbers = bboxSites$site[1:3],
#                           asDateTime = FALSE)
```

---

readNGWMNlevels *Retrieve groundwater levels from the National Ground Water Monitoring Network.*

Description

Retrieve groundwater levels from the National Ground Water Monitoring Network.

Usage

```
readNGWMNlevels(siteNumbers, asDateTime = TRUE, tz = "UTC")
```

Arguments

siteNumbers	character Vector of feature IDs formatted with agency code and site number separated by a period or semicolon, e.g. USGS.404159100494601.
asDateTime	logical Should dates and times be converted to date/time objects, or returned as character? Defaults to TRUE. Must be set to FALSE if a site contains non-standard dates.
tz	character to set timezone attribute of dateTime. Default is "UTC", and converts the date times to UTC, properly accounting for daylight savings times based on the data's provided time zone offset. Possible values to provide are "America/New_York", "America/Chicago", "America/Denver", "America/Los_Angeles", "America/Anchorage", as well as the following which do not use daylight savings time: "America/Honolulu", "America/Jamaica", "America/Managua", "America/Phoenix", and "America/Metlakatla". See also OlsonNames() for more information on time zones.

Examples

```
# one site
site <- "USGS.430427089284901"
# oneSite <- readNGWMNlevels(siteNumbers = site)

# multiple sites
sites <- c("USGS:272838082142201", "USGS:404159100494601", "USGS:401216080362703")
# multiSiteData <- readNGWMNlevels(sites)

# non-USGS site
site <- "MBMG.103306"
# data <- readNGWMNlevels(siteNumbers = site, asDateTime = FALSE)

# site with no data returns empty data frame
noDataSite <- "UTGS.401544112060301"
# noDataSite <- readNGWMNlevels(siteNumbers = noDataSite)
```

`readNGWMNsites` *Retrieve site data from the National Ground Water Monitoring Network.*

Description

Retrieve site data from the National Ground Water Monitoring Network.

Usage

```
readNGWMNsites(siteNumbers)
```

Arguments

`siteNumbers`	character Vector of feature IDs formatted with agency code and site number separated by a period or semicolon, e.g. `USGS.404159100494601`.

Value

A data frame the following columns: #'

Name	Type	Description
site	char	Site FID
description	char	Site description
dec_lat_va, dec_lon_va	numeric	Site latitude and longitude

Examples

```
# one site
site <- "USGS.430427089284901"
#oneSite <- readNGWMNsites(siteNumbers = site)

# non-USGS site
site <- "MBMG.103306"
#siteInfo <- readNGWMNsites(siteNumbers = site)
```

readNWISdata *General Data Import from NWIS*

Description

Returns data from the NWIS web service. Arguments to the function should be based on https://waterservices.usgs.gov service calls. See examples below for ideas of constructing queries.

Usage

```
readNWISdata(..., asDateTime = TRUE, convertType = TRUE, tz = "UTC")
```

Arguments

...	see https://waterservices.usgs.gov/docs/site-service/ for a complete list of options. A list of arguments can also be supplied. One important argument to include is "service". Possible values are "iv" (for instantaneous), "dv" (for daily values), "site" (for site service), and "stat" (for statistics service). The statistics service has a limited selection of arguments (see https://waterservices.usgs.gov/docs/site-service/).
asDateTime	logical, if TRUE returns date and time as POSIXct, if FALSE, Date
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates, datetimes, numerics based on a standard algorithm. If false, everything is returned as a character
tz	character to set timezone attribute of dateTime. Default is "UTC", and converts the date times to UTC, properly accounting for daylight savings times based on the data's provided tz_cd column. Possible values to provide are "America/New_York", "America/Chicago", "America/Denver", "America/Los_Angeles", "America/Anchorage", as well as the following which do not use daylight savings time: "America/Honolulu", "America/Jamaica", "America/Managua", "America/Phoenix", and "America/Metlakatla". See also OlsonNames() for more information on time zones.

Details

This function requires users to create their own arguments based on the NWIS web services. It is a more complicated function to use compared to other NWIS functions such as readNWISdv(), readNWISuv(), etc. However, this function adds a lot of flexibility to the possible queries. This function will also behave exactly as NWIS when it comes to date queries. NWIS by default will only return the latest value for the daily and instantaneous services. So if you do not provide a starting date, you will only get back the latest value. If you want the full period of record, you can use "startDate = '1900-01-01'". Other options for dates are periods, such as "period = 'P7D'" which translates to a period of 7 days. For period, use only a positive ISO-8601 duration format, which should not be expressed in periods of less than a day, or in increments of months M or years Y. period returns data for a site generally from now to a time in the past. Note that when period is used all data up to the most recent value are returned.

Value

A data frame with the following columns:

Name	Type	Description
agency	character	The NWIS code for the agency reporting the data
site	character	The USGS site number
dateTime	POSIXct	The date and time (if applicable) of the measurement, converted to UTC for unit value data. R only al
tz_cd	character	The time zone code for dateTime column
code	character	Any codes that qualify the corresponding value
value	numeric	The numeric value for the parameter

Note that code and value are repeated for the parameters requested. The names are of the form X_D_P_S, where X is literal, D is an option description of the parameter, P is the parameter code, and S is the statistic code (if applicable).

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
siteInfo	data.frame	A data frame containing information on the requested sites
variableInfo	data.frame	A data frame containing information on the requested parameters
statisticInfo	data.frame	A data frame containing information on the requested statistics on the data
queryTime	POSIXct	The time the data was returned

See Also

read_waterdata()

Examples

```
# Examples not run for time considerations

instFlow <- readNWISdata(
  sites = "05114000", service = "iv",
  parameterCd = "00060",
  startDate = "2014-05-01T00:00Z", endDate = "2014-05-01T12:00Z"
)

instFlowCDT <- readNWISdata(
  sites = "05114000", service = "iv",
  parameterCd = "00060",
  startDate = "2014-05-01T00:00", endDate = "2014-05-01T12:00",
  tz = "America/Chicago"
)

multiSite <- readNWISdata(
  sites = c("04025500", "040263491"),
```

```
  service = "iv", parameterCd = "00060"
)

waterYearStat <- readNWISdata(
  site = c("01646500"),
  service = "stat",
  statReportType = "annual",
  statYearType = "water",
  missingData = "on"
)
monthlyStat <- readNWISdata(
  site = c("01646500"),
  service = "stat",
  statReportType = "monthly"
)

dailyStat <- readNWISdata(
  site = c("01646500"),
  service = "stat",
  statReportType = "daily",
  statType = c("p25", "p50", "p75", "min", "max"),
  parameterCd = "00060"
)

arg.list <- list(
  site = "03111548",
  statReportType = "daily",
  statType = c("p25", "p50", "p75", "min", "max"),
  parameterCd = "00060"
)
allDailyStats_2 <- readNWISdata(arg.list, service = "stat")


site_id <- "01594440"
rating_curve <- readNWISdata(service = "rating", site_no = site_id, file_type = "base")
all_sites_base <- readNWISdata(service = "rating", file_type = "base")
all_sites_core <- readNWISdata(service = "rating", file_type = "corr")
all_sites_exsa <- readNWISdata(service = "rating", file_type = "exsa")
all_sites_24hrs <- readNWISdata(service = "rating", file_type = "exsa", period = 24)

peak_data <- readNWISdata(
  service = "peak",
  site_no = c("01594440", "040851325"),
  range_selection = "data_range"
)

peak_data <- readNWISdata(
  service = "peak",
  state_cd = "PA"
)

peak_data <- readNWISdata(
  service = "peak",
```

```
  huc2_cd = "20"
)
```

`readNWISdv` *Daily Value USGS NWIS Data Retrieval*

Description

Imports data from NWIS daily web service. This function gets the data from here: https://waterservices.usgs.gov/docs/dv-service/daily-values-service-details/ Inputs to this function are just USGS site ids, USGS parameter codes, USGS statistic codes, and start and end date. For a more complex query, use `readNWISdata()`, with an argument service = "dv". Data coming the daily web services are aggregates of the instantaneous (sensor) web services. Not all statistical codes are available for all data. Use the function `whatNWISdata()` to discover what data is available for a USGS site. The column data_type_cd with the values "dv" returned from `whatNWISdata()`) are available from this service.

Usage

```
readNWISdv(
  siteNumbers,
  parameterCd,
  startDate = "",
  endDate = "",
  statCd = "00003"
)
```

Arguments

`siteNumbers`	character USGS site number. This is usually an 8 digit number. Multiple sites can be requested with a character vector.
`parameterCd`	character of USGS parameter code(s). This is usually an 5 digit number.
`startDate`	character starting date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the earliest possible record. Date arguments are always specified in local time.
`endDate`	character ending date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the latest possible record. Date arguments are always specified in local time.
`statCd`	character USGS statistic code. This is usually 5 digits. Daily mean (00003) is the default.

Details

More information on the web service can be found here: https://waterservices.usgs.gov/test-tools, choosing the "Daily Value Service".

Value

A data frame with the following columns:

Name	Type	Description
agency	character	The NWIS code for the agency reporting the data
site	character	The USGS site number
Date	Date	The date of the value
code	character	Any codes that qualify the corresponding value
value	numeric	The numeric value for the parameter

Note that code and value are repeated for the parameters requested. The names are of the form X_D_P_S, where X is literal, D is an option description of the parameter, P is the parameter code, and S is the statistic code (if applicable).

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
siteInfo	data.frame	A data frame containing information on the requested sites
variableInfo	data.frame	A data frame containing information on the requested parameters
statisticInfo	data.frame	A data frame containing information on the requested statistics on the data
queryTime	POSIXct	The time the data was returned

See Also

read_waterdata_daily()

Examples

```
# see ?read_waterdata_daily

#site_id <- "04085427"
#startDate <- "2012-01-01"
#endDate <- "2012-06-30"
#pCode <- "00060"
#
#rawDailyQ <- readNWISdv(site_id, pCode, startDate, endDate)
```

---

readNWISpCode *USGS Parameter Data Retrieval*

Description

Imports data from NWIS about measured parameter based on user-supplied parameter code or codes.

Usage

```
readNWISpCode(parameterCd)
```

Arguments

parameterCd	character of USGS parameter codes (or multiple parameter codes). These are 5 digit number codes. To get a complete list of all current parameter codes in the USGS, use "all" as the input.

Value

parameterData data frame with the following information:

Name	Type	Description
parameter_cd	character	5-digit USGS parameter code
parameter_group_nm	character	USGS parameter group name
parameter_nm	character	USGS parameter name
casrn	character	Chemical Abstracts Service (CAS) Registry Number
srsname	character	Substance Registry Services Name
parameter_units	character	Parameter units

See Also

importRDB1()

---

readNWISpeak *Peak flow data from USGS (NWIS)*

Description

Reads peak flow from NWISweb.

Usage

```
readNWISpeak(
  siteNumbers,
  startDate = "",
  endDate = "",
  asDateTime = TRUE,
  convertType = TRUE
)
```

Arguments

siteNumbers	character USGS site number(or multiple sites). This is usually an 8 digit number.
startDate	character starting date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the earliest possible record.
endDate	character ending date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the latest possible record.
asDateTime	logical default to TRUE. When TRUE, the peak_dt column is converted to a Date object, and incomplete dates are removed. When FALSE, no columns are removed, but no dates are converted.
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates, datetimes, numerics based on a standard algorithm. If false, everything is returned as a character

Details

In some cases, the specific date of the peak data is not know. This function will default to converting complete dates to a "Date" object, and converting incomplete dates to "NA". If those incomplete dates are needed, set the asDateTime argument to FALSE. No dates will be converted to R Date objects.

Value

A data frame with the following columns:

Name	Type	Description
agency_cd	character	The NWIS code for the agency reporting the data
site_no	character	The USGS site number
peak_dt	Date	Date of peak streamflow
peak_tm	character	Time of peak streamflow as character
peak_va	numeric	Annual peak streamflow value in cfs
peak_cd	character	Peak Discharge-Qualification codes (see comment for more information)
gage_ht	numeric	Gage height for the associated peak streamflow in feet
gage_ht_cd	character	Gage height qualification codes
year_last_pk	numeric	Peak streamflow reported is the highest since this year
ag_dt	Date	Date of maximum gage-height for water year (if not concurrent with peak)
ag_tm	character	Time of maximum gage-height for water year (if not concurrent with peak)
ag_gage_ht	numeric	maximum Gage height for water year in feet (if not concurrent with peak)
ag_gage_ht_cd	character	maximum Gage height code

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
queryTime	POSIXct	The time the data was returned
comment	character	Header comments from the RDB file
siteInfo	data.frame	A data frame containing information on the requested sites

See Also

constructNWISURL(), importRDB1()

Examples

```
site_ids <- c("01594440", "040851325")

data <- readNWISpeak(site_ids)
data2 <- readNWISpeak(site_ids, asDateTime = FALSE)
stations <- c("06011000")
peakdata <- readNWISpeak(stations, convertType = FALSE)
```

---

readNWISrating *Rating table for an active USGS streamgage retrieval*

Description

Reads current rating table for an active USGS streamgage from NWISweb.

Usage

```
readNWISrating(siteNumber, type = "base", convertType = TRUE)
```

Arguments

siteNumber	character USGS site number. This is usually an 8 digit number
type	character can be "base", "corr", or "exsa"
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates, datetimes, numerics based on a standard algorithm. If false, everything is returned as a character

Value

A data frame. If type is "base, " then the columns are INDEP, typically the gage height, in feet; DEP, typically the streamflow, in cubic feet per second; and STOR, where "*" indicates that the pair are a fixed point of the rating curve. If type is "exsa, " then an additional column, SHIFT, is included that indicates the current shift in the rating for that value of INDEP. If type is "corr, " then the columns are INDEP, typically the gage height, in feet; CORR, the correction for that value; and CORRINDEP, the corrected value for CORR.
If type is "base, " then the data frame has an attribute called "RATING" that describes the rating curve is included.

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
queryTime	POSIXct	The time the data was returned
comment	character	Header comments from the RDB file
siteInfo	data.frame	A data frame containing information on the requested sites
RATING	character	Rating information

Note

Not all active USGS streamgages have traditional rating curves that relate flow to stage.

See Also

`constructNWISURL()`, `importRDB1()`

Examples

```
site_id <- "01594440"

data <- readNWISrating(site_id, "base")
attr(data, "RATING")
```

---

readNWISsite *USGS Site File Data Retrieval*

Description

Imports data from USGS site file site. This function gets data from here: https://waterservices.usgs.gov/

Usage

```
readNWISsite(siteNumbers)
```

Arguments

siteNumbers	character USGS site number (or multiple sites). This is usually an 8 digit number

Value

A data frame with at least the following columns:

Name	Type	Description
agency_cd	character	The NWIS code for the agency reporting the data
site_no	character	The USGS site number

station_nm	character	Site name
site_tp_cd	character	Site type
lat_va	numeric	DMS latitude
long_va	numeric	DMS longitude
dec_lat_va	numeric	Decimal latitude
dec_long_va	numeric	Decimal longitude
coord_meth_cd	character	Latitude-longitude method
coord_acy_cd	character	Latitude-longitude accuracy
coord_datum_cd	character	Latitude-longitude datum
dec_coord_datum_cd	character	Decimal Latitude-longitude datum
district_cd	character	District code
state_cd	character	State code
county_cd	character	County code
country_cd	character	Country code
land_net_ds	character	Land net location description
map_nm	character	Name of location map
map_scale_fc	character	Scale of location map
alt_va	numeric	Altitude of Gage/land surface
alt_meth_cd	character	Method altitude determined
alt_acy_va	numeric	Altitude accuracy
alt_datum_cd	character	Altitude datum
huc_cd	character	Hydrologic unit code
basin_cd	character	Drainage basin code
topo_cd	character	Topographic setting code
instruments_cd	character	Flags for instruments at site
construction_dt	character	Date of first construction
inventory_dt	character	Date site established or inventoried
drain_area_va	numeric	Drainage area
contrib_drain_area_va	numeric	Contributing drainage area
tz_cd	character	Time Zone abbreviation
local_time_fg	character	Site honors Daylight Savings Time
reliability_cd	character	Data reliability code
gw_file_cd	character	Data-other GW files
nat_aqfr_cd	character	National aquifer code
aqfr_cd	character	Local aquifer code
aqfr_type_cd	character	Local aquifer type code
well_depth_va	numeric	Well depth
hole_depth_va	numeric	Hole depth
depth_src_cd	character	Source of depth data
project_no	character	Project number

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
queryTime	POSIXct	The time the data was returned
comment	character	Header comments from the RDB file

See Also

read_waterdata_monitoring_location()

Examples

```
# see ?read_waterdata_monitoring_location
# siteINFOMulti <- readNWISsite(c("05114000", "09423350"))
```

---

readNWISstat *Site statistics retrieval from USGS (NWIS)*

Description

Retrieves site statistics from the USGS Statistics Web Service beta.

Usage

```
readNWISstat(
  siteNumbers,
  parameterCd,
  startDate = "",
  endDate = "",
  convertType = TRUE,
  statReportType = "daily",
  statType = "mean"
)
```

Arguments

siteNumbers	character USGS site number (or multiple sites). This is usually an 8 digit number.
parameterCd	character USGS parameter code. This is usually a 5 digit number.
startDate	character starting date for data retrieval in the form YYYY, YYYY-MM, or YYYY-MM-DD. Dates cannot be more specific than the statReportType, i.e. startDate for monthly statReportTypes cannot include days, and annual statReportTypes cannot include days or months. Months and days are optional for the daily statReportType. Default is "" which indicates retrieval for the earliest possible record. For daily data, this indicates the start of the period the statistics will be computed over.
endDate	character ending date for data retrieval in the form YYYY, YYYY-MM, or YYYY-MM-DD. Default is "" which indicates retrieval for the latest possible record. For daily data, this indicates the end of the period the statistics will be computed over. The same restrictions as startDate apply.

convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to numerics based on a standard algorithm. Years, months, and days (if appliccable) are also returned as numerics in separate columns. If convertType is false, everything is returned as a character.
statReportType	character time division for statistics: daily, monthly, or annual. Default is daily. Note that daily provides statistics for each calendar day over the specified range of water years, i.e. no more than 366 data points will be returned for each site/parameter. Use readNWISdata or readNWISdv for daily averages. Also note that 'annual' returns statistics for the calendar year. Use readNWISdata for water years. Monthly and yearly provide statistics for each month and year within the range indivually.
statType	character type(s) of statistics to output for daily values. Default is mean, which is the only option for monthly and yearly report types.

Value

A data frame with the following columns:

Name	Type	Description
agency_cd	character	The NWIS code for the agency repor
site_no	character	The USGS site number
parameter_cd	character	The USGS parameter code

Other columns will be present depending on statReportType and statType

See Also

constructNWISURL(), importRDB1()

Examples

