Package ‘inldata’

April 15, 2024

Title  Collection of Datasets for the USGS-INL Monitoring Networks

Version  1.1.5

Description  A collection of analysis-ready datasets for the
U.S. Geological Survey - Idaho National Laboratory (USGS-INL)
groundwater and surface-water monitoring networks, administered by the
USGS-INL Project Office in cooperation with the U.S. Department of Energy.
The data collected from wells and surface-water stations at the
Idaho National Laboratory and surrounding areas have been used to describe
the effects of waste disposal on water contained in the
eastern Snake River Plain aquifer, located in the southeastern part of
Idaho, and the availability of water for long-term consumptive and
industrial use. The package includes long-term monitoring records dating
back to measurements from 1949. Geospatial data describing the areas from
which samples were collected or observations were made are also included in
the package. Bundling this data into a single package significantly reduces
the magnitude of data processing for researchers and provides a way to
distribute the data along with its documentation in a standard format.
Geospatial datasets are made available in a common projection and datum, and
geo-hydrologic data have been structured to facilitate analysis.

Depends  R (>= 4.1)

Imports  checkmate,
sf,
stats,
stringi,
terra,
tools,
utils

Suggests  arrow,
connectapi,
covr,
dataRetrieval,
DiagrammeR,
DiagrammeRsvg,
dm,
fontawesome,
htmltools,
htmlwidgets,
httr,
inlcolor,
jsonlite,
knitr,
pkgload,
pkgbuild,
pkgdown,
plotrix,
rappdirs,
reactable,
renv,
rmarkdown,
roxygen2,
rconnect,
tinytest,
V8,
webmap,
writeXl,
xml2

License CC0


BugReports https://code.usgs.gov/inl/inldata/-/issues

Copyright This software is in the public domain because it contains materials that originally came from the United States Geological Survey (USGS), an agency of the United States Department of Interior. For more information, see the official USGS copyright policy at https://www.usgs.gov/information-policies-and-instructions/copyrights-and-credits

Encoding UTF-8

LazyData true

LazyDataCompression xz

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.1

R topics documented:

background .................................................. 3
benchmarks .................................................. 4
cities ...................................................... 6
counties ..................................................... 6
crs .......................................................... 7
dem ........................................................ 8
Description

Water-quality background concentrations for selected radionuclides, organic compounds, and chemical constituents that were analyzed for in water from the eastern Snake River Plain aquifer at and near the Idaho National Laboratory (INL). The background concentrations are defined as groundwater influenced by western tributary recharge in the western INL and by eastern regional recharge in the eastern INL. These concentrations are either naturally occurring or anthropogenic (substances present in the environment as a result of human activities) and are not influenced by waste and wastewater disposal at the INL, according to Bartholomay and Hall (2016).

Usage

background

Format

A data frame with columns:

- `srsname` Substance Registry Services (SRS) name for constituents, such as "Strontium-90".
- `pcode` U.S. Geological Survey 5-digit parameter code used to identify the constituent measured, see `parameters` dataset for details. For example, the parameter code for Tritium is "07000".
- `parm_unit` Parameter units of measure, see `units` dataset for unit descriptions.
**example**

```r
str(background)
```

<table>
<thead>
<tr>
<th>benchmarks</th>
<th>Benchmark Concentrations</th>
</tr>
</thead>
</table>

**Description**

Water-quality benchmark concentrations of selected radionuclides, organic compounds, and chemical constituents. These benchmarks include the United States Environmental Protection Agency (EPA) Maximum Contaminant Levels (MCLs), Human Health Benchmarks for Pesticides (HHBPs), and U.S. Geological Survey (USGS) Health-Based Screening Levels (HBSLs).

**Usage**

`benchmarks`
### Format

A data frame with columns:

- **srsname**: Substance Registry Services (SRS) name for constituents, such as "Strontium-90".
- **pcode**: U.S. Geological Survey 5-digit parameter code used to identify the constituent measured, see parameters dataset for details. For example, the parameter code for Tritium is "07000".
- **parm_unit**: Parameter units of measure, see units dataset for unit descriptions.
- **mcl**: Maximum Contaminant Levels (MCLs), a standard set by the EPA for drinking water quality. An MCL is the legal threshold limit on the amount of a substance that is allowed in public water systems under the Safe Drinking Water Act.
- **hhbp_noncancer**: EPA Chronic Noncancer HHBPs.
- **hhbp_cancer_min**: EPA Carcinogenic HHBPs for a one-in-one million cancer risk.
- **hhbp_cancer_max**: EPA Carcinogenic HHBPs for a one-in-ten thousand cancer risk.
- **hbsl_noncancer**: USGS Noncancer HBSLs.
- **hbsl_cancer_min**: USGS Cancer HBSLs for a one-in-one million cancer risk.
- **hbsl_cancer_max**: USGS Cancer HBSLs for a one-in-ten thousand cancer risk.
- **remark**: Benchmark remarks that provide additional information about some MCLs, HHBPs, and HBSLs.

### Source

Many of the water-quality benchmarks were accessed from the U.S. Geological Survey Health-Based Screening Levels database, accessed on April 4, 2023, from [https://water.usgs.gov/water-resources/hbsl/](https://water.usgs.gov/water-resources/hbsl/). Benchmarks for total Trihalomethanes, Tritium, and Strontium-90 were provided by the U.S. Environmental Protection Agency (2015). Note that MCL benchmark values reported in millirem per year were substituted with a 50 picocuries per liter screening level.