```
x1 <- readNWISstat(
  siteNumbers = c("02319394"),
  parameterCd = c("00060"),
  statReportType = "annual"
)

# all the annual mean discharge data for two sites
x2 <- readNWISstat(
  siteNumbers = c("02319394", "02171500"),
  parameterCd = c("00010", "00060"),
  statReportType = "annual"
)

# Request p25, p75, and mean values for temperature and discharge for the 2000s
# Note that p25 and p75 were not available for temperature, and return NAs
x <- readNWISstat(
  siteNumbers = c("02171500"),
  parameterCd = c("00010", "00060"),
  statReportType = "daily",
```

```
  statType = c("mean", "median"),
  startDate = "2000", endDate = "2010"
)
```

`readNWISuse` *Water use data retrieval from USGS (NWIS)*

Description

Retrieves water use data from USGS Water Use Data for the Nation.

Usage

```
readNWISuse(
  stateCd,
  countyCd,
  years = "ALL",
  categories = "ALL",
  convertType = TRUE,
  transform = FALSE
)
```

Arguments

`stateCd`	could be character (full name, abbreviation, id), or numeric (id). Only one is accepted per query.
`countyCd`	could be character (name, with or without "County", or "ALL"), numeric (id), or `NULL`, which will return state or national data depending on the stateCd argument. "ALL" may also be supplied, which will return data for every county in a state. Can be a vector of counties in the same state.
`years`	integer Years for data retrieval. Must be years ending in 0 or 5. Default is all available years.
`categories`	character categories of water use. Defaults to "ALL". Specific categories must be supplied as two- letter abbreviations as seen in the URL when using the NWIS water use web interface. Note that there are different codes for national and state level data.
`convertType`	logical defaults to `TRUE`. If `TRUE`, the function will convert the data to numerics based on a standard algorithm. Years, months, and days (if appliccable) are also returned as numerics in separate columns. If convertType is false, everything is returned as a character.
`transform`	logical only intended for use with national data. Defaults to `FALSE`, with data being returned as presented by the web service. If `TRUE`, data will be transformed and returned with column names, which will reformat national data to be similar to state data.

Value

A data frame with at least the year of record, and all available statistics for the given geographic parameters. County and state fields will be included as appropriate.

`readNWISuv` *Instantaneous value data retrieval from USGS (NWIS)*

Description

Imports data from NWIS web service. Inputs to this function are USGS site ids, USGS parameter codes, and start and end date. For a more complex query, use `readNWISdata()`, including an arguement service="uv". Not all parameter codes are available for all data. Use the function `whatNWISdata()` to discover what data is available for a USGS site. The column data_type_cd with the values "uv" returned from `whatNWISdata()`) are available from this service.

Usage

```
readNWISuv(siteNumbers, parameterCd, startDate = "", endDate = "", tz = "UTC")
```

Arguments

`siteNumbers`	character USGS site number (or multiple sites). This is usually an 8 digit number
`parameterCd`	character USGS parameter code. This is usually an 5 digit number.
`startDate`	character starting date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the earliest possible record. Simple date arguments are specified in local time.
`endDate`	character ending date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the latest possible record. Simple date arguments are specified in local time.
`tz`	character to set timezone attribute of dateTime. Default is "UTC", and converts the date times to UTC, properly accounting for daylight savings times based on the data provided tz_cd column. Possible values to provide are "America/New_York", "America/Chicago", "America/Denver", "America/Los_Angeles", "America/Anchorage", as well as the following which do not use daylight savings time: "America/Honolulu", "America/Jamaica", "America/Managua", "America/Phoenix", and "America/Metlakatla". See also `OlsonNames()` for more information on time zones.

Value

A data frame with the following columns:

Name	Type	Description
agency_cd	character	The NWIS code for the agency reporting the data
site_no	character	The USGS site number

dateTime	POSIXct	The date and time of the value converted to UTC
tz_cd	character	The time zone code for dateTime
code	character	Any codes that qualify the corresponding value
value	numeric	The numeric value for the parameter

Note that code and value are repeated for the parameters requested. The names are of the form: X_D_P_S, where X is literal, D is an option description of the parameter, P is the parameter code, and S is the statistic code (if applicable).

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
siteInfo	data.frame	A data frame containing information on the requested sites
variableInfo	data.frame	A data frame containing information on the requested parameters
statisticInfo	data.frame	A data frame containing information on the requested statistics on the data
queryTime	POSIXct	The time the data was returned

See Also

renameNWISColumns(), importWaterML1()

Examples

```
site_id <- "05114000"
parameterCd <- "00060"
startDate <- "2014-10-10"
endDate <- "2014-10-10"


rawData <- readNWISuv(site_id, parameterCd, startDate, endDate)

rawData_today <- readNWISuv(site_id, parameterCd, Sys.Date(), Sys.Date())

timeZoneChange <- readNWISuv(
  c("04024430", "04024000"), parameterCd,
  "2013-11-03", "2013-11-03"
)

centralTime <- readNWISuv(site_id, parameterCd,
  "2014-10-10T12:00", "2014-10-10T23:59",
  tz = "America/Chicago"
)

# Adding 'Z' to the time indicates to the web service to call the data with UTC time:
GMTdata <- readNWISuv(
  site_id, parameterCd,
  "2014-10-10T00:00Z", "2014-10-10T23:59Z"
)
```

readWQPdata *General Data Import from Water Quality Portal*

Description

Imports data from Water Quality Portal web service. This function gets the data from here: https://www.waterqualitydata.us.

Usage

```
readWQPdata(
  ...,
  service = "Result",
  querySummary = FALSE,
  tz = "UTC",
  ignore_attributes = FALSE,
  convertType = TRUE
)
```

Arguments

...	see https://www.waterqualitydata.us/webservices_documentation for a complete list of options. A list of arguments can also be supplied. For more information see the above description for this help file. One way to figure out how to construct a WQP query is to go to the "Advanced" form in the Water Quality Portal. Use the form to discover what parameters are available. Once the query is set in the form, scroll down to the "Query URL". You will see the parameters after "https://www.waterqualitydata.us/#". For example, if you chose "Nutrient" in the Characteristic Group dropdown, you will see characteristicType=Nutrient in the Query URL. The corresponding argument for dataRetrieval is characteristicType = "Nutrient". dataRetrieval users do not need to include mimeType, and providers is optional (these arguments are picked automatically).
service	character. See Details for more information.
querySummary	logical to only return the number of records and unique sites that will be returned from this query. Choosing TRUE is deprecated, readWQPsummary is recommended instead.
tz	character to set timezone attribute of dateTime. Default is "UTC", and converts the date times to UTC, properly accounting for daylight savings times based on the data's provided tz_cd column. Possible values to provide are "America/New_York","America/Chicago", "America/Denver","America/Los_Angeles", "America/Anchorage", as well as the following which do not use daylight savings time: "America/Honolulu", "America/Jamaica","America/Managua","America/Phoenix", and "America/Metlakatla". See also OlsonNames() for more information on time zones.

`ignore_attributes`
: logical to choose to ignore fetching site and status attributes. Default is `FALSE`.

`convertType`
: logical, defaults to `TRUE`. If `TRUE`, the function will convert the data to dates, datetimes, numerics based on a standard algorithm. If false, everything is returned as a character.

Details

This function uses ... as a query input, which can be very flexible, but also has a steeper learning curve. For a quick overview, scroll down to the Examples in this help file to see many query options.

There are currently 10 legacy options for data provided by the Water Quality Portal:

Legacy:

WQP Radio Button	service argument	Base URL
Sample Results	Result	/data/Result/search
Site Data Only	Station	/data/Station/search
Sampling Activity	Activity	/data/Activity/search
Sampling Activity Metrics	ActivityMetric	/data/ActivityMetric/search
Site Summary (not advertised on WQP)	SiteSummary	/data/summary/monitoringLocation/s
Project Data	Project	/data/Project/search
Project Monitoring Location Weighting Data	ProjectMonitoringLocationWeighting	/data/ProjectMonitoringLocationWeig
Result Detection Quantitation Limit Data	ResultDetectionQuantitationLimit	/data/ResultDetectionQuantitationLim
Biological Habitat Metrics	BiologicalMetric	/data/BiologicalMetric/search
Organization Data	Organization	/data/Organization/search

There are 4 WQX3 options. These are still in-development, and should be used with caution.

WQP Radio Button	service argument	Base URL	dataProfile
Monitoring Locations	StationWQX3	/wqx3/Station/search	
Full Physical Chemical	ResultWQX3	/wqx3/Result/search	fullPhysChem
Narrow	ResultWQX3	/wqx3/Result/search	narrow
Basic Physical Chemical	ResultWQX3	/wqx3/Result/search	basicPhysChem
Sampling Activity	ActivityWQX3	/wqx3/Activity/search	

Value

A data frame, the specific columns will depend on the "service" and/or "dataProfile".

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
siteInfo	data.frame	A data frame containing information on the requested sites
headerInfo	data.frame	A data frame returned from the WQP status service
queryTime	POSIXct	The time the data was returned

Examples

```
# Legacy:
nameToUse <- "pH"
pHData <- readWQPdata(siteid = "USGS-04024315",
                      characteristicName = nameToUse)
ncol(pHData)
attr(pHData, "siteInfo")
attr(pHData, "queryTime")
attr(pHData, "url")

# WQX3:
pHData_wqx3 <- readWQPdata(siteid = "USGS-04024315",
                           characteristicName = nameToUse,
                           service = "ResultWQX3",
                           dataProfile = "basicPhysChem")
attr(pHData_wqx3, "url")

# More examples:
# querying by county
DeWitt <- readWQPdata(
  statecode = "Illinois",
  countycode = "DeWitt",
  characteristicName = "Nitrogen"
)

attr(DeWitt, "url")

DeWitt_wqx3 <- readWQPdata(
   statecode = "Illinois",
   countycode = "DeWitt",
   characteristicName = "Nitrogen",
   service = "ResultWQX3",
   dataProfile = "basicPhysChem",
   ignore_attributes = TRUE)

attr(DeWitt_wqx3, "url")

# Data profile: "Sampling Activity"
activity <- readWQPdata(
  siteid = "USGS-04024315",
  service = "Activity"
)
attr(activity, "url")

# activity_wqx3 <- readWQPdata(
#   siteid = "USGS-04024315",
#   service = "ActivityWQX3"
# )
# attr(activity_wqx3, "url")
```

```
Dane_activity <- readWQPdata(
  statecode = "Wisconsin",
  countycode = "Dane",
  startDateLo = "2023-01-01",
  startDateHi = "2023-12-31",
  service = "Activity"
)
attr(Dane_activity, "url")

# Dane_activity_wqx3 <- readWQPdata(
#   statecode = "Wisconsin",
#   countycode = "Dane",
#   startDateLo = "2023-01-01",
#   startDateHi = "2023-12-31",
#   service = "ActivityWQX3"
# )
# attr(Dane_activity_wqx3, "url")

########################################################
# Additional examples:


# Data profiles: "Organization Data"
org_data <- readWQPdata(
  statecode = "WI",
  countycode = "Dane",
  service = "Organization"
)

# Data profiles: "Project Data"
project_data <- readWQPdata(
  statecode = "WI",
  countycode = "Dane",
  service = "Project"
)

# Data profiles: "Project Monitoring Location Weighting Data"
proj_mlwd <- readWQPdata(
  statecode = "WI",
  countycode = "Dane",
  service = "ProjectMonitoringLocationWeighting"
)

# Data profiles: "Sample Results (physical/chemical metadata)"
samp_data <- readWQPdata(
  siteid = "USGS-04024315",
  dataProfile = "resultPhysChem",
  service = "Result"
)

# Data profiles: "Sample Results (biological metadata)"
samp_bio <- readWQPdata(
  siteid = "USGS-04024315",
```

```
  dataProfile = "biological",
  service = "Result"
)


# Data profiles: "Sample Results (narrow)"
samp_narrow <- readWQPdata(
  siteid = "USGS-04024315",
  service = "Result",
  dataProfile = "narrowResult"
)

# samp_narrow_wqx3 <- readWQPdata(
#   siteid = "USGS-04024315",
#   service = "ResultWQX3",
#   dataProfile = "narrow"
# )


# Data profiles: "Sampling Activity"
samp_activity <- readWQPdata(
  siteid = "USGS-04024315",
  dataProfile = "activityAll",
  service = "Activity"
)

# Data profile: "Sampling Activity Metrics"
act_metrics <- readWQPdata(
  statecode = "WI",
  countycode = "Dane",
  service = "ActivityMetric"
)

# Data profile: "Result Detection Quantitation Limit Data"
dl_data <- readWQPdata(
  siteid = "USGS-04024315",
  service = "ResultDetectionQuantitationLimit"
)

# other options:
Phosphorus <- readWQPdata(
  statecode = "WI", countycode = "Dane",
  characteristicName = "Phosphorus",
  startDateLo = "2023-01-01",
  ignore_attributes = TRUE,
  convertType = FALSE
)

rawPHsites_legacy <- readWQPdata(siteid = c("USGS-05406450", "USGS-05427949", "WIDNR_WQX-133040"),
                         characteristicName = "pH",
                         service = "Result",
                         dataProfile = "narrowResult" )
```

```
# rawPHsites <- readWQPdata(siteid = c("USGS-05406450", "USGS-05427949", "WIDNR_WQX-133040"),
#                           characteristicName = "pH",
#                           service = "ResultWQX3",
#                           dataProfile = "narrow" )
```

---

readWQPqw *Raw Data Import for Water Quality Portal*

Description

Imports data from the Water Quality Portal. This function gets the data from here: https://www.waterqualitydata.us. There are four required input arguments: siteNumbers, parameterCd, startDate, and endDate. parameterCd can either be a USGS 5-digit code, or a characteristic name. The sites can be either USGS, or other Water Quality Portal offered sites. It is required to use the 'full' site name, such as 'USGS-01234567'.

Usage

```
readWQPqw(
  siteNumbers,
  parameterCd,
  startDate = "",
  endDate = "",
  tz = "UTC",
  legacy = TRUE,
  querySummary = FALSE,
  ignore_attributes = FALSE,
  convertType = TRUE
)
```

Arguments

siteNumbers	character site number. This needs to include the full agency code prefix.
parameterCd	vector of USGS 5-digit parameter code or characteristicNames. Leaving this blank will return all of the measured values during the specified time period.
startDate	character starting date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the earliest possible record. Date arguments are always specified in local time.
endDate	character ending date for data retrieval in the form YYYY-MM-DD. Default is "" which indicates retrieval for the latest possible record. Date arguments are always specified in local time.

tz	character to set timezone attribute of dateTime. Default is "UTC", and converts the date times to UTC, properly accounting for daylight savings times based on the data provided tz_cd column. Possible values to provide are "America/New_York","America/Chicago", "America/Denver","America/Los_Angeles", "America/Anchorage", as well as the following which do not use daylight savings time: "America/Honolulu", "America/Jamaica","America/Managua","America/Phoenix", and "America/Metlakatla". See also `OlsonNames()` for more information on time zones.
legacy	Logical. If TRUE, uses legacy WQP services. Default is TRUE. Setting legacy = FALSE uses WQX3.0 WQP services, which are in-development, use with caution.
querySummary	logical to look at number of records and unique sites that will be returned from this query.
ignore_attributes	logical to choose to ignore fetching site and parameter attributes. Default is `FALSE`.
convertType	logical, defaults to `TRUE`. If `TRUE`, the function will convert the data to dates, datetimes, numerics based on a standard algorithm. If false, everything is returned as a character.

Value

A data frame derived from the default data profile.

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
siteInfo	data.frame	A data frame containing information on the requested sites
variableInfo	data.frame	A data frame containing information on the requested parameters
queryTime	POSIXct	The time the data was returned

See Also

`readWQPdata()`, `whatWQPsites()`, and `importWQP()`

Examples

```
rawPcode <- readWQPqw("USGS-01594440", "01075", "", "")

attr(rawPcode, "siteInfo")
attr(rawPcode, "queryTime")
attr(rawPcode, "url")

rawCharacteristicName <- readWQPqw("WIDNR_WQX-10032762", "Specific conductance", "", "")
pHsites_legacy <- readWQPqw(c("USGS-05406450", "USGS-05427949", "WIDNR_WQX-133040"),
                      "pH", "", "")
ncol(pHsites_legacy)
```

```
attr(pHsites_legacy, "url")

# pHsites_modern <- readWQPqw(c("USGS-05406450", "USGS-05427949", "WIDNR_WQX-133040"),
#                          "pH", "", "", legacy = FALSE)
# ncol(pHsites_modern)
# attr(pHsites_modern, "url")

nwisEx <- readWQPqw("USGS-04024000", c("34247", "30234", "32104", "34220"), "", "2022-12-20")

DO <- readWQPqw(siteNumbers = "USGS-05288705",
                parameterCd = "00300",
                convertType = FALSE)
```

`readWQPsummary` *Summary of Data Available from Water Quality Portal*

Description

Returns a list of sites with year-by-year information on what data is available. The function gets the data from: https://www.waterqualitydata.us. Arguments to the function should be based on https://www.waterqualitydata.us/webservices_documentation. The information returned from this function describes the available data at the WQP sites, and some metadata on the sites themselves.

Usage

```
readWQPsummary(...)
```

Arguments

`...`	see https://www.waterqualitydata.us/webservices_documentation for a complete list of options. A list of arguments can also be supplied. One way to figure out how to construct a WQP query is to go to the "Advanced" form in the Water Quality Portal: https://www.waterqualitydata.us/#mimeType=csv&providers=NWIS&providers=STORET Use the form to discover what parameters are available. Once the query is set in the form, scroll down to the "Query URL". You will see the parameters after "https://www.waterqualitydata.us/#". For example, if you chose "Nutrient" in the Characteristic Group dropdown, you will see characteristicType=Nutrient in the Query URL. The corresponding argument for dataRetrieval is characteristicType = "Nutrient". dataRetrieval users do not need to include mimeType, and providers is optional (these arguments are picked automatically).

Value

A data frame from the data returned from the Water Quality Portal about the data available for the query parameters.

See Also

whatWQPsites whatWQPdata

Examples

```
# Summary of a single site for the last 5 years:
site_5 <- readWQPsummary(
  siteid = "USGS-07144100",
  summaryYears = 5
)

# Summary of a single site for the full period of record:
site_all <- readWQPsummary(
  siteid = "USGS-07144100",
  summaryYears = "all"
)

# Summary of the data available from streams in a single county:
dane_county_data <- readWQPsummary(
  countycode = "US:55:025",
  summaryYears = 5,
  siteType = "Stream"
)

# Summary of the data all available from lakes in a single county:
lake_sites <- readWQPsummary(
  siteType = "Lake, Reservoir, Impoundment",
  countycode = "US:55:025"
)

# Summary of the data available for the last 5 years in New Jersey:
state1 <- readWQPsummary(
  statecode = "NJ",
  summaryYears = 5,
  siteType = "Stream"
)
```

---

read_waterdata *Generalized USGS Water Data API retrieval function*

Description

Function that allows complex CQL queries. See https://api.waterdata.usgs.gov/docs/ogcapi/complex-queries/ for more information.

Usage

```
read_waterdata(service, CQL, ..., convertType = TRUE)
```

Arguments

service	character, can be any existing collection.
CQL	A string in a Common Query Language format.
...	Additional arguments to send to the request.
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates and qualifier to string vector.

Examples

```
cql <- '{
"op": "and",
"args": [
  {
    "op": "in",
    "args": [
      { "property": "parameter_code" },
      [ "00060", "00065" ]
    ]
  },
 {
    "op": "in",
    "args": [
      { "property": "monitoring_location_id" },
      [ "USGS-07367300", "USGS-03277200" ]
    ]
  }
]
}'

dv_data <- read_waterdata(service = "daily",
                          CQL = cql,
                          time = c("2023-01-01", "2024-01-01"))

# A wildcard in CQL2 is %
# Here's how to get HUCs that fall within 02070010
cql_huc_wildcard <- '{
"op": "like",
"args": [
  { "property": "hydrologic_unit_code" },
  "02070010%"
]
}'

what_huc_sites <- read_waterdata(service = "monitoring-locations",
                                 CQL = cql_huc_wildcard)
```

read_waterdata_channel *Get USGS Channel Measurements*

Description

Channel measurements taken as part of streamflow field measurements.

Usage

```
read_waterdata_channel(
  monitoring_location_id = NA_character_,
  field_visit_id = NA_character_,
  measurement_number = NA_character_,
  time = NA_character_,
  channel_name = NA_character_,
  channel_flow = NA_character_,
  channel_flow_unit = NA_character_,
  channel_width = NA_character_,
  channel_width_unit = NA_character_,
  channel_area = NA_character_,
  channel_area_unit = NA_character_,
  channel_velocity = NA_character_,
  channel_velocity_unit = NA_character_,
  channel_location_distance = NA_character_,
  channel_location_distance_unit = NA_character_,
  channel_location_direction = NA_character_,
  channel_stability = NA_character_,
  channel_material = NA_character_,
  channel_evenness = NA_character_,
  horizontal_velocity_description = NA_character_,
  vertical_velocity_description = NA_character_,
  longitudinal_velocity_description = NA_character_,
  measurement_type = NA_character_,
  last_modified = NA_character_,
  channel_measurement_type = NA_character_,
  properties = NA_character_,
  skipGeometry = NA,
  bbox = NA,
  limit = NA,
  convertType = TRUE,
  no_paging = FALSE
)
```

Arguments

`monitoring_location_id`
: A unique identifier representing a single monitoring location. This corresponds to the `id` field in the `monitoring-locations` endpoint. Monitoring location IDs are created by combining the agency code of the agency responsible for the monitoring location (e.g. USGS) with the ID number of the monitoring location (e.g. 02238500), separated by a hyphen (e.g. USGS-02238500).

  Multiple monitoring_location_ids can be requested as a character vector.

`field_visit_id`
: A universally unique identifier (UUID) for the field visit. Multiple measurements may be made during a single field visit.

`measurement_number`
: Measurement number.

`time`
: The date an observation represents. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

  - A date-time: "2018-02-12T23:20:50Z"
  - A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
  - Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
  - Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

  Only features that have a `time` that intersects the value of datetime are selected. If a feature has multiple temporal properties, it is the decision of the server whether only a single temporal property is used to determine the extent or all relevant temporal properties.

`channel_name`
: The channel name.

`channel_flow`
: Channel discharge.

`channel_flow_unit`
: The units for channel discharge.

`channel_width`
: The channel width.

`channel_width_unit`
: The units for channel width.

`channel_area`
: The channel area.

`channel_area_unit`
: The units for channel area.

`channel_velocity`
: The mean channel velocity.

`channel_velocity_unit`
: The units for channel velocity.

`channel_location_distance`
: The channel location distance.

`channel_location_distance_unit`
: The units for channel location distance.

`channel_location_direction`
: Location of the measurement from the gage.

`channel_stability`
: The stability of the channel material.

`channel_material`
: The channel material.

`channel_evenness`
: The channel evenness from bank to bank.

`horizontal_velocity_description`
: The horizontal velocity description.

`vertical_velocity_description`
: The vertical velocity description.

`longitudinal_velocity_description`
: The longitudinal velocity description.

`measurement_type`
: The measurement type.

`last_modified`
: The last time a record was refreshed in our database. This may happen due to regular operational processes and does not necessarily indicate anything about the measurement has changed. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

  - A date-time: "2018-02-12T23:20:50Z"
  - A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
  - Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
  - Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

  Only features that have a `last_modified` that intersects the value of datetime are selected.

```
                                        field_visit_id = NA_character_,
```

`channel_measurement_type`
: The channel measurement type.

`properties`
: A vector of requested columns to be returned from the query. Available options are: geometry, channel_measurements_id, monitoring_location_id, field_visit_id, measurement_number, time, channel_name, channel_flow, channel_flow_unit, channel_width, channel_width_unit, channel_area, channel_area_unit, channel_velocity, channel_velocity_unit, channel_location_distance, channel_location_distance_unit, channel_stability, channel_material, channel_evenness, horizontal_velocity_description, vertical_velocity_description, longitudinal_velocity_description, measurement_type, last_modified, channel_measurement_type, channel_location_direction. The default (`NA`) will return all columns of the data.

`skipGeometry`
: This option can be used to skip response geometries for each feature. The returning object will be a data frame with no spatial information.

`bbox`
: Only features that have a geometry that intersects the bounding box are selected.The bounding box is provided as four or six numbers, depending on whether the coordinate reference system includes a vertical axis (height or depth). Coordinates are assumed to be in crs 4326. The expected format is a numeric

vector structured: c(xmin,ymin,xmax,ymax). Another way to think of it is c(Western-most longitude, Southern-most latitude, Eastern-most longitude, Northern-most longitude).

`limit` The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (NA) will set the limit to the maximum allowable limit for the service.

`convertType` logical, defaults to TRUE. If TRUE, the function will convert the data to dates and qualifier to string vector.

`no_paging` logical, defaults to FALSE. If TRUE, the data will be requested from a native csv format. This can be dangerous because the data will cut off at 50,000 rows without indication that more data is available. Use TRUE with caution.

Details

You can also use a vector of length 2 for any time queries (such as time or last_modified). The first value is the starting date (or datetime), the second value is the ending date(or datetime). NA's within the vector indicate a half-bound date. For example, `time = c("2024-01-01", NA)` will return all data starting at 2024-01-01. `time = c(NA, "2024-01-01")` will return all data from the beginning of the timeseries until 2024-01-01. By default, time is assumed UTC, although time zone attributes will be accommodated. As an example, setting `time = as.POSIXct(c("2021-01-01 12:00:00","2021-01-01 14:00"), tz = "America/New_York")` will request data that between noon and 2pm eastern time on 2021-01-01. All time values RETURNED from the service are UTC with the exception of daily data, which returns time values in local dates.

Examples

```
site <- "USGS-02238500"
df <- read_waterdata_channel(monitoring_location_id = site)
```

---

read_waterdata_combined_meta

Get USGS Monitoring Location Metadata

Description

This endpoint combines metadata from timeseries and field measurements collections by site.

Usage

```
read_waterdata_combined_meta(
  monitoring_location_id = NA_character_,
  parameter_code = NA_character_,
  parameter_name = NA_character_,
  unit_of_measure = NA_character_,
  statistic_id = NA_character_,
  parameter_description = NA_character_,
  data_type = NA_character_,
  computation_identifier = NA_character_,
  computation_period_identifier = NA_character_,
  thresholds = NA_character_,
  sublocation_identifier = NA_character_,
  primary = NA_character_,
  web_description = NA_character_,
  parent_time_series_id = NA_character_,
  begin = NA_character_,
  end = NA_character_,
  last_modified = NA_character_,
  agency_code = NA_character_,
  agency_name = NA_character_,
  monitoring_location_number = NA_character_,
  monitoring_location_name = NA_character_,
  district_code = NA_character_,
  country_code = NA_character_,
  country_name = NA_character_,
  state_code = NA_character_,
  state_name = NA_character_,
  county_code = NA_character_,
  county_name = NA_character_,
  minor_civil_division_code = NA_character_,
  site_type_code = NA_character_,
  site_type = NA_character_,
  hydrologic_unit_code = NA_character_,
  basin_code = NA_character_,
  altitude = NA_character_,
  altitude_accuracy = NA_character_,
  altitude_method_code = NA_character_,
  altitude_method_name = NA_character_,
  vertical_datum = NA_character_,
  vertical_datum_name = NA_character_,
  horizontal_positional_accuracy_code = NA_character_,
  horizontal_positional_accuracy = NA_character_,
  horizontal_position_method_code = NA_character_,
  horizontal_position_method_name = NA_character_,
  original_horizontal_datum = NA_character_,
  original_horizontal_datum_name = NA_character_,
  drainage_area = NA_character_,
```

```
  contributing_drainage_area = NA_character_,
  time_zone_abbreviation = NA_character_,
  uses_daylight_savings = NA_character_,
  construction_date = NA_character_,
  aquifer_code = NA_character_,
  national_aquifer_code = NA_character_,
  aquifer_type_code = NA_character_,
  well_constructed_depth = NA_character_,
  hole_constructed_depth = NA_character_,
  depth_source_code = NA_character_,
  properties = NA_character_,
  skipGeometry = NA,
  bbox = NA,
  limit = NA,
  convertType = TRUE,
  no_paging = FALSE
)
```

Arguments

`monitoring_location_id`
: A unique identifier representing a single monitoring location. This corresponds to the `id` field in the `monitoring-locations` endpoint. Monitoring location IDs are created by combining the agency code of the agency responsible for the monitoring location (e.g. USGS) with the ID number of the monitoring location (e.g. 02238500), separated by a hyphen (e.g. USGS-02238500).

  Multiple monitoring_location_ids can be requested as a character vector.

`parameter_code`
: Parameter codes are 5-digit codes used to identify the constituent measured and the units of measure. A complete list of parameter codes and associated groupings can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/parameter-codes/items.

  Multiple parameter_codes can be requested as a character vector.

`parameter_name`
: A human-understandable name corresponding to `parameter_code`. Multiple parameter_names can be requested as a character vector.

`unit_of_measure`
: A human-readable description of the units of measurement associated with an observation.

`statistic_id`
: A code corresponding to the statistic an observation represents. Example codes include 00001 (max), 00002 (min), and 00003 (mean). A complete list of codes and their descriptions can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/statistic-codes/items.

  Multiple statistic_ids can be requested as a character vector.

`parameter_description`
: A description of what the parameter code represents, as used by WDFN and other USGS data dissemination products.

`data_type`
: The computational period type of data collected at the monitoring location.

`computation_identifier`
: Indicates whether the data from this time series represent a specific statistical computation. Multiple computation_identifiers can be requested as a character vector.

`computation_period_identifier`
: Multiple computation_period_identifiers can be requested as a character vector.

`thresholds`
: Thresholds represent known numeric limits for a time series, for example the historic maximum value for a parameter or a level below which a sensor is non-operative. These thresholds are sometimes used to automatically determine if an observation is erroneous due to sensor error, and therefore shouldn't be included in the time series.

`sublocation_identifier`
: An optional human-readable identifier used to specify where measurements are recorded at a monitoring location.

`primary`
: A flag identifying if the time series is a "primary" time series. "Primary" time series (which have this flag) are standard observations which undergo Bureau review and approval processes. Non-primary time series, which will have missing values for "primary", are provisional datasets made available to meet the need for timely best science and to assist with daily operations which need real-time information. Non-primary time series data are only retained by this system for 120 days.

`web_description`
: A description of what this time series represents, as used by WDFN and other USGS data dissemination products.

`parent_time_series_id`
: The unique identifier representing the parent or "upchain" time series that a daily values time series is generated from. Daily values time series have one and only one parent time series.

`begin`
: The datetime of the earliest observation in the time series. Together with `end`, this field represents the period of record of a time series. Note that some time series may have large gaps in their collection record. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

  - A date-time: "2018-02-12T23:20:50Z"
  - A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
  - Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
  - Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

  Only features that have a `begin` that intersects the value of datetime are selected.

  See also Details below for more information.

`end`
: The datetime of the most recent observation in the time series. Data returned by this endpoint updates at most once per day, and potentially less frequently than that, and as such there may be more recent observations within a time series than the time series `end` value reflects. Together with `begin`, this field represents the period of record of a time series. It is additionally used to determine whether

a time series is "active". You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a end that intersects the value of datetime are selected.

See also Details below for more information.

`last_modified` The last time a record was refreshed in our database. This may happen due to regular operational processes and does not necessarily indicate anything about the measurement has changed. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `last_modified` that intersects the value of datetime are selected.

See also Details below for more information.

`agency_code` The agency that is reporting the data. Agency codes are fixed values assigned by the National Water Information System (NWIS). A list of agency codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/agency-codes/items.

`agency_name` The name of the agency that is reporting the data.

`monitoring_location_number`
Each monitoring location in the USGS data base has a unique 8- to 15-digit identification number.

`monitoring_location_name`
This is the official name of the monitoring location in the database. For well information this can be a district-assigned local number.

`district_code` The Water Science Centers (WSCs) across the United States use the FIPS state code as the district code. In some case, monitoring locations and samples may be managed by a water science center that is adjacent to the state in which the monitoring location actually resides. For example a monitoring location may have a district code of 30 which translates to Montana, but the state code could be 56 for Wyoming because that is where the monitoring location actually is located.

`country_code` The code for the country in which the monitoring location is located.

`country_name` The name of the country in which the monitoring location is located.

state_code State code. A two-digit ANSI code (formerly FIPS code) as defined by the American National Standards Institute, to define States and equivalents. A three-digit ANSI code is used to define counties and county equivalents. A lookup table is available. The only countries with political subdivisions other than the US are Mexico and Canada. The Mexican states have US state codes ranging from 81-86 and Canadian provinces have state codes ranging from 90-98.

state_name The name of the state or state equivalent in which the monitoring location is located.

county_code The code for the county or county equivalent (parish, borough, etc.) in which the monitoring location is located. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/counties/items.

county_name The name of the county or county equivalent (parish, borough, etc.) in which the monitoring location is located. [A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/counties/items.

minor_civil_division_code
Codes for primary governmental or administrative divisions of the county or county equivalent in which the monitoring location is located.

site_type_code A code describing the hydrologic setting of the monitoring location. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/site-types/items.

site_type A description of the hydrologic setting of the monitoring location. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/site-types/items.

hydrologic_unit_code
The United States is divided and sub-divided into successively smaller hydrologic units which are classified into four levels: regions, sub-regions, accounting units, and cataloging units. The hydrologic units are arranged within each other, from the smallest (cataloging units) to the largest (regions). Each hydrologic unit is identified by a unique hydrologic unit code (HUC) consisting of two to eight digits based on the four levels of classification in the hydrologic unit system.

basin_code The Basin Code or "drainage basin code" is a two-digit code that further subdivides the 8-digit hydrologic-unit code. The drainage basin code is defined by the USGS State Office where the monitoring location is located.

altitude Altitude of the monitoring location referenced to the specified Vertical Datum.

altitude_accuracy
Accuracy of the altitude, in feet. An accuracy of +/- 0.1 foot would be entered as ".1". Many altitudes are interpolated from the contours on topographic maps; accuracies determined in this way are generally entered as one-half of the contour interval.

altitude_method_code
Codes representing the method used to measure altitude.

altitude_method_name
The name of the method used to measure altitude.

vertical_datum The datum used to determine altitude and vertical position at the monitoring location. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/altitude-datums/items.

vertical_datum_name
: The datum used to determine altitude and vertical position at the monitoring location. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/altitude-datums/items.

horizontal_positional_accuracy_code
: Indicates the accuracy of the latitude longitude values. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-accuracy-codes/items.

horizontal_positional_accuracy
: Indicates the accuracy of the latitude longitude values. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-accuracy-codes/items.

horizontal_position_method_code
: Indicates the method used to determine latitude longitude values. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-method-codes/items.

horizontal_position_method_name
: Indicates the method used to determine latitude longitude values. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-method-codes/items.

original_horizontal_datum
: Coordinates are published in EPSG:4326 / WGS84 / World Geodetic System 1984. This field indicates the original datum used to determine coordinates before they were converted. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-datum-codes/items.

original_horizontal_datum_name
: Coordinates are published in EPSG:4326 / WGS84 / World Geodetic System 1984. This field indicates the original datum used to determine coordinates before they were converted. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-datum-codes/items.

drainage_area
: The area enclosed by a topographic divide from which direct surface runoff from precipitation normally drains by gravity into the stream above that point.

contributing_drainage_area
: The contributing drainage area of a lake, stream, wetland, or estuary monitoring location, in square miles. This item should be present only if the contributing area is different from the total drainage area. This situation can occur when part of the drainage area consists of very porous soil or depressions that either allow all runoff to enter the groundwater or traps the water in ponds so that rainfall does not contribute to runoff. A transbasin diversion can also affect the total drainage area.

time_zone_abbreviation
: A short code describing the time zone used by a monitoring location.

uses_daylight_savings
: A flag indicating whether or not a monitoring location uses daylight savings.

construction_date
: Date the well was completed.

`aquifer_code` Local aquifers in the USGS water resources data base are identified by a geohydrologic unit code (a three-digit number related to the age of the formation, followed by a 4 or 5 character abbreviation for the geologic unit or aquifer name).

`national_aquifer_code`
National aquifers are the principal aquifers or aquifer systems in the United States, defined as regionally extensive aquifers or aquifer systems that have the potential to be used as a source of potable water. Not all groundwater monitoring locations can be associated with a National Aquifer. Such monitoring locations will not be retrieved using this search criteria. A list of National aquifer codes and names is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/national-aquifer-codes/items.

`aquifer_type_code`
Describes the confinement status of an aquifer at the monitoring location. A confined aquifer is an aquifer below the land surface that is saturated with water. A water table–or unconfined–aquifer is an aquifer whose upper water surface (water table) is at atmospheric pressure, and thus is able to rise and fall.

`well_constructed_depth`
The depth of the finished well, in feet below land surface datum. Note: Not all groundwater monitoring locations have information on Well Depth. Such monitoring locations will not be retrieved using this search criteria.

`hole_constructed_depth`
The total depth to which the hole is drilled, in feet below land surface datum. Note: Not all groundwater monitoring locations have information on Hole Depth. Such monitoring locations will not be retrieved using this search criteria.

`depth_source_code`
A code indicating the source of water-level data.

`properties` A vector of requested columns to be returned from the query. Available options are: geometry, monitoring_location_id, agency_code, agency_name, monitoring_location_number, monitoring_location_name, district_code, country_code, country_name, state_code, state_name, county_code, county_name, minor_civil_division_code, site_type_code, site_type, hydrologic_unit_code, basin_code, altitude, altitude_accuracy, altitude_method_code, altitude_method_name, vertical_datum, vertical_datum_name, horizontal_positional_accuracy_code, horizontal_positional_accuracy, horizontal_position_method_code, horizontal_position_method_name, original_horizontal_datum, original_horizontal_datum_name, drainage_area, contributing_drainage_area, time_zone_abbreviation, uses_daylight_savings, construction_date, aquifer_code, national_aquifer_code, aquifer_type_code, well_constructed_depth, hole_constructed_depth, depth_source_code, field_measurement_id, unit_of_measure, parameter_name, parameter_code, statistic_id, last_modified, begin, end, data_type, computation_identifier, thresholds, sublocatio The default (`NA`) will return all columns of the data.

`skipGeometry` This option can be used to skip response geometries for each feature. The returning object will be a data frame with no spatial information.

`bbox` Only features that have a geometry that intersects the bounding box are selected.The bounding box is provided as four or six numbers, depending on whether the coordinate reference system includes a vertical axis (height or depth). Coordinates are assumed to be in crs 4326. The expected format is a numeric vector structured: c(xmin,ymin,xmax,ymax). Another way to think of it is

	c(Western-most longitude, Southern-most latitude, Eastern-most longitude, Northern-most longitude).
limit	The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (NA) will set the limit to the maximum allowable limit for the service.
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates and qualifier to string vector.
no_paging	logical, defaults to FALSE. If TRUE, the data will be requested from a native csv format. This can be dangerous because the data will cut off at 50,000 rows without indication that more data is available. Use TRUE with caution.

Details

You can also use a vector of length 2 for any time queries (such as time or last_modified). The first value is the starting date (or datetime), the second value is the ending date(or datetime). NA's within the vector indicate a half-bound date. For example, `time = c("2024-01-01", NA)` will return all data starting at 2024-01-01. `time = c(NA, "2024-01-01")` will return all data from the beginning of the timeseries until 2024-01-01. By default, time is assumed UTC, although time zone attributes will be accommodated. As an example, setting `time = as.POSIXct(c("2021-01-01 12:00:00","2021-01-01 14:00"), tz = "America/New_York")` will request data that between noon and 2pm eastern time on 2021-01-01. All time values RETURNED from the service are UTC with the exception of daily data, which returns time values in local dates.

Examples

```
site <- "USGS-05407000"
available_data_sf <- read_waterdata_combined_meta(monitoring_location_id = site)

groundwater <- read_waterdata_combined_meta(monitoring_location_id = "USGS-375907091432201")

date_wi_data <- read_waterdata_combined_meta(state_name = "Wisconsin",
                                     county_name = "Dane County")


multi_site <- read_waterdata_combined_meta(
                             monitoring_location_id =  c("USGS-451605097071701",
                                                         "USGS-263819081585801"),
                             parameter_code = c("62611", "72019"))


surface_water <- read_waterdata_combined_meta(
                         monitoring_location_id = c("USGS-07069000",
                                                    "USGS-07064000",
                                                    "USGS-07068000"),
                         end = "P1M",
                         parameter_code = "00060")
```

```
hucs <- read_waterdata_combined_meta(
                        hydrologic_unit_code = c("11010008", "11010009"),
                        site_type = c("Stream", "Spring")
)

site_list <- read_waterdata_combined_meta(
                        monitoring_location_id = hucs$monitoring_location_id
)
```

`read_waterdata_continuous`

Get Continuous USGS Water Data

Description

Continuous data are collected via automated sensors installed at a monitoring location. They are collected at a high frequency and often at a fixed 15-minute interval. Depending on the specific monitoring location, the data may be transmitted automatically via telemetry and be available on WDFN within minutes of collection, while other times the delivery of data may be delayed if the monitoring location does not have the capacity to automatically transmit data. Continuous data are described by parameter name and parameter code (pcode). These data might also be referred to as "instantaneous values" or "IV".

Currently, the services only allow up to 3 years of data to be requested with a single request. If no "time" is specified, the service will return the last single year of data. If this is a bottleneck, please check back for new direct download functions that are expected to be available sometime in 2026.

Geometry output is not supported in the continuous data API endpoint.

Usage

```
read_waterdata_continuous(
  monitoring_location_id = NA_character_,
  parameter_code = NA_character_,
  properties = NA_character_,
  time_series_id = NA_character_,
  approval_status = NA_character_,
  unit_of_measure = NA_character_,
  qualifier = NA_character_,
  value = NA,
  last_modified = NA_character_,
  time = NA_character_,
  limit = NA,
  convertType = TRUE,
  no_paging = FALSE
)
```

Arguments

`monitoring_location_id`
: A unique identifier representing a single monitoring location. This corresponds to the `id` field in the `monitoring-locations` endpoint. Monitoring location IDs are created by combining the agency code of the agency responsible for the monitoring location (e.g. USGS) with the ID number of the monitoring location (e.g. 02238500), separated by a hyphen (e.g. USGS-02238500).

  Multiple monitoring_location_ids can be requested as a character vector.

`parameter_code`
: Parameter codes are 5-digit codes used to identify the constituent measured and the units of measure. A complete list of parameter codes and associated groupings can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/parameter-codes/items.

  Multiple parameter_codes can be requested as a character vector.

`properties`
: A vector of requested columns to be returned from the query. Available options are: geometry, continuous_id, time_series_id, monitoring_location_id, parameter_code, statistic_id, time, value, unit_of_measure, approval_status, qualifier, last_modified. The default (`NA`) will return all columns of the data.

`time_series_id`
: A unique identifier representing a single time series. This corresponds to the `id` field in the `time-series-metadata` endpoint.

  Multiple time_series_ids can be requested as a character vector.

`approval_status`
: Some of the data that you have obtained from this U.S. Geological Survey database may not have received Director's approval. Any such data values are qualified as provisional and are subject to revision. Provisional data are released on the condition that neither the USGS nor the United States Government may be held liable for any damages resulting from its use. This field reflects the approval status of each record, and is either "Approved", meaining processing review has been completed and the data is approved for publication, or "Provisional" and subject to revision. For more information about provisional data, go to https://waterdata.usgs.gov/provisional-data-statement/.

`unit_of_measure`
: A human-readable description of the units of measurement associated with an observation.

`qualifier`
: This field indicates any qualifiers associated with an observation, for instance if a sensor may have been impacted by ice or if values were estimated.

`value`
: The value of the observation. Values are transmitted as strings in the JSON response format in order to preserve precision.

`last_modified`
: The last time a record was refreshed in our database. This may happen due to regular operational processes and does not necessarily indicate anything about the measurement has changed. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

  - A date-time: "2018-02-12T23:20:50Z"
  - A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
  - Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"

- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `last_modified` that intersects the value of datetime are selected.

See also Details below for more information.

`time` The date an observation represents. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `time` that intersects the value of datetime are selected. If a feature has multiple temporal properties, it is the decision of the server whether only a single temporal property is used to determine the extent or all relevant temporal properties.

See also Details below for more information.

`limit` The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (`NA`) will set the limit to the maximum allowable limit for the service.

`convertType` logical, defaults to `TRUE`. If `TRUE`, the function will convert the data to dates and qualifier to string vector, and sepcifically order the returning data frame by time and monitoring_location_id.

`no_paging` logical, defaults to `FALSE`. If `TRUE`, the data will be requested from a native csv format. This can be dangerous because the data will cut off at 50,000 rows without indication that more data is available. Use `TRUE` with caution.

Details

You can also use a vector of length 2 for any time queries (such as time or last_modified). The first value is the starting date (or datetime), the second value is the ending date(or datetime). NA's within the vector indicate a half-bound date. For example, `time = c("2024-01-01", NA)` will return all data starting at 2024-01-01. `time = c(NA, "2024-01-01")` will return all data from the beginning of the timeseries until 2024-01-01. By default, time is assumed UTC, although time zone attributes will be accommodated. As an example, setting `time = as.POSIXct(c("2021-01-01 12:00:00","2021-01-01 14:00"), tz = "America/New_York")` will request data that between noon and 2pm eastern time on 2021-01-01. All time values RETURNED from the service are UTC with the exception of daily data, which returns time values in local dates.

Examples

```
site <- "USGS-451605097071701"
```

```
pcode <- "72019"

uv_data_trim <- read_waterdata_continuous(monitoring_location_id = site,
                           parameter_code = pcode,
                           properties = c("value", "time"),
                           time = as.POSIXct(c("2026-02-07 12:00",
                                               "2026-02-08 12:00"),
                                             tz = "America/Chicago"))

uv_data <- read_waterdata_continuous(monitoring_location_id = site,
                           parameter_code = pcode,
                           time = "P2D")


# Only return data that has been modified in last 7 days
multi_site2 <- read_waterdata_continuous(monitoring_location_id =  c("USGS-451605097071701",
                                                                     "USGS-14181500"),
                                         parameter_code = c("00060", "72019"),
                                         last_modified = "P7D")

# how to split up request into roughly 3 year chunks

site <- "USGS-0208458892"
pcode <- "00095" # Specific conductance
start <- as.Date("2013-01-01")
end <- as.Date("2025-12-31")

n_days <- difftime(end, start, units = "days")

# create a vector of dates that are about 3 years apart:
time_chunks <- seq(from = start,
                   to = end,
                   length.out = ceiling(n_days/(3*365.25)) + 1)


# create a list where each element starts at the beginning
# of a chunk, and ends the day before the next chunk:
time_df <- data.frame(start = time_chunks[-length(time_chunks)],
                      end = time_chunks[-1]-1)

#all_data <- data.frame()
#for(i in seq_along(time_df$start)){
#   sub_df <- read_waterdata_continuous(monitoring_location_id = site,
#                                       parameter_code = pcode,
#                                      time = c(time_df$start[i],
#                                               time_df$end[i]))
#   all_data <- rbind(all_data, sub_df)
# }

# Set the time to Eastern:
# all_data$time <- lubridate::force_tz(all_data$time, "America/New_York")
```

`read_waterdata_daily` *Get USGS Daily Data*

Description

Daily data provide one data value to represent water conditions for the day. Throughout much of the history of the USGS, the primary water data available was daily data collected manually at the monitoring location once each day. With improved availability of computer storage and automated transmission of data, the daily data published today are generally a statistical summary or metric of the continuous data collected each day, such as the daily mean, minimum, or maximum value. Daily data are automatically calculated from the continuous data of the same parameter code and are described by parameter code and a statistic code. These data have also been referred to as "daily values" or "DV".

Usage

```
read_waterdata_daily(
  monitoring_location_id = NA_character_,
  parameter_code = NA_character_,
  statistic_id = NA_character_,
  properties = NA_character_,
  time_series_id = NA_character_,
  approval_status = NA_character_,
  unit_of_measure = NA_character_,
  qualifier = NA_character_,
  value = NA,
  last_modified = NA_character_,
  skipGeometry = NA,
  time = NA_character_,
  bbox = NA,
  limit = NA,
  convertType = TRUE,
  no_paging = FALSE
)
```

Arguments

`monitoring_location_id`

A unique identifier representing a single monitoring location. This corresponds to the `id` field in the `monitoring-locations` endpoint. Monitoring location IDs are created by combining the agency code of the agency responsible for the monitoring location (e.g. USGS) with the ID number of the monitoring location (e.g. 02238500), separated by a hyphen (e.g. USGS-02238500).

Multiple monitoring_location_ids can be requested as a character vector.

parameter_code Parameter codes are 5-digit codes used to identify the constituent measured and the units of measure. A complete list of parameter codes and associated groupings can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/parameter-codes/items.

Multiple parameter_codes can be requested as a character vector.

statistic_id A code corresponding to the statistic an observation represents. Example codes include 00001 (max), 00002 (min), and 00003 (mean). A complete list of codes and their descriptions can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/statistic-codes/items.

Multiple statistic_ids can be requested as a character vector.

properties A vector of requested columns to be returned from the query. Available options are: geometry, daily_id, time_series_id, monitoring_location_id, parameter_code, statistic_id, time, value, unit_of_measure, approval_status, qualifier, last_modified. The default (NA) will return all columns of the data.

time_series_id A unique identifier representing a single time series. This corresponds to the id field in the time-series-metadata endpoint.

Multiple time_series_ids can be requested as a character vector.

approval_status

Some of the data that you have obtained from this U.S. Geological Survey database may not have received Director's approval. Any such data values are qualified as provisional and are subject to revision. Provisional data are released on the condition that neither the USGS nor the United States Government may be held liable for any damages resulting from its use. This field reflects the approval status of each record, and is either "Approved", meaining processing review has been completed and the data is approved for publication, or "Provisional" and subject to revision. For more information about provisional data, go to https://waterdata.usgs.gov/provisional-data-statement/.

unit_of_measure

A human-readable description of the units of measurement associated with an observation.

qualifier This field indicates any qualifiers associated with an observation, for instance if a sensor may have been impacted by ice or if values were estimated.

value The value of the observation. Values are transmitted as strings in the JSON response format in order to preserve precision.

last_modified The last time a record was refreshed in our database. This may happen due to regular operational processes and does not necessarily indicate anything about the measurement has changed. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `last_modified` that intersects the value of datetime are selected.

See also Details below for more information.

`skipGeometry` This option can be used to skip response geometries for each feature. The returning object will be a data frame with no spatial information.

`time` The date an observation represents. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `time` that intersects the value of datetime are selected. If a feature has multiple temporal properties, it is the decision of the server whether only a single temporal property is used to determine the extent or all relevant temporal properties.

See also Details below for more information.

`bbox` Only features that have a geometry that intersects the bounding box are selected.The bounding box is provided as four or six numbers, depending on whether the coordinate reference system includes a vertical axis (height or depth). Coordinates are assumed to be in crs 4326. The expected format is a numeric vector structured: c(xmin,ymin,xmax,ymax). Another way to think of it is c(Western-most longitude, Southern-most latitude, Eastern-most longitude, Northern-most longitude).

`limit` The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (`NA`) will set the limit to the maximum allowable limit for the service.

`convertType` logical, defaults to `TRUE`. If `TRUE`, the function will convert the data to dates and qualifier to string vector.

`no_paging` logical, defaults to `FALSE`. If `TRUE`, the data will be requested from a native csv format. This can be dangerous because the data will cut off at 50,000 rows without indication that more data is available. Use `TRUE` with caution.

Details

You can also use a vector of length 2 for any time queries (such as time or last_modified). The first value is the starting date (or datetime), the second value is the ending date(or datetime). NA's within the vector indicate a half-bound date. For example, `time = c("2024-01-01", NA)` will return all data starting at 2024-01-01. `time = c(NA, "2024-01-01")` will return all data from the beginning of the timeseries until 2024-01-01. By default, time is assumed UTC, although time zone attributes will be accommodated. As an example, setting `time = as.POSIXct(c("2021-01-01 12:00:00","2021-01-01 14:00"), tz = "America/New_York")` will request data that between noon and 2pm eastern time on 2021-01-01. All time values RETURNED from the service are UTC with the exception of daily data, which returns time values in local dates.

Examples

```
site <- "USGS-02238500"
dv_data_sf <- read_waterdata_daily(monitoring_location_id = site,
                              parameter_code = "00060",
                              time = c("2021-01-01", "2022-01-01"))

dv_data_last_modified <- read_waterdata_daily(monitoring_location_id = site,
                              parameter_code = "00060",
                              last_modified = "P7D")

dv_data_trim <- read_waterdata_daily(monitoring_location_id = site,
                            parameter_code = "00060",
                            properties = c("value",
                                           "time"),
                            time = c("2021-01-01", "2022-01-01"))

dv_data <- read_waterdata_daily(monitoring_location_id = site,
                             parameter_code = "00060",
                             skipGeometry = TRUE)

dv_data_period <- read_waterdata_daily(monitoring_location_id = site,
                                  parameter_code = "00060",
                                  time = "P7D")

multi_site <- read_waterdata_daily(monitoring_location_id =  c("USGS-01491000",
                                                               "USGS-01645000"),
                              parameter_code = c("00060", "00010"),
                              time = c("2023-01-01", "2024-01-01"))

dv_data_quick <- read_waterdata_daily(monitoring_location_id = site,
                                   parameter_code = "00060",
                                   no_paging = TRUE)

dv_post <- read_waterdata_daily(monitoring_location_id = site,
                                approval_status = c("Approved", "Provisional"))
# Don't attach "request" attribute:
options("dataRetrieval.attach_request" = FALSE)
dv_data_no_request <- read_waterdata_daily(monitoring_location_id = site,
                              parameter_code = "00060",
                              time = c("2021-01-01", "2022-01-01"))
```

`read_waterdata_field_measurements`

Get USGS Field Measurement Water Data

Description

Field measurements are physically measured values collected during a visit to the monitoring location. Field measurements consist of measurements of gage height and discharge, and readings of groundwater levels, and are primarily used as calibration readings for the automated sensors collecting continuous data. They are collected at a low frequency, and delivery of the data in WDFN may be delayed due to data processing time.

Usage

```
read_waterdata_field_measurements(
  monitoring_location_id = NA_character_,
  parameter_code = NA_character_,
  observing_procedure_code = NA_character_,
  properties = NA_character_,
  field_visit_id = NA_character_,
  approval_status = NA_character_,
  unit_of_measure = NA_character_,
  qualifier = NA_character_,
  value = NA,
  last_modified = NA_character_,
  observing_procedure = NA_character_,
  vertical_datum = NA_character_,
  measuring_agency = NA_character_,
  control_condition = NA_character_,
  measurement_rated = NA_character_,
  skipGeometry = NA,
  time = NA_character_,
  bbox = NA,
  limit = NA,
  convertType = TRUE,
  no_paging = FALSE
)
```

Arguments

`monitoring_location_id`
: A unique identifier representing a single monitoring location. This corresponds to the `id` field in the `monitoring-locations` endpoint. Monitoring location IDs are created by combining the agency code of the agency responsible for the monitoring location (e.g. USGS) with the ID number of the monitoring location (e.g. 02238500), separated by a hyphen (e.g. USGS-02238500).

  Multiple monitoring_location_ids can be requested as a character vector.

`parameter_code`
: Parameter codes are 5-digit codes used to identify the constituent measured and the units of measure. A complete list of parameter codes and associated groupings can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/parameter-codes/items.

  Multiple parameter_codes can be requested as a character vector.

observing_procedure_code
: A short code corresponding to the observing procedure for the field measurement.

properties
: A vector of requested columns to be returned from the query. Available options are: geometry, field_measurement_id, field_visit_id, parameter_code, monitoring_location_id, observing_procedure_code, observing_procedure, value, unit_of_measure, time, qualifier, vertical_datum, approval_status, measuring_agency, last_modified, control_condition, measurement_rated. The default (NA) will return all columns of the data.

field_visit_id
: A universally unique identifier (UUID) for the field visit. Multiple measurements may be made during a single field visit.

approval_status
: Some of the data that you have obtained from this U.S. Geological Survey database may not have received Director's approval. Any such data values are qualified as provisional and are subject to revision. Provisional data are released on the condition that neither the USGS nor the United States Government may be held liable for any damages resulting from its use. This field reflects the approval status of each record, and is either "Approved", meaining processing review has been completed and the data is approved for publication, or "Provisional" and subject to revision. For more information about provisional data, go to https://waterdata.usgs.gov/provisional-data-statement/.

unit_of_measure
: A human-readable description of the units of measurement associated with an observation.

qualifier
: This field indicates any qualifiers associated with an observation, for instance if a sensor may have been impacted by ice or if values were estimated.

value
: The value of the observation. Values are transmitted as strings in the JSON response format in order to preserve precision.

last_modified
: The last time a record was refreshed in our database. This may happen due to regular operational processes and does not necessarily indicate anything about the measurement has changed. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

  - A date-time: "2018-02-12T23:20:50Z"
  - A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
  - Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
  - Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

  Only features that have a last_modified that intersects the value of datetime are selected.

  See also Details below for more information.

observing_procedure
: Water measurement or water-quality observing procedure descriptions.

vertical_datum
: The datum used to determine altitude and vertical position at the monitoring location. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/altitude-datums/items.

`measuring_agency`
: The agency performing the measurement.

`control_condition`
: What and where the control of flow is for the gage pool.

`measurement_rated`
: Rated measurement based on the hydrologic/hydraulic conditions in which the measurement was made (excellent (2 percent), good (5 percent), fair (8 percent), or poor (more than 8 percent). percent)

`skipGeometry`
: This option can be used to skip response geometries for each feature. The returning object will be a data frame with no spatial information.

`time`
: The date an observation represents. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

  - A date-time: "2018-02-12T23:20:50Z"
  - A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
  - Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
  - Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

  Only features that have a `time` that intersects the value of datetime are selected. If a feature has multiple temporal properties, it is the decision of the server whether only a single temporal property is used to determine the extent or all relevant temporal properties.

  See also Details below for more information.

`bbox`
: Only features that have a geometry that intersects the bounding box are selected.The bounding box is provided as four or six numbers, depending on whether the coordinate reference system includes a vertical axis (height or depth). Coordinates are assumed to be in crs 4326. The expected format is a numeric vector structured: c(xmin,ymin,xmax,ymax). Another way to think of it is c(Western-most longitude, Southern-most latitude, Eastern-most longitude, Northern-most longitude).

`limit`
: The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (`NA`) will set the limit to the maximum allowable limit for the service.

`convertType`
: logical, defaults to `TRUE`. If `TRUE`, the function will convert the data to dates and qualifier to string vector.

`no_paging`
: logical, defaults to `FALSE`. If `TRUE`, the data will be requested from a native csv format. This can be dangerous because the data will cut off at 50,000 rows without indication that more data is available. Use `TRUE` with caution.

Details

You can also use a vector of length 2 for any time queries (such as time or last_modified). The first value is the starting date (or datetime), the second value is the ending date(or datetime). NA's within the vector indicate a half-bound date. For example, `time = c("2024-01-01", NA)` will return all

data starting at 2024-01-01. `time = c(NA, "2024-01-01")` will return all data from the beginning of the timeseries until 2024-01-01. By default, time is assumed UTC, although time zone attributes will be accommodated. As an example, setting `time = as.POSIXct(c("2021-01-01 12:00:00","2021-01-01 14:00"), tz = "America/New_York")` will request data that between noon and 2pm eastern time on 2021-01-01. All time values RETURNED from the service are UTC with the exception of daily data, which returns time values in local dates.

Examples

```
site <- "USGS-02238500"
field_data_sf <- read_waterdata_field_measurements(monitoring_location_id = site)

groundwater <- read_waterdata_field_measurements(monitoring_location_id = "USGS-375907091432201")

field_data <- read_waterdata_field_measurements(monitoring_location_id = "USGS-02238500",
                           parameter_code = "00060",
                           time = as.POSIXct(c("2024-02-26 15:00:00",
                                               "2025-08-27 12:00:00"),
                                             tz = "America/Chicago"),
                           skipGeometry = TRUE)

gwl_data_period <- read_waterdata_field_measurements(
                                  monitoring_location_id = "USGS-375907091432201",
                                  parameter_code = "72019",
                                  time = "P20Y")

multi_site <- read_waterdata_field_measurements(
                              monitoring_location_id =  c("USGS-451605097071701",
                                                          "USGS-263819081585801"),
                              parameter_code = c("62611", "72019"))

old_df <- read_waterdata_field_measurements(monitoring_location_id = "USGS-425957088141001",
                                            time = c("1980-01-01", NA))

surface_water <- read_waterdata_field_measurements(
                         monitoring_location_id = c("USGS-07069000",
                                                    "USGS-07064000",
                                                    "USGS-07068000"),
                         time = "2024-07-01T00:00:00Z/..",
                         parameter_code = "00060")
```

`read_waterdata_field_meta`

Get USGS Field Measurement Metadata

Description

This endpoint provides metadata about field measurement collections, including when the earliest and most recent observations for a parameter occurred at a monitoring location and its units.

Usage

```
read_waterdata_field_meta(
  monitoring_location_id = NA_character_,
  parameter_code = NA_character_,
  parameter_name = NA_character_,
  parameter_description = NA_character_,
  begin = NA_character_,
  end = NA_character_,
  last_modified = NA_character_,
  properties = NA_character_,
  skipGeometry = NA,
  bbox = NA,
  limit = NA,
  convertType = TRUE,
  no_paging = FALSE
)
```

Arguments

`monitoring_location_id`
: A unique identifier representing a single monitoring location. This corresponds to the `id` field in the `monitoring-locations` endpoint. Monitoring location IDs are created by combining the agency code of the agency responsible for the monitoring location (e.g. USGS) with the ID number of the monitoring location (e.g. 02238500), separated by a hyphen (e.g. USGS-02238500).

  Multiple monitoring_location_ids can be requested as a character vector.

`parameter_code`
: Parameter codes are 5-digit codes used to identify the constituent measured and the units of measure. A complete list of parameter codes and associated groupings can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/parameter-codes/items.

  Multiple parameter_codes can be requested as a character vector.

`parameter_name`
: A human-understandable name corresponding to `parameter_code`. Multiple parameter_names can be requested as a character vector.

`parameter_description`
: A description of what the parameter code represents, as used by WDFN and other USGS data dissemination products.

  Multiple parameter_descriptions can be requested as a character vector.

`begin`
: The datetime of the earliest observation in the time series. Together with end, this field represents the period of record of a time series. Note that some time series may have large gaps in their collection record. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `begin` that intersects the value of datetime are selected.

See also Details below for more information.

`end` The datetime of the most recent observation in the time series. Data returned by this endpoint updates at most once per day, and potentially less frequently than that, and as such there may be more recent observations within a time series than the time series `end` value reflects. Together with `begin`, this field represents the period of record of a time series. It is additionally used to determine whether a time series is "active". You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `end` that intersects the value of datetime are selected.

See also Details below for more information.

`last_modified` The last time a record was refreshed in our database. This may happen due to regular operational processes and does not necessarily indicate anything about the measurement has changed. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `last_modified` that intersects the value of datetime are selected.

See also Details below for more information.

`properties` A vector of requested columns to be returned from the query. Available options are: geometry, field_measurement_id, monitoring_location_id, parameter_code, parameter_name, parameter_description, begin, end, last_modified. The default (`NA`) will return all columns of the data.

`skipGeometry` This option can be used to skip response geometries for each feature. The returning object will be a data frame with no spatial information.

`bbox` Only features that have a geometry that intersects the bounding box are selected.The bounding box is provided as four or six numbers, depending on whether the coordinate reference system includes a vertical axis (height or depth).

Coordinates are assumed to be in crs 4326. The expected format is a numeric vector structured: c(xmin,ymin,xmax,ymax). Another way to think of it is c(Western-most longitude, Southern-most latitude, Eastern-most longitude, Northern-most longitude).

`limit` The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (`NA`) will set the limit to the maximum allowable limit for the service.

`convertType` logical, defaults to `TRUE`. If `TRUE`, the function will convert the data to dates and qualifier to string vector.

`no_paging` logical, defaults to `FALSE`. If `TRUE`, the data will be requested from a native csv format. This can be dangerous because the data will cut off at 50,000 rows without indication that more data is available. Use `TRUE` with caution.

Details

You can also use a vector of length 2 for any time queries (such as time or last_modified). The first value is the starting date (or datetime), the second value is the ending date(or datetime). NA's within the vector indicate a half-bound date. For example, `time = c("2024-01-01", NA)` will return all data starting at 2024-01-01. `time = c(NA, "2024-01-01")` will return all data from the beginning of the timeseries until 2024-01-01. By default, time is assumed UTC, although time zone attributes will be accommodated. As an example, setting `time = as.POSIXct(c("2021-01-01 12:00:00","2021-01-01 14:00"), tz = "America/New_York")` will request data that between noon and 2pm eastern time on 2021-01-01. All time values RETURNED from the service are UTC with the exception of daily data, which returns time values in local dates.

Examples

```
site <- "USGS-02238500"
field_data_sf <- read_waterdata_field_meta(monitoring_location_id = site)

groundwater <- read_waterdata_field_meta(monitoring_location_id = "USGS-375907091432201")

gwl_data <- read_waterdata_field_meta(monitoring_location_id = "USGS-02238500",
                           parameter_code = "00060",
                           begin = as.POSIXct(c(NA,
                                                "2025-08-27 12:00:00"),
                                              tz = "America/Chicago"),
                           skipGeometry = TRUE)

gwl_data_period <- read_waterdata_field_meta(
                                  monitoring_location_id = "USGS-375907091432201",
                                  parameter_code = "72019",
                                  last_modified = "P1Y")

multi_site <- read_waterdata_field_meta(
                               monitoring_location_id =  c("USGS-451605097071701",
                                                           "USGS-263819081585801"),
```

```
                              parameter_code = c("62611", "72019"))

surface_water <- read_waterdata_field_meta(
                          monitoring_location_id = c("USGS-07069000",
                                                     "USGS-07064000",
                                                     "USGS-07068000"),
                          end = "2024-07-01T00:00:00Z/..",
                          parameter_code = "00060")
```

`read_waterdata_latest_continuous` *Get Latest Continuous USGS Water Data*

Description

This endpoint provides the most recent observation for each time series of continuous data. Continuous data are collected via automated sensors installed at a monitoring location. They are collected at a high frequency and often at a fixed 15-minute interval. Depending on the specific monitoring location, the data may be transmitted automatically via telemetry and be available on WDFN within minutes of collection, while other times the delivery of data may be delayed if the monitoring location does not have the capacity to automatically transmit data. Continuous data are described by parameter name and parameter code. These data might also be referred to as "instantaneous values" or "IV"

Usage

```
read_waterdata_latest_continuous(
  monitoring_location_id = NA_character_,
  parameter_code = NA_character_,
  properties = NA_character_,
  time_series_id = NA_character_,
  approval_status = NA_character_,
  unit_of_measure = NA_character_,
  qualifier = NA_character_,
  value = NA,
  last_modified = NA_character_,
  skipGeometry = NA,
  time = NA_character_,
  bbox = NA,
  limit = NA,
  convertType = TRUE,
  no_paging = FALSE
)
```

Arguments

`monitoring_location_id`
: A unique identifier representing a single monitoring location. This corresponds to the `id` field in the `monitoring-locations` endpoint. Monitoring location IDs are created by combining the agency code of the agency responsible for the monitoring location (e.g. USGS) with the ID number of the monitoring location (e.g. 02238500), separated by a hyphen (e.g. USGS-02238500).

  Multiple monitoring_location_ids can be requested as a character vector.

`parameter_code`
: Parameter codes are 5-digit codes used to identify the constituent measured and the units of measure. A complete list of parameter codes and associated groupings can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/parameter-codes/items.

  Multiple parameter_codes can be requested as a character vector.

`properties`
: A vector of requested columns to be returned from the query. Available options are: geometry, latest_continuous_id, time_series_id, monitoring_location_id, parameter_code, statistic_id, time, value, unit_of_measure, approval_status, qualifier, last_modified. The default (`NA`) will return all columns of the data.

`time_series_id`
: A unique identifier representing a single time series. This corresponds to the `id` field in the `time-series-metadata` endpoint.

  Multiple time_series_ids can be requested as a character vector.

`approval_status`
: Some of the data that you have obtained from this U.S. Geological Survey database may not have received Director's approval. Any such data values are qualified as provisional and are subject to revision. Provisional data are released on the condition that neither the USGS nor the United States Government may be held liable for any damages resulting from its use. This field reflects the approval status of each record, and is either "Approved", meaining processing review has been completed and the data is approved for publication, or "Provisional" and subject to revision. For more information about provisional data, go to https://waterdata.usgs.gov/provisional-data-statement/.

`unit_of_measure`
: A human-readable description of the units of measurement associated with an observation.

`qualifier`
: This field indicates any qualifiers associated with an observation, for instance if a sensor may have been impacted by ice or if values were estimated.

`value`
: The value of the observation. Values are transmitted as strings in the JSON response format in order to preserve precision.

`last_modified`
: The last time a record was refreshed in our database. This may happen due to regular operational processes and does not necessarily indicate anything about the measurement has changed. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

  - A date-time: "2018-02-12T23:20:50Z"
  - A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
  - Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"

- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `last_modified` that intersects the value of datetime are selected.

See also Details below for more information.

`skipGeometry` This option can be used to skip response geometries for each feature. The returning object will be a data frame with no spatial information.

`time` The date an observation represents. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `time` that intersects the value of datetime are selected. If a feature has multiple temporal properties, it is the decision of the server whether only a single temporal property is used to determine the extent or all relevant temporal properties.

See also Details below for more information.

`bbox` Only features that have a geometry that intersects the bounding box are selected.The bounding box is provided as four or six numbers, depending on whether the coordinate reference system includes a vertical axis (height or depth). Coordinates are assumed to be in crs 4326. The expected format is a numeric vector structured: c(xmin,ymin,xmax,ymax). Another way to think of it is c(Western-most longitude, Southern-most latitude, Eastern-most longitude, Northern-most longitude).

`limit` The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (`NA`) will set the limit to the maximum allowable limit for the service.

`convertType` logical, defaults to `TRUE`. If `TRUE`, the function will convert the data to dates and qualifier to string vector.

`no_paging` logical, defaults to `FALSE`. If `TRUE`, the data will be requested from a native csv format. This can be dangerous because the data will cut off at 50,000 rows without indication that more data is available. Use `TRUE` with caution.

Details

You can also use a vector of length 2 for any time queries (such as time or last_modified). The first value is the starting date (or datetime), the second value is the ending date(or datetime). NA's within the vector indicate a half-bound date. For example, `time = c("2024-01-01", NA)` will return all data starting at 2024-01-01. `time = c(NA, "2024-01-01")` will return all data from the beginning of the timeseries until 2024-01-01. By default, time is assumed UTC, although time zone attributes will be accommodated. As an example, setting `time = as.POSIXct(c("2021-01-01`

12:00:00","2021-01-01 14:00"), tz = "America/New_York") will request data that between noon and 2pm eastern time on 2021-01-01. All time values RETURNED from the service are UTC with the exception of daily data, which returns time values in local dates.

Examples

```
site <- "USGS-451605097071701"
pcode <- "72019"
uv_data_sf <- read_waterdata_latest_continuous(monitoring_location_id = site,
                              parameter_code = pcode)

uv_data_trim <- read_waterdata_latest_continuous(monitoring_location_id = site,
                           parameter_code = pcode,
                           properties = c("monitoring_location_id",
                                          "value",
                                          "time"))

uv_data <- read_waterdata_latest_continuous(monitoring_location_id = site,
                            parameter_code = pcode,
                            skipGeometry = TRUE)

uv_data_period <- read_waterdata_latest_continuous(monitoring_location_id = site,
                                   parameter_code = pcode,
                                   time = "P7D")

multi_site <- read_waterdata_latest_continuous(monitoring_location_id = c("USGS-451605097071701",
                                                          "USGS-14181500"),
                              parameter_code = c("00060", "72019"),
                              skipGeometry = TRUE)

# Only return data that has been modified in last 7 days
multi_site2 <- read_waterdata_latest_continuous(monitoring_location_id = c("USGS-451605097071701",
                                                                           "USGS-14181500"),
                                                parameter_code = c("00060", "72019"),
                                                last_modified = "P7D")
```

---

read_waterdata_latest_daily

Get Latest USGS Daily Data

Description

Daily data provide one data value to represent water conditions for the day. Throughout much of the history of the USGS, the primary water data available was daily data collected manually at the monitoring location once each day. With improved availability of computer storage and automated

transmission of data, the daily data published today are generally a statistical summary or metric of the continuous data collected each day, such as the daily mean, minimum, or maximum value. Daily data are automatically calculated from the continuous data of the same parameter code and are described by parameter code and a statistic code. These data have also been referred to as "daily values" or "DV".

Usage

```
read_waterdata_latest_daily(
  monitoring_location_id = NA_character_,
  parameter_code = NA_character_,
  statistic_id = NA_character_,
  properties = NA_character_,
  time_series_id = NA_character_,
  approval_status = NA_character_,
  unit_of_measure = NA_character_,
  qualifier = NA_character_,
  value = NA,
  last_modified = NA_character_,
  skipGeometry = NA,
  time = NA_character_,
  bbox = NA,
  limit = NA,
  convertType = TRUE,
  no_paging = FALSE
)
```

Arguments

`monitoring_location_id`
: A unique identifier representing a single monitoring location. This corresponds to the `id` field in the `monitoring-locations` endpoint. Monitoring location IDs are created by combining the agency code of the agency responsible for the monitoring location (e.g. USGS) with the ID number of the monitoring location (e.g. 02238500), separated by a hyphen (e.g. USGS-02238500).

  Multiple monitoring_location_ids can be requested as a character vector.

`parameter_code`
: Parameter codes are 5-digit codes used to identify the constituent measured and the units of measure. A complete list of parameter codes and associated groupings can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/parameter-codes/items.

  Multiple parameter_codes can be requested as a character vector.

`statistic_id`
: A code corresponding to the statistic an observation represents. Example codes include 00001 (max), 00002 (min), and 00003 (mean). A complete list of codes and their descriptions can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/statistic-codes/items.

  Multiple statistic_ids can be requested as a character vector.

`properties`
: A vector of requested columns to be returned from the query. Available options are: geometry, latest_daily_id, time_series_id, monitoring_location_id, param-

eter_code, statistic_id, time, value, unit_of_measure, approval_status, qualifier, last_modified. The default (NA) will return all columns of the data.

`time_series_id` A unique identifier representing a single time series. This corresponds to the `id` field in the `time-series-metadata` endpoint.

Multiple time_series_ids can be requested as a character vector.

`approval_status`

Some of the data that you have obtained from this U.S. Geological Survey database may not have received Director's approval. Any such data values are qualified as provisional and are subject to revision. Provisional data are released on the condition that neither the USGS nor the United States Government may be held liable for any damages resulting from its use. This field reflects the approval status of each record, and is either "Approved", meaining processing review has been completed and the data is approved for publication, or "Provisional" and subject to revision. For more information about provisional data, go to https://waterdata.usgs.gov/provisional-data-statement/.

`unit_of_measure`

A human-readable description of the units of measurement associated with an observation.

`qualifier` This field indicates any qualifiers associated with an observation, for instance if a sensor may have been impacted by ice or if values were estimated.

`value` The value of the observation. Values are transmitted as strings in the JSON response format in order to preserve precision.

`last_modified` The last time a record was refreshed in our database. This may happen due to regular operational processes and does not necessarily indicate anything about the measurement has changed. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `last_modified` that intersects the value of datetime are selected.

See also Details below for more information.

`skipGeometry` This option can be used to skip response geometries for each feature. The returning object will be a data frame with no spatial information.

`time` The date an observation represents. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a time that intersects the value of datetime are selected. If a feature has multiple temporal properties, it is the decision of the server whether only a single temporal property is used to determine the extent or all relevant temporal properties.

See also Details below for more information.

bbox — Only features that have a geometry that intersects the bounding box are selected.The bounding box is provided as four or six numbers, depending on whether the coordinate reference system includes a vertical axis (height or depth). Coordinates are assumed to be in crs 4326. The expected format is a numeric vector structured: c(xmin,ymin,xmax,ymax). Another way to think of it is c(Western-most longitude, Southern-most latitude, Eastern-most longitude, Northern-most longitude).

limit — The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (NA) will set the limit to the maximum allowable limit for the service.

convertType — logical, defaults to TRUE. If TRUE, the function will convert the data to dates and qualifier to string vector.

no_paging — logical, defaults to FALSE. If TRUE, the data will be requested from a native csv format. This can be dangerous because the data will cut off at 50,000 rows without indication that more data is available. Use TRUE with caution.

Details

You can also use a vector of length 2 for any time queries (such as time or last_modified). The first value is the starting date (or datetime), the second value is the ending date(or datetime). NA's within the vector indicate a half-bound date. For example, `time = c("2024-01-01", NA)` will return all data starting at 2024-01-01. `time = c(NA, "2024-01-01")` will return all data from the beginning of the timeseries until 2024-01-01. By default, time is assumed UTC, although time zone attributes will be accommodated. As an example, setting `time = as.POSIXct(c("2021-01-01 12:00:00","2021-01-01 14:00"), tz = "America/New_York")` will request data that between noon and 2pm eastern time on 2021-01-01. All time values RETURNED from the service are UTC with the exception of daily data, which returns time values in local dates.

Examples

```
site <- "USGS-02238500"
pcode <- "00060"
dv_data_sf <- read_waterdata_latest_daily(monitoring_location_id = site,
                              parameter_code = pcode)

dv_data_last_modified <- read_waterdata_latest_daily(monitoring_location_id = site,
                              parameter_code = pcode,
                              last_modified = "P7D")

dv_data_trim <- read_waterdata_latest_daily(monitoring_location_id = site,
                            parameter_code = pcode,
```

```
                          properties = c("monitoring_location_id",
                                         "value",
                                         "time"))

dv_data <- read_waterdata_latest_daily(monitoring_location_id = site,
                           parameter_code = pcode,
                           skipGeometry = TRUE)

multi_site <- read_waterdata_latest_daily(monitoring_location_id =  c("USGS-01491000",
                                                                      "USGS-01645000"),
                              parameter_code = c("00060", "00010"))
```

`read_waterdata_metadata` *Generalized USGS Water Meta Data API retrieval function*

Description

Function to get metadata from Water Data API. These are useful to get the human readable words and other metadata associated with USGS codes.

Usage

```
read_waterdata_metadata(collection, limit = NA, ...)
```

Arguments

`collection`	character, can be any existing collection such as "parameter-codes", "agency-codes", "altitude-datums", "aquifer-codes", "aquifer-types", "coordinate-accuracy-codes", "coordinate-datum-codes", "coordinate-method-codes", "hydrologic-unit-codes", "medium-codes", "national-aquifer-codes", "reliability-codes", "site-types", "statistic-codes", "topographic-codes", "time-zone-codes".
`limit`	The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (`NA`) will set the limit to the maximum allowable limit for the service.
`...`	Optional arguments to pass to the query. Available parameters can be found with the `get_ogc_params` function.

Examples

```
agency_codes <- read_waterdata_metadata("agency-codes")
altitude_datums <- read_waterdata_metadata("altitude-datums")
```

```
aquifer_codes <- read_waterdata_metadata("aquifer-codes")
aquifer_types <- read_waterdata_metadata("aquifer-types")
counties <- read_waterdata_metadata("counties")
us_counties <- read_waterdata_metadata("counties", country_code = "US")
coordinate_accuracy_codes <- read_waterdata_metadata("coordinate-accuracy-codes")
coordinate_datum_codes <- read_waterdata_metadata("coordinate-datum-codes")
coordinate_method_codes <- read_waterdata_metadata("coordinate-method-codes")
huc_codes <- read_waterdata_metadata("hydrologic-unit-codes")
national_aquifer_codes <- read_waterdata_metadata("national-aquifer-codes")
parameter_codes <- read_waterdata_metadata("parameter-codes")
reliability_codes <- read_waterdata_metadata("reliability-codes")
site_types <- read_waterdata_metadata("site-types")
states <- read_waterdata_metadata("states")
us_states_territories <- read_waterdata_metadata("states", country_code = "US")
statistic_codes <- read_waterdata_metadata("statistic-codes")
topographic_codes <- read_waterdata_metadata("topographic-codes")
time_zone_codes <- read_waterdata_metadata("time-zone-codes")
time_zone_limited <- read_waterdata_metadata("time-zone-codes",
                  time_zone_description = c("Alaska", "Hawaii", "Pacific North America"))
```

`read_waterdata_monitoring_location` *Get USGS Monitoring Location Data*

Description

Location information is basic information about the monitoring location including the name, identifier, agency responsible for data collection, and the date the location was established. It also includes information about the type of location, such as stream, lake, or groundwater, and geographic information about the location, such as state, county, latitude and longitude, and hydrologic unit code (HUC).

Usage

```
read_waterdata_monitoring_location(
  monitoring_location_id = NA_character_,
  agency_code = NA_character_,
  agency_name = NA_character_,
  monitoring_location_number = NA_character_,
  monitoring_location_name = NA_character_,
  district_code = NA_character_,
  country_code = NA_character_,
  country_name = NA_character_,
  state_code = NA_character_,
  state_name = NA_character_,
  county_code = NA_character_,
  county_name = NA_character_,
```

```
  minor_civil_division_code = NA_character_,
  site_type_code = NA_character_,
  site_type = NA_character_,
  hydrologic_unit_code = NA_character_,
  basin_code = NA_character_,
  altitude = NA_character_,
  altitude_accuracy = NA_character_,
  altitude_method_code = NA_character_,
  altitude_method_name = NA_character_,
  vertical_datum = NA_character_,
  vertical_datum_name = NA_character_,
  horizontal_positional_accuracy_code = NA_character_,
  horizontal_positional_accuracy = NA_character_,
  horizontal_position_method_code = NA_character_,
  horizontal_position_method_name = NA_character_,
  original_horizontal_datum = NA_character_,
  original_horizontal_datum_name = NA_character_,
  drainage_area = NA_character_,
  contributing_drainage_area = NA_character_,
  time_zone_abbreviation = NA_character_,
  uses_daylight_savings = NA_character_,
  construction_date = NA_character_,
  aquifer_code = NA_character_,
  national_aquifer_code = NA_character_,
  aquifer_type_code = NA_character_,
  well_constructed_depth = NA_character_,
  hole_constructed_depth = NA_character_,
  depth_source_code = NA_character_,
  properties = NA_character_,
  bbox = NA,
  limit = NA,
  skipGeometry = NA
)
```

Arguments

`monitoring_location_id`
: A unique identifier representing a single monitoring location. This corresponds to the `id` field in the `monitoring-locations` endpoint. Monitoring location IDs are created by combining the agency code of the agency responsible for the monitoring location (e.g. USGS) with the ID number of the monitoring location (e.g. 02238500), separated by a hyphen (e.g. USGS-02238500).

  Multiple monitoring_location_ids can be requested as a character vector.

`agency_code`
: The agency that is reporting the data. Agency codes are fixed values assigned by the National Water Information System (NWIS). A list of agency codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/agency-codes/items.

`agency_name`
: The name of the agency that is reporting the data.

monitoring_location_number
: Each monitoring location in the USGS data base has a unique 8- to 15-digit identification number.

monitoring_location_name
: This is the official name of the monitoring location in the database. For well information this can be a district-assigned local number.

district_code
: The Water Science Centers (WSCs) across the United States use the FIPS state code as the district code. In some case, monitoring locations and samples may be managed by a water science center that is adjacent to the state in which the monitoring location actually resides. For example a monitoring location may have a district code of 30 which translates to Montana, but the state code could be 56 for Wyoming because that is where the monitoring location actually is located.

country_code
: The code for the country in which the monitoring location is located.

country_name
: The name of the country in which the monitoring location is located.

state_code
: State code. A two-digit ANSI code (formerly FIPS code) as defined by the American National Standards Institute, to define States and equivalents. A three-digit ANSI code is used to define counties and county equivalents. A lookup table is available. The only countries with political subdivisions other than the US are Mexico and Canada. The Mexican states have US state codes ranging from 81-86 and Canadian provinces have state codes ranging from 90-98.

state_name
: The name of the state or state equivalent in which the monitoring location is located.

county_code
: The code for the county or county equivalent (parish, borough, etc.) in which the monitoring location is located. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/counties/items.

county_name
: The name of the county or county equivalent (parish, borough, etc.) in which the monitoring location is located. [A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/counties/items.

minor_civil_division_code
: Codes for primary governmental or administrative divisions of the county or county equivalent in which the monitoring location is located.

site_type_code
: A code describing the hydrologic setting of the monitoring location. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/site-types/items.

site_type
: A description of the hydrologic setting of the monitoring location. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/site-types/items.

hydrologic_unit_code
: The United States is divided and sub-divided into successively smaller hydrologic units which are classified into four levels: regions, sub-regions, accounting units, and cataloging units. The hydrologic units are arranged within each other, from the smallest (cataloging units) to the largest (regions). Each hydrologic unit is identified by a unique hydrologic unit code (HUC) consisting of two to eight digits based on the four levels of classification in the hydrologic unit system.

basin_code
: The Basin Code or "drainage basin code" is a two-digit code that further subdivides the 8-digit hydrologic-unit code. The drainage basin code is defined by the USGS State Office where the monitoring location is located.

altitude
: Altitude of the monitoring location referenced to the specified Vertical Datum.

altitude_accuracy
: Accuracy of the altitude, in feet. An accuracy of +/- 0.1 foot would be entered as ".1". Many altitudes are interpolated from the contours on topographic maps; accuracies determined in this way are generally entered as one-half of the contour interval.

altitude_method_code
: Codes representing the method used to measure altitude.

altitude_method_name
: The name of the method used to measure altitude.

vertical_datum
: The datum used to determine altitude and vertical position at the monitoring location. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/altitude-datums/items.

vertical_datum_name
: The datum used to determine altitude and vertical position at the monitoring location. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/altitude-datums/items.

horizontal_positional_accuracy_code
: Indicates the accuracy of the latitude longitude values. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-accuracy-codes/items.

horizontal_positional_accuracy
: Indicates the accuracy of the latitude longitude values. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-accuracy-codes/items.

horizontal_position_method_code
: Indicates the method used to determine latitude longitude values. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-method-codes/items.

horizontal_position_method_name
: Indicates the method used to determine latitude longitude values. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-method-codes/items.

original_horizontal_datum
: Coordinates are published in EPSG:4326 / WGS84 / World Geodetic System 1984. This field indicates the original datum used to determine coordinates before they were converted. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-datum-codes/items.

original_horizontal_datum_name
: Coordinates are published in EPSG:4326 / WGS84 / World Geodetic System 1984. This field indicates the original datum used to determine coordinates before they were converted. A list of codes is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/coordinate-datum-codes/items.

drainage_area The area enclosed by a topographic divide from which direct surface runoff from precipitation normally drains by gravity into the stream above that point.

contributing_drainage_area
The contributing drainage area of a lake, stream, wetland, or estuary monitoring location, in square miles. This item should be present only if the contributing area is different from the total drainage area. This situation can occur when part of the drainage area consists of very porous soil or depressions that either allow all runoff to enter the groundwater or traps the water in ponds so that rainfall does not contribute to runoff. A transbasin diversion can also affect the total drainage area.

time_zone_abbreviation
A short code describing the time zone used by a monitoring location.

uses_daylight_savings
A flag indicating whether or not a monitoring location uses daylight savings.

construction_date
Date the well was completed.

aquifer_code Local aquifers in the USGS water resources data base are identified by a geohydrologic unit code (a three-digit number related to the age of the formation, followed by a 4 or 5 character abbreviation for the geologic unit or aquifer name).

national_aquifer_code
National aquifers are the principal aquifers or aquifer systems in the United States, defined as regionally extensive aquifers or aquifer systems that have the potential to be used as a source of potable water. Not all groundwater monitoring locations can be associated with a National Aquifer. Such monitoring locations will not be retrieved using this search criteria. A list of National aquifer codes and names is available at https://api.waterdata.usgs.gov/ogcapi/v0/collections/national-aquifer-codes/items.

aquifer_type_code
Describes the confinement status of an aquifer at the monitoring location. A confined aquifer is an aquifer below the land surface that is saturated with water. A water table–or unconfined–aquifer is an aquifer whose upper water surface (water table) is at atmospheric pressure, and thus is able to rise and fall.

well_constructed_depth
The depth of the finished well, in feet below land surface datum. Note: Not all groundwater monitoring locations have information on Well Depth. Such monitoring locations will not be retrieved using this search criteria.

hole_constructed_depth
The total depth to which the hole is drilled, in feet below land surface datum. Note: Not all groundwater monitoring locations have information on Hole Depth. Such monitoring locations will not be retrieved using this search criteria.

depth_source_code
A code indicating the source of water-level data.

properties A vector of requested columns to be returned from the query. Available options are: geometry, monitoring_location_id, agency_code, agency_name, monitoring_location_number, monitoring_location_name, district_code, country_code, country_name, state_code, state_name, county_code, county_name, minor_civil_division_code,

site_type_code, site_type, hydrologic_unit_code, basin_code, altitude, altitude_accuracy, altitude_method_code, altitude_method_name, vertical_datum, vertical_datum_name, horizontal_positional_accuracy_code, horizontal_positional_accuracy, horizontal_position_method_code, horizontal_position_method_name, original_horizontal_datum, original_horizontal_datum_name, drainage_area, contributing_drainage_area, time_zone_abbreviation, uses_daylight_savings, construction_date, aquifer_code, national_aquifer_code, aquifer_type_code, well_constructed_depth, hole_constructed_depth, depth_source_code, revision_note, revision_created, revision_modified. The default (NA) will return all columns of the data.

`bbox` Only features that have a geometry that intersects the bounding box are selected.The bounding box is provided as four or six numbers, depending on whether the coordinate reference system includes a vertical axis (height or depth). Coordinates are assumed to be in crs 4326. The expected format is a numeric vector structured: c(xmin,ymin,xmax,ymax). Another way to think of it is c(Western-most longitude, Southern-most latitude, Eastern-most longitude, Northern-most longitude).

`limit` The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (NA) will set the limit to the maximum allowable limit for the service.

`skipGeometry` This option can be used to skip response geometries for each feature. The returning object will be a data frame with no spatial information.

Examples

```
site <- "USGS-02238500"
site_info <- read_waterdata_monitoring_location(monitoring_location_id = site)

site_slim <- read_waterdata_monitoring_location(
                            monitoring_location_id = c("USGS-05114000",
                                                       "USGS-09423350"),
                            properties = c("monitoring_location_id",
                                           "state_name",
                                           "county_name",
                                           "country_name"))

site_slim_no_sf_slim <- read_waterdata_monitoring_location(state_name = "Wisconsin",
                                        properties = c("monitoring_location_id",
                                                       "state_name",
                                                       "county_name",
                                                       "country_name",
                                                       "site_type"),
                                        site_type = "Well",
                                        skipGeometry = TRUE)

site_info_no_sf <- read_waterdata_monitoring_location(
     monitoring_location_id = site_slim_no_sf_slim$monitoring_location_id[1:1000],
     skipGeometry = TRUE)
```

```
bbox_vals = c(-94.00, 35.0, -93.5, 35.5)
multi_site <- read_waterdata_monitoring_location(bbox = bbox_vals)
```

`read_waterdata_parameter_codes`
Get USGS Parameter Code Information

Description

Parameter codes are 5-digit codes and associated descriptions used to identify the constituent measured and the units of measure. Some parameter code definitions include information about the sampling matrix, fraction, and methods used to measure the constituent. Some parameters are fixed-value (fxd) numeric codes having textual meaning (for example: parameter code 00041 is a weather code parameter, code of 60 means rain), but more commonly represent a numeric value for chemical, physical, or biological data.

Usage

```
read_waterdata_parameter_codes(
  parameter_code = NA_character_,
  parameter_name = NA_character_,
  unit_of_measure = NA_character_,
  parameter_group_code = NA_character_,
  parameter_description = NA_character_,
  medium = NA_character_,
  statistical_basis = NA_character_,
  weight_basis = NA_character_,
  sample_fraction = NA_character_,
  temperature_basis = NA_character_,
  epa_equivalence = NA_character_,
  properties = NA_character_,
  limit = NA
)
```

Arguments

`parameter_code`	Parameter code.
`parameter_name`	Parameter short name.
`unit_of_measure`	Parameter reporting units defined to cooperate with descriptions by USEPA.
`parameter_group_code`	Categorical groupings of parameters by water-quality data type for display and report ordering

`parameter_description`
: Parameter description.

`medium`
: Parameter medium.

`statistical_basis`
: Parameter statistical basis.

`weight_basis`
: Parameter weight basis.

`sample_fraction`
: Parameter fraction.

`temperature_basis`
: Parameter temperature basis.

`epa_equivalence`
: Indicates the relationship of the USGS parameter code to the EPA code.

`properties`
: A vector of requested columns to be returned from the query. Available options are: geometry, parameter_code_id, parameter_name, unit_of_measure, parameter_group_code, parameter_description, medium, statistical_basis, time_basis, weight_basis, particle_size_basis, sample_fraction, temperature_basis, epa_equivalence. The default (`NA`) will return all columns of the data.

`limit`
: The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (`NA`) will set the limit to the maximum allowable limit for the service.

Examples

```
pcode <- "00060"
pcode_info <- read_waterdata_parameter_codes(parameter_code = pcode)

pcodes <- read_waterdata_parameter_codes(parameter_code = c("00660", "00060"))

# equivalent to read_waterdata_metadata("parameter-codes")
all_pcodes <- read_waterdata_parameter_codes()

total_nutrients <- read_waterdata_parameter_codes(parameter_group_code = "NUT",
                                          sample_fraction = "Total")

group_of_nutrients <- read_waterdata_parameter_codes(parameter_group_code = "NUT",
                                                     unit_of_measure = c("mg/l",
                                                     "mg/l as N", "mg/l NO3",
                                                     "mg/l asNO2"))
```

read_waterdata_samples *USGS Samples Data*

Description

This function creates the call and gets the data for discrete water quality samples data service described at https://waterdata.usgs.gov/download-samples/.

Usage

```
read_waterdata_samples(
  monitoringLocationIdentifier = NA,
  siteTypeCode = NA,
  boundingBox = NA,
  hydrologicUnit = NA,
  activityMediaName = NA,
  characteristicGroup = NA,
  characteristic = NA,
  characteristicUserSupplied = NA,
  activityStartDateLower = NA,
  activityStartDateUpper = NA,
  countryFips = NA,
  stateFips = NA,
  countyFips = NA,
  projectIdentifier = NA,
  recordIdentifierUserSupplied = NA,
  siteTypeName = NA,
  usgsPCode = NA,
  pointLocationLatitude = NA,
  pointLocationLongitude = NA,
  pointLocationWithinMiles = NA,
  dataType = "results",
  dataProfile = NA,
  tz = "UTC",
  convertType = TRUE
)

read_USGS_samples(
  monitoringLocationIdentifier = NA,
  siteTypeCode = NA,
  boundingBox = NA,
  hydrologicUnit = NA,
  activityMediaName = NA,
  characteristicGroup = NA,
  characteristic = NA,
  characteristicUserSupplied = NA,
```

```
  activityStartDateLower = NA,
  activityStartDateUpper = NA,
  countryFips = NA,
  stateFips = NA,
  countyFips = NA,
  projectIdentifier = NA,
  recordIdentifierUserSupplied = NA,
  siteTypeName = NA,
  usgsPCode = NA,
  pointLocationLatitude = NA,
  pointLocationLongitude = NA,
  pointLocationWithinMiles = NA,
  dataType = "results",
  dataProfile = NA,
  tz = "UTC",
  convertType = TRUE
)
```

Arguments

`monitoringLocationIdentifier`
: A monitoring location identifier has two parts: the agency code and the location number, separated by a dash (-). Location identifiers should be separated with commas, for example: AZ014-320821110580701, CAX01-15304600, USGS-040851385. Location numbers without an agency prefix are assumed to have the prefix USGS.

`siteTypeCode`
: Site type code query parameter. See available options by running `check_waterdata_sample_params("s`

`boundingBox`
: North and South are latitude values; East and West are longitude values. A vector of 4 (west, south, east, north) is expected. An example would be: c(-92.8, 44.2, -88.9, 46.0).

`hydrologicUnit`
: Hydrologic Unit Codes (HUCs) identify physical areas within the US that drain to a certain portion of the stream network. This filter accepts values containing 2, 4, 6, 8, 10 or 12 digits.

`activityMediaName`
: Sample media refers to the environmental medium that was sampled or analyzed.

`characteristicGroup`
: Characteristic group is a broad category describing the sample. See available options by running `check_waterdata_sample_params("characteristicgroup")$characteristicGrou`

`characteristic`
: Characteristic is a specific category describing the sample. See available options by running `check_waterdata_sample_params("characteristics")$characteristicName`.

`characteristicUserSupplied`
: Observed property is the USGS term for the constituent sampled and the property name gives a detailed description of what was sampled. Observed property is mapped to characteristicUserSupplied and replaces the parameter name and pcode USGS previously used to describe discrete sample data. Find more information in the Observed Properties and Parameter Codes section of the Code Dictionary found here: https://waterdata.usgs.gov/code-dictionary/.

activityStartDateLower
: The service will return records with dates earlier than the value entered for activityStartDateUpper. Can be an R Date object, or a string with format YYYY-MM-DD. The logic is inclusive, i.e. it will also return records that match the date.

activityStartDateUpper
: The service will return records with dates later than the value entered for activityStartDateLower. Can be an R Date object, or a string with format YYYY-MM-DD. The logic is inclusive, i.e. it will also return records that match the date.

countryFips
: Country query parameter. Do not set redundant parameters. If another query parameter contains the country information, leave this parameter set to the default NA. See available options by running `check_waterdata_sample_params("countries")`, where the "id" field contains the value to use in the countryFips input.

stateFips
: State query parameter. To get a list of available state fips, run `check_waterdata_sample_params("state`
The "fips" can be created using the function `stateCdLookup` - for example: `stateCdLookup("WI", "fips")`. FIPs codes for states take the format: CountryAbbrev:StateNumber, like US:55 for Wisconsin.

countyFips
: County query parameter. To get a list of available counties, run `check_waterdata_sample_params("cou`
The "Fips" can be created using the function `countyCdLookup` - for example: `countyCdLookup("WI", "Dane", "fips")` for Dane County, WI. FIPs codes for counties take the format: CountryAbbrev:StateNumber:CountyNumber, like US:55:025 for Dane County, WI.

projectIdentifier
: Project identifier query parameter. This information would be needed from prior project information.

recordIdentifierUserSupplied
: Record identifier, user supplied identifier. This information would be needed from the data supplier.

siteTypeName
: Site type name query parameter. See available options by running `check_param("sitetype")$typeName`

usgsPCode
: USGS parameter code. See available options by running `check_waterdata_sample_params("characte`

pointLocationLatitude
: Latitude for a point/radius query (decimal degrees). Must be used with pointLocationLongitude and pointLocationWithinMiles.

pointLocationLongitude
: Longitude for a point/radius query (decimal degrees). Must be used with pointLocationLatitude and pointLocationWithinMiles.

pointLocationWithinMiles
: Radius for a point/radius query. Must be used with pointLocationLatitude and pointLocationLongitude

dataType
: Options include: "Results", "Monitoring locations", "Activities", "Projects", and "Organizations".

dataProfile
: Profile depends on type. Options for "results" dataType are: "fullphyschem", "basicphyschem", "fullbio", "basicbio", "narrow", "resultdetectionquantitationlimit", "labsampleprep", "count". Options for "locations" are: "site" and "count".

	Options for "activities" are "sampact", "actmetric", "actgroup", and "count". Options for "projects" are: "project" and "projectmonitoringlocationweight". Options for "organizations" are: "organization" and "count".
tz	character to set timezone attribute of datetime. Default is UTC (properly accounting for daylight savings times based on the data's provided tz_cd column). Possible values include "America/New_York","America/Chicago", "America/Denver","America/Los_An "America/Anchorage","America/Honolulu","America/Jamaica","America/Managua", "America/Phoenix", and "America/Metlakatla"
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates, datetimes, numerics based on a standard algorithm. If false, everything is returned as a character.

Examples

```
ph_data <- read_waterdata_samples(
               monitoringLocationIdentifier = "USGS-04074950",
               characteristicUserSupplied = "pH, water, unfiltered, field",
               activityStartDateUpper = "2000-01-01",
               dataProfile = "narrow")

nameToUse <- "pH"
pHData <- read_waterdata_samples(monitoringLocationIdentifier = "USGS-04024315",
                          characteristic = nameToUse)
ncol(pHData)
attr(pHData, "url")
attr(pHData, "queryTime")

summary_data <- read_waterdata_samples(monitoringLocationIdentifier = "USGS-04024315",
                                dataType = "projects")
```

---

read_waterdata_stats_por

Get USGS daily data statistics

Description

This service provides endpoints for access to computations on the historical record regarding water conditions. For more information regarding the calculation of statistics and other details, please visit the Statistics documentation page.

Note: The /statistics API is under active beta development and subject to change. Improved handling of significant figures will be addressed in a future release.

`read_waterdata_stats_por` Returns day-of-year and month-of-year statistics matching your query.

`read_waterdata_stats_daterange` Returns monthly and annual statistics matching your query.

Usage

```
read_waterdata_stats_por(
  approval_status = NA,
  computation_type = NA_character_,
  country_code = NA_character_,
  state_code = NA_character_,
  county_code = NA_character_,
  start_date = NA_character_,
  end_date = NA_character_,
  monitoring_location_id = NA_character_,
  parent_time_series_id = NA_character_,
  site_type_code = NA_character_,
  site_type_name = NA_character_,
  parameter_code = NA_character_,
  page_size = NA
)

read_waterdata_stats_daterange(
  approval_status = NA,
  computation_type = NA_character_,
  country_code = NA_character_,
  state_code = NA_character_,
  county_code = NA_character_,
  start_date = NA_character_,
  end_date = NA_character_,
  monitoring_location_id = NA_character_,
  parent_time_series_id = NA_character_,
  site_type_code = NA_character_,
  site_type_name = NA_character_,
  parameter_code = NA_character_,
  page_size = NA
)
```

Arguments

approval_status	Whether to include approved and/or provisional observations. At this time, only approved observations are returned.
computation_type	Desired statistical computation method. Available values: "arithmetic_mean", "maximum", "median", "minimum", "percentile".
country_code	Country Query Parameter. Accepts multiple values (see examples). If one of country, county, or state code is supplied then the other two arguments do not need to be specified.
state_code	State Query Parameter. Accepts multiple values in a character vector.
county_code	County Query Parameter. Accepts multiple values in a character vector.

start_date Start Date Query Parameter. The logic is inclusive i.e., it will also return records that match the date. If an end date is supplied, but no start date is supplied, then statistics will be supplied for the entire period of record ending with the end date. If an end date is not supplied, but a start date is supplied then statistics will be supplied for the period of record following the start date. If no start or end date are supplied then statistics will be supplied for the entire period of record.

end_date End Date Query Parameter. The logic is inclusive i.e., it will also return records that match the date.

monitoring_location_id Each monitoring location has been assigned a unique station number that places them in downstream order. Accepts multiple values in a character vector.

parent_time_series_id The parent_time_series_id returns statistics tied to a particular database entry. Accepts multiple values in a character vector. If no parent time series identifier is supplied, then all records matching the rest of the provided criteria will be returned.

site_type_code Site Type Code Query Parameter. Accepts multiple values in a character vector. If no Site Type code is specified, statistics of all site types with the matching Monitoring Location Identifier will be returned.

site_type_name Site Type Name Query Parameter. If no Site Type name is specified, statistics of all site types with the matching Monitoring Location Identifier will be returned.

parameter_code USGS Parameter Code Query Parameter. Accepts multiple values in a character vector. If no USGS parameter code is specified, but a Monitoring Location Identifier is supplied, then all statistics and their parameter codes with a matching monitoring location identifier will be returned. All statistics within the period of record will be returned if no parameter code or monitoring location identifier are specified.

page_size Return a defined number of results (default: 1000).

See Also

https://api.waterdata.usgs.gov/statistics/v0/docs

Examples

```
# All day-of-year and month-of-year statistics for two sites
x1 <- read_waterdata_stats_por(
  monitoring_location_id = c("USGS-02319394", "USGS-02171500")
)

# Request temperature percentiles for specific month-day range
# Returns:
# - Day-of-year temperature percentiles for each day between June 1 through June 15.
# - Month-of-year percentiles for June, computed using
#    all June data (not just June 1 through June 15).
# Note: the month-of-year percentiles are returned only when the month-day range includes
```

```
#  the beginning of the month (e.g., "06-01")
x2 <- read_waterdata_stats_por(
  monitoring_location_id = c("USGS-02319394", "USGS-02171500"),
  parameter_code = "00010",
  start_date = "06-01", end_date = "06-15",
  computation_type = "percentile"
)

# All calendar month, calendar year, and water year statistics for two sites
x3 <- read_waterdata_stats_daterange(
  monitoring_location_id = c("USGS-02319394", "USGS-02171500")
)

# Request specific gage height and discharge summaries for a limited date range
# Returns:
# - calendar month summaries for each month between January, 2010 through December, 2011
# - calendar year summaries for 2010 and 2011
# - water year summaries for WY2010, WY2011, and WY2012
x4 <- read_waterdata_stats_daterange(
  monitoring_location_id = c("USGS-02319394", "USGS-02171500"),
  parameter_code = c("00065", "00060"),
  start_date = "2010-01-01", end_date = "2011-12-31",
  computation_type = c("minimum", "median", "maximum")
)
```

`read_waterdata_ts_meta`

Get USGS Time Series Metadata

Description

Daily data and continuous measurements are grouped into time series, which represent a collection of observations of a single parameter, potentially aggregated using a standard statistic, at a single monitoring location. This endpoint provides metadata about those time series, including their operational thresholds, units of measurement, and when the earliest and most recent observations in a time series occurred.

Usage

```
read_waterdata_ts_meta(
  monitoring_location_id = NA_character_,
  parameter_code = NA_character_,
  parameter_name = NA_character_,
  properties = NA_character_,
  statistic_id = NA_character_,
  last_modified = NA_character_,
```

```
  begin_utc = NA_character_,
  end_utc = NA_character_,
  hydrologic_unit_code = NA_character_,
  state_name = NA_character_,
  unit_of_measure = NA_character_,
  computation_period_identifier = NA_character_,
  computation_identifier = NA_character_,
  thresholds = NA,
  sublocation_identifier = NA_character_,
  primary = NA_character_,
  parent_time_series_id = NA_character_,
  time_series_id = NA_character_,
  web_description = NA_character_,
  skipGeometry = NA,
  limit = NA,
  max_results = NA,
  bbox = NA,
  begin = NA_character_,
  end = NA_character_,
  convertType = TRUE,
  no_paging = FALSE
)
```

Arguments

`monitoring_location_id`
: A unique identifier representing a single monitoring location. This corresponds to the `id` field in the `monitoring-locations` endpoint. Monitoring location IDs are created by combining the agency code of the agency responsible for the monitoring location (e.g. USGS) with the ID number of the monitoring location (e.g. 02238500), separated by a hyphen (e.g. USGS-02238500).

  Multiple monitoring_location_ids can be requested as a character vector.

`parameter_code`
: Parameter codes are 5-digit codes used to identify the constituent measured and the units of measure. A complete list of parameter codes and associated groupings can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/parameter-codes/items.

  Multiple parameter_codes can be requested as a character vector.

`parameter_name`
: A human-understandable name corresponding to `parameter_code`.

`properties`
: A vector of requested columns to be returned from the query. Available options are: geometry, time_series_id, unit_of_measure, parameter_name, parameter_code, statistic_id, hydrologic_unit_code, state_name, last_modified, begin, end, begin_utc, end_utc, computation_period_identifier, computation_identifier, thresholds, sublocation_identifier, primary, monitoring_location_id, web_description, parameter_description, parent_time_series_id. The default (`NA`) will return all columns of the data.

`statistic_id`
: A code corresponding to the statistic an observation represents. Example codes include 00001 (max), 00002 (min), and 00003 (mean). A complete list of codes

and their descriptions can be found at https://api.waterdata.usgs.gov/ogcapi/v0/collections/statistic-codes/items.

Multiple statistic_ids can be requested as a character vector.

`last_modified` The last time a record was refreshed in our database. This may happen due to regular operational processes and does not necessarily indicate anything about the measurement has changed. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `last_modified` that intersects the value of datetime are selected.

See also Details below for more information.

`begin_utc` The datetime of the earliest observation in the time series. Together with `end`, this field represents the period of record of a time series. Note that some time series may have large gaps in their collection record. You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `begin` that intersects the value of datetime are selected.

#' See also Details below for more information.

`end_utc` The datetime of the most recent observation in the time series. Data returned by this endpoint updates at most once per day, and potentially less frequently than that, and as such there may be more recent observations within a time series than the time series `end` value reflects. Together with `begin`, this field represents the period of record of a time series. It is additionally used to determine whether a time series is "active". You can query this field using date-times or intervals, adhering to RFC 3339, or using ISO 8601 duration objects. Intervals may be bounded or half-bounded (double-dots at start or end). Examples:

- A date-time: "2018-02-12T23:20:50Z"
- A bounded interval: "2018-02-12T00:00:00Z/2018-03-18T12:31:12Z"
- Half-bounded intervals: "2018-02-12T00:00:00Z/.." or "../2018-03-18T12:31:12Z"
- Duration objects: "P1M" for data from the past month or "PT36H" for the last 36 hours

Only features that have a `end` that intersects the value of datetime are selected.

See also Details below for more information.

`hydrologic_unit_code`
: The United States is divided and sub-divided into successively smaller hydrologic units which are classified into four levels: regions, sub-regions, accounting units, and cataloging units. The hydrologic units are arranged within each other, from the smallest (cataloging units) to the largest (regions). Each hydrologic unit is identified by a unique hydrologic unit code (HUC) consisting of two to eight digits based on the four levels of classification in the hydrologic unit system.

`state_name`
: The name of the state or state equivalent in which the monitoring location is located.

`unit_of_measure`
: A human-readable description of the units of measurement associated with an observation.

`computation_period_identifier`
: Indicates the period of data used for any statistical computations. Multiple computation_period_identifiers can be requested as a character vector.

`computation_identifier`
: Indicates whether the data from this time series represent a specific statistical computation. Multiple computation_identifiers can be requested as a character vector.

`thresholds`
: Thresholds represent known numeric limits for a time series, for example the historic maximum value for a parameter or a level below which a sensor is non-operative. These thresholds are sometimes used to automatically determine if an observation is erroneous due to sensor error, and therefore shouldn't be included in the time series.

`sublocation_identifier`
: An optional human-readable identifier used to specify where measurements are recorded at a monitoring location.

`primary`
: A flag identifying if the time series is a "primary" time series. "Primary" time series (which have this flag) are standard observations which undergo Bureau review and approval processes. Non-primary time series, which will have missing values for "primary", are provisional datasets made available to meet the need for timely best science and to assist with daily operations which need real-time information. Non-primary time series data are only retained by this system for 120 days.

`parent_time_series_id`
: The unique identifier representing the parent or "upchain" time series that a daily values time series is generated from. Daily values time series have one and only one parent time series.

`time_series_id`
: A unique identifier representing a single time series. This corresponds to the `id` field in the `time-series-metadata` endpoint.

`web_description`
: A description of what this time series represents, as used by WDFN and other USGS data dissemination products.

`skipGeometry`
: This option can be used to skip response geometries for each feature. The returning object will be a data frame with no spatial information.

limit	The optional limit parameter is used to control the subset of the selected features that should be returned in each page. The maximum allowable limit is 50000. It may be beneficial to set this number lower if your internet connection is spotty. The default (NA) will set the limit to the maximum allowable limit for the service.
max_results	The optional maximum number of rows to return. This value must be less than the requested limit.
bbox	Only features that have a geometry that intersects the bounding box are selected.The bounding box is provided as four or six numbers, depending on whether the coordinate reference system includes a vertical axis (height or depth). Coordinates are assumed to be in crs 4326. The expected format is a numeric vector structured: c(xmin,ymin,xmax,ymax). Another way to think of it is c(Western-most longitude, Southern-most latitude, Eastern-most longitude, Northern-most longitude).
begin	This field contains the same information as "begin_utc", but in the local time of the monitoring location. It is retained for backwards compatibility, but will be removed in V1 of these APIs.
end	This field contains the same information as "end_utc", but in the local time of the monitoring location. It is retained for backwards compatibility, but will be removed in V1 of these APIs.
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates and qualifier to string vector.
no_paging	logical, defaults to FALSE. If TRUE, the data will be requested from a native csv format. This can be dangerous because the data will cut off at 50,000 rows without indication that more data is available. Use TRUE with caution.

Details

You can also use a vector of length 2 for any time queries (such as time or last_modified). The first value is the starting date (or datetime), the second value is the ending date(or datetime). NA's within the vector indicate a half-bound date. For example, `time = c("2024-01-01", NA)` will return all data starting at 2024-01-01. `time = c(NA, "2024-01-01")` will return all data from the beginning of the timeseries until 2024-01-01. By default, time is assumed UTC, although time zone attributes will be accommodated. As an example, setting `time = as.POSIXct(c("2021-01-01 12:00:00","2021-01-01 14:00"), tz = "America/New_York")` will request data that between noon and 2pm eastern time on 2021-01-01. All time values RETURNED from the service are UTC with the exception of daily data, which returns time values in local dates.

Examples

```
site <- "USGS-02238500"
meta_1 <- read_waterdata_ts_meta(monitoring_location_id = site)

meta_multi <- read_waterdata_ts_meta(monitoring_location_id =  c("USGS-01491000",
                                                                 "USGS-01645000"),
                              parameter_code = c("00060", "00010"),
                              properties = c("monitoring_location_id",
```

```
                                          "parameter_code",
                                          "begin_utc",
                                          "end_utc",
                                          "time_series_id"),
                              skipGeometry = TRUE)

meta_wi <- read_waterdata_ts_meta(state_name = "Wisconsin")
```

---

renameNWISColumns *renameColumns*

Description

Rename columns coming back from NWIS data retrievals. Daily and unit value columns have names derived from their data descriptor, parameter, and statistic codes. This function reads information from the header and the arguments in the call to to rename those columns.

Usage

```
renameNWISColumns(
  rawData,
  p00010 = "Wtemp",
  p00045 = "Precip",
  p00060 = "Flow",
  p00065 = "GH",
  p00095 = "SpecCond",
  p00300 = "DO",
  p00400 = "pH",
  p62611 = "GWL",
  p63680 = "Turb",
  p72019 = "WLBLS",
  ...
)
```

Arguments

rawData	the daily- or unit-values datset retrieved from NWISweb.
p00010	the base name for parameter code 00010.
p00045	the base name for parameter code 00045.
p00060	the base name for parameter code 00060.
p00065	the base name for parameter code 00065.
p00095	the base name for parameter code 00095.
p00300	the base name for parameter code 00300.
p00400	the base name for parameter code 00400.

p62611	the base name for parameter code 62611.
p63680	the base name for parameter code 63680.
p72019	the base name for parameter code 72019.
...	named arguments for the base name for any other parameter code. The form of the name must be like pXXXXX, where XXXXX is the parameter code.

Value

A dataset like data with selected columns renamed.

Note

The following statistics codes are converted by renameNWISColumns.

00000 Instantaneous Value, suffix: Inst

00001 Maximum value, suffix: Max

00002 Minimum value, suffix: Min

00003 Mean value, no suffix

00006 Sum of values, suffix: Sum

00007 Modal value, suffix: Mode

00008 Median value, suffix: Median

00012 Equivalent mean value, suffix: EqMean

00021 Tidal high-high value, suffix: HiHiTide

00022 Tidal low-high value, suffix: LoHiTide

00023 Tidal high-low value, suffix: HiLoTide

00024 Tidal low-low value, suffix: LoLoTide

See Also

readNWISdv(), readNWISuv()

Examples

```
siteWithTwo <- "01480015"
startDate <- "2012-09-01"
endDate <- "2012-10-01"

twoResults <- readNWISdv(siteWithTwo, "00060", startDate, endDate)
names(twoResults)
renamedCols <- renameNWISColumns(twoResults)
names(renamedCols)
# Custom names:
newNames <- renameNWISColumns(twoResults, p00060 = "Discharge")
names(newNames)
```

setAccess *Set data endpoint*

Description

access Indicate which dataRetrieval access code you want to use options: `c('public','internal')`

Usage

```
setAccess(access = "public")
```

Arguments

access	code for data access. Options are: "public","internal","cooperator", or "USGS". • "internal" represents Access=3 ...for a single water science center • "USGS" represents Access=2 ...for all water science centers • "cooperator" represents Access=1 • "public" represents Access=0, public access

Author(s)

Luke Winslow, Jordan S Read

Examples

```
setAccess("internal")

setAccess("public")
```

stateCd *US State Code Lookup Table*

Description

Classic lookup table for states. Has been replaced in functions with `check_waterdata_sample_params("states")`.

Value

stateCd data frame.

Name	Type	Description
STATE	character	FIPS State Code
STUSAB	character	Official United States Postal Service (USPS) Code
STATE_NAME	character	State Name
STATENS	character	Geographic Names Information System Identifier (GNISID)

Examples

```
head(stateCd)
```

---

stateCdLookup *State code look up*

Description

Function to simplify finding state and state code definitions. Used in `readNWISdata` and `readWQPdata`.

Usage

```
stateCdLookup(input, outputType = "postal", country = "US")
```

Arguments

`input`	could be character (full name, abbreviation, id), or numeric (id)
`outputType`	character can be "postal", "fullName", "tableIndex", or "id".
`country`	description

Examples

```
fullName <- stateCdLookup("wi", "fullName")
abbriev <- stateCdLookup("Wisconsin", "postal")
id <- stateCdLookup("WI", "id")
name <- stateCdLookup(55, "fullName")
fips <- stateCdLookup("WI", "fips")
canada_st <- stateCdLookup(13, "fullName", country = "CA")
mexico_st <- stateCdLookup(13, "fullName", country = "MX")
stateCdLookup(c("West Virginia", "Wisconsin", 200, 55, "MN"))
```

summarize_waterdata_samples *USGS Samples Summary Data*

Description

This function creates the call and gets the data for discrete water quality samples summary data service described at https://api.waterdata.usgs.gov/samples-data/docs.

Usage

```
summarize_waterdata_samples(monitoringLocationIdentifier)

summarize_USGS_samples(monitoringLocationIdentifier)
```

Arguments

monitoringLocationIdentifier	A monitoring location identifier has two parts, separated by a dash (-): the agency code and the location number. Location identifiers should be separated with commas, for example: AZ014-320821110580701, CAX01-15304600, USGS-040851385. Location numbers without an agency prefix are assumed to have the prefix USGS.

Value

data frame with summary of data available based on the monitoringLocationIdentifier

Examples

```
monitoringLocationIdentifier <- "USGS-04074950"

what_data <- summarize_waterdata_samples(monitoringLocationIdentifier)
```

whatNWISdata *USGS data availability*

Description

Imports a table of available parameters, period of record, and count. See https://waterservices.usgs.gov/docs/site-service/ for more information.

Usage

```
whatNWISdata(..., convertType = TRUE)
```

Arguments

...	see https://waterservices.usgs.gov/docs/site-service/ for a complete list of options. A list of arguments can also be supplied.
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates, datetimes, numerics based on a standard algorithm. If false, everything is returned as a character

Details

This function requires users to create their own arguments based on the NWIS web services. It is a more complicated function to use compared to other NWIS functions such as readNWISdv(), readNWISuv(), etc. However, this function adds a lot of flexibility to the possible queries. If the "service" argument is included, the results will be filtered to the proper data_type_cd. This is a great function to use before a large data set, by filtering down the number of sites that have useful data.

Value

A data frame with the following columns:

Name	Type	Description
agency_cd	character	The NWIS code for the agency reporting the data
site_no	character	The USGS site number
station_nm	character	Site name
site_tp_cd	character	Site type
dec_lat_va	numeric	Decimal latitude
dec_long_va	numeric	Decimal longitude
coord_acy_cd	character	Latitude-longitude accuracy
dec_coord_datum_cd	character	Decimal Latitude-longitude datum
alt_va	character	Altitude of Gage or land surface
alt_acy_va	character	Altitude accuracy
alt_datum_cd	character	Altitude datum
huc_cd	character	Hydrologic unit code
data_type_cd	character	Data type
parm_cd	character	Parameter code
stat_cd	character	Statistical code
dd_nu	character	Internal database key
loc_web_ds	character	Additional measurement description
medium_grp_cd	character	Medium group code
parm_grp_cd	character	Parameter group code
srs_id	character	SRS ID
access_cd	character	Access code
begin_date	Date	Begin date
end_date	Date	End date
count_nu	integer	Record count

parameter_group_nm	character	Parameter group name
parameter_nm	character	Parameter name
casrn	character	Chemical Abstracts Service (CAS) Registry Number
srsname	character	Substance Registry Services
parameter_units	character	Parameter units

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
comment	character	Header comments from the RDB file
queryTime	POSIXct	The time the data was returned

See Also

read_waterdata_ts_meta()

Examples

```
# see ?read_waterdata_ts_meta

#site1 <- whatWQPsamples(siteid = "USGS-01594440")

#type <- "Stream"

#sites <- whatWQPsamples(countycode = "US:55:025", siteType = type)

#lakeSites_samples <- whatWQPsamples(siteType = "Lake, Reservoir, Impoundment",
#                                    countycode = "US:55:025")
```

---

whatNWISsites *Site Data Import from NWIS*

Description

Returns a list of sites from the NWIS web service. This function gets the data from: https://waterservices.usgs.gov/docs/site-service/. Mapper format is used

Usage

```
whatNWISsites(...)
```

Arguments

...	see https://waterservices.usgs.gov/docs/site-service/ for a complete list of options. A list (or lists) can also be supplied.

Value

A data frame with at least the following columns:

Name	Type	Description
agency_cd	character	The NWIS code for the agency reporting the data
site_no	character	The USGS site number
station_nm	character	Station name
site_tp_cd	character	Site type code
dec_lat_va	numeric	Decimal latitude
dec_long_va	numeric	Decimal longitude
queryTime	POSIXct	Query time

There are also several useful attributes attached to the data frame:

Name	Type	Description
url	character	The url used to generate the data
queryTime	POSIXct	The time the data was returned

See Also

read_waterdata_monitoring_location()

Examples

```
# see ?read_waterdata_monitoring_location
#siteListPhos <- whatNWISsites(stateCd = "OH", parameterCd = "00665")
#oneSite <- whatNWISsites(sites = "05114000")
```

---

whatWQPdata *Data Available from Water Quality Portal*

Description

Returns a list of sites from the Water Quality Portal web service. This function gets the data from: https://www.waterqualitydata.us. Arguments to the function should be based on https://www.waterqualitydata.us/webservices_documentation. The information returned from whatWQPdata describes the available data at the WQP sites, and some metadata on the sites themselves. For example, a row is returned for each individual site that fulfills this query. In that we can learn how many sampling activities and results are available for the query. It does not break those results down by any finer grain. For example, if you ask for "Nutrients" (characteristicGroup), you will not learn what specific nutrients are available at that site. For that kind of data discovery see readWQPsummary.

Usage

```
whatWQPdata(..., convertType = TRUE)
```

Arguments

...	see https://www.waterqualitydata.us/webservices_documentation for a complete list of options. A list of arguments can also be supplied. One way to figure out how to construct a WQP query is to go to the "Advanced" form in the Water Quality Portal: https://www.waterqualitydata.us/#mimeType=csv&providers=NWIS&providers=STORET Use the form to discover what parameters are available. Once the query is set in the form, scroll down to the "Query URL". You will see the parameters after "https://www.waterqualitydata.us/#". For example, if you chose "Nutrient" in the Characteristic Group dropdown, you will see characteristicType=Nutrient in the Query URL. The corresponding argument for dataRetrieval is characteristicType = "Nutrient". dataRetrieval users do not need to include mimeType, and providers is optional (these arguments are picked automatically).
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates, datetimes, numerics based on a standard algorithm. If false, everything is returned as a character.

Value

A data frame that returns basic data availability such as sites, number of results, and number of sampling activities from the query parameters for the Water Quality Portal.

See Also

whatWQPsites readWQPsummary readWQPdata

Examples

```
site1 <- whatWQPdata(siteid = "USGS-01594440")

type <- "Stream"
sites <- whatWQPdata(countycode = "US:55:025", siteType = type)

lakeSites <- whatWQPdata(siteType = "Lake, Reservoir, Impoundment",
                         countycode = "US:55:025")
lakeSites_chars <- whatWQPdata(
  siteType = "Lake, Reservoir, Impoundment",
  countycode = "US:55:025", convertType = FALSE)


bbox <- c(-86.9736, 34.4883, -86.6135, 34.6562)
what_bb <- whatWQPdata(bBox = bbox)
```

whatWQPsamples *Site Data Import from Water Quality Portal*

Description

Returns a list of sites from the Water Quality Portal web service. This function gets the data from: https://www.waterqualitydata.us. Arguments to the function should be based on https://www.waterqualitydata.us/webservices_documentation. The return from this function returns the basic metadata on WQP sites. It is generally faster than the whatWQPdata() function, but does not return information on what data was collected at the site.

Usage

```
whatWQPsamples(..., convertType = TRUE, legacy = TRUE)

whatWQPmetrics(..., convertType = TRUE)

whatWQPsites(..., legacy = TRUE, convertType = TRUE)
```

Arguments

...	see https://www.waterqualitydata.us/webservices_documentation for a complete list of options. A list of arguments can also be supplied. One way to figure out how to construct a WQP query is to go to the "Advanced" form in the Water Quality Portal: https://www.waterqualitydata.us/#mimeType=csv&providers=NWIS&providers=STORET Use the form to discover what parameters are available. Once the query is set in the form, scroll down to the "Query URL". You will see the parameters after "https://www.waterqualitydata.us/#". For example, if you chose "Nutrient" in the Characteristic Group dropdown, you will see characteristicType=Nutrient in the Query URL. The corresponding argument for dataRetrieval is characteristicType = "Nutrient". dataRetrieval users do not need to include mimeType, and providers is optional (these arguments are picked automatically).
convertType	logical, defaults to TRUE. If TRUE, the function will convert the data to dates, datetimes, numerics based on a standard algorithm. If false, everything is returned as a character.
legacy	Logical. If TRUE, uses legacy WQP services. Default is TRUE. Setting legacy = FALSE uses WQX3.0 WQP services, which are in-development, use with caution.

Value

A data frame with information on the sampling activity available from the Water Quality Portal for the query parameters.

data frame that includes information on site metadata.

See Also

whatWQPdata readWQPsummary

whatNWISdata

Examples

```
site1 <- whatWQPsamples(siteid = "USGS-01594440")

type <- "Stream"

sites <- whatWQPsamples(countycode = "US:55:025", siteType = type)

lakeSites_samples <- whatWQPsamples(siteType = "Lake, Reservoir, Impoundment",
                                    countycode = "US:55:025")



type <- "Stream"

sites <- whatWQPmetrics(countycode = "US:55:025", siteType = type)
lakeSites_metrics <- whatWQPmetrics(siteType = "Lake, Reservoir, Impoundment",
                                    countycode = "US:55:025")




site1 <- whatWQPsites(siteid = "USGS-01594440")

type <- "Stream"
sites <- whatWQPsites(
  countycode = "US:55:025",
  characteristicName = "Phosphorus",
  siteType = type
)
```

---

wqp_check_status *Get WQP service metadata*

Description

The information from this request is only available for a limited time after the original query from the WQP. In the readWQPdata and readWQPqw functions, the results from this function will be attached as an attribute to the data.

Usage

```
wqp_check_status(wqp_request_id)
```

Arguments

wqp_request_id A character returned from the header of a WQP request.

Value

a list generated from the WQP describing what data was returned.

Examples

```
rawPcode <- readWQPqw("USGS-01594440", "01075",
                      ignore_attributes = TRUE, legacy = FALSE)
headerInfo <- attr(rawPcode, "headerInfo")
wqp_request_id <- headerInfo$`wqp-request-id`
count_info <- wqp_check_status(wqp_request_id)
count_info[["dataProviders"]]
```

---

zeroPad *Pad string with leading zeros*

Description

Function to pad a string with leading zeros. Useful for parameter codes and USGS site IDs.

Usage

```
zeroPad(x, padTo)
```

Arguments

x	character
padTo	number Final desired length of the character

Value

x character returned with leading zeros

Examples

```
pCode <- "10"
correctPCode <- zeroPad(pCode, 5)
pCodes <- c("100", "1000", "0", "12345", "1565465465465465")
correctPCodes <- zeroPad(pCodes, 5)
pCodeNA <- c(1, 2, NA)
padPCodeNA <- zeroPad(pCodeNA, 4)
```

Index