### References


### Examples

```r
str(benchmarks)
```
**cities**  
*Cities and Towns*

**Description**
Cities and towns (populated places) in the vicinity of Idaho National Laboratory, eastern Idaho.

**Usage**
cities

**Format**
A simple feature with fields:
- name  City name.
- id    Unique identifier.
- geometry Zero-dimensional geometry containing a single point.

**Source**
Spatial point extracts were obtained from the Master Address File / Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) Database (MTDB), 2023 data collection, released November 22, 2023. Which is a part of the U.S. Department of Commerce, U.S. Census Bureau, Geography Division/Cartographic Products Branch. The centroids of these extracts were cropped to the study area, and any non-essential data was removed.

**Examples**

```r
print(cities)

plot(cities["name"], pch = 19)
```

---

**counties**  
*County Boundaries*

**Description**
County boundaries in the vicinity of Idaho National Laboratory, eastern Idaho.

**Usage**
counties
Format

A simple feature with fields:

- **name**: County name.
- **id**: Unique identifier.
- **geometry**: Polygon geometry with a positive area (two-dimensional); sequence of points that form a closed, non-self-intersecting ring; the first ring denotes the exterior ring, zero or more subsequent rings denote holes in this exterior ring.

Source

Spatial polygon extracts were obtained from the Master Address File / Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) Database (MTDB), 2023 data collection, released November 22, 2023. Which is a part of the U.S. Department of Commerce, U.S. Census Bureau, Geography Division/Cartographic Products Branch. These polygons were cropped to study area, and any non-essential data was removed.

Examples

```r
print(counties)
plot(counties["name"])
```

---

**crs**

*Coordinate Reference System*

Description

Coordinate reference system (CRS) used by the U.S. Geological Survey Idaho National Laboratory Project Office. The CRS is based on the following attributes: Albers equal-area conic projection; latitude of first and second standard parallel is 42.83 and 44.16 decimal degrees, respectively; latitude and longitude of false origin is 41.5 and -113 decimal degrees, respectively; easting and northing of false origin is 200,000 and 0 meters, respectively; Clarke (1966) reference ellipsoid; North American Datum of 1983; and units of meters. The CRS is represented using an updated version of the well-known text (WKT2) strings.

Usage

```r
crs
```

Format

A list with the following elements representing the CRS: *input* is the PROJ.4 string, and *wkt* is the WKT2 strings.

Source

Idaho National Laboratory Project Office
Examples

    print(crs)

---

dem  

Digital Elevation Model

Description

The digital elevation model (DEM) for the Idaho National Laboratory and its vicinity in eastern Idaho. A DEM is a representation of the land surface that uses a digital grid to describe the elevation values of the terrain. The spatial grid is composed of 1,142 rows and 1,145 columns, and has cell sizes that are constant at 100 meters by 100 meters.

Usage
dem

Format

A compressed SpatRaster class object with layer:


Source

The National Map (TNM) 1/3-arc-second DEM (Gesch, 2007; Gesch and others, 2002), accessed on August 4, 2020. This dataset can be downloaded in an Esri ArcGRID TM format using TNM Download. Elevation datasets are distributed in geographic coordinates in units of decimal degrees, and in conformance with the NAD 83. Elevation values are in feet above the NAVD 88.

References


See Also

make_shade function for computing the hill shade from the DEM.
Examples

```r
elevation <- terra::unwrap(dem)
print(elevation)

col <- inlcolor::get_colors(n = 256, scheme = "dem2", bias = 0.9)
terra::plot(elevation, col = col)
```

---

**Laboratory Detection Limits**

**Description**

Analytical method detection limits of selected radionuclides based on laboratory procedures.

**Usage**

`dl`

**Format**

A data frame with columns:

- `srsname` Substance Registry Services (SRS) name for constituents, such as "Strontium-90".
- `pcode` U.S. Geological Survey 5-digit parameter code used to identify the constituent measured, see `parameters` dataset for details. For example, the parameter code for Tritium is "07000".
- `parm_unit` Parameter units of measure, see `units` dataset for unit descriptions.
- `lab_det_lim_va` Laboratory detection limit concentration.
- `sdate` Date that the detection limit went into effect.
- `reference` Source of detection limit. Reference citations are as follows: "Bartholomay and others (2003, table 9)", "Bartholomay and others (2014, table D1)", "Bartholomay and others (2021, table 4.1)", and "Bodnar and Percival (1982)"

**Source**

Idaho National Laboratory Project Office

**References**


Examples

str(dl)

---

esrp  
_Eastern Snake River Plain Boundary_

Description

Boundary of the eastern Snake River Plain (ESPR), Idaho. The ESPR is a structural basin that extends about 200 miles in a northeast direction and is 50-70 miles wide. The basin is bounded by faults on the northwest and by down warping and faulting on the southeast. It has been filled with basaltic lava flows interbedded with terrestrial sediments. The combination of basaltic rock and sedimentary deposits forms the ESPR aquifer, which is the primary source of groundwater in the basin.

Usage

esrp

Format

A simple feature with fields:

(geometry Polygon geometry with a positive area (two-dimensional); sequence of points that form a closed, non-self-intersecting ring; the first ring denotes the exterior ring, zero or more subsequent rings denote holes in this exterior ring.

Source

The ESPR boundary was digitized from 7.5’ quads within and in the vicinity of the Idaho National Laboratory by the U.S. Geological Survey Idaho National Laboratory Project Office The digitized polygons were made into shapefiles and converted into GeoJSON files.

Examples

print(esrp)

plot(esrp, col = "grey")
**facilities**

**Idaho National Laboratory Facilities**

**Description**

Federal research facilities at the Idaho National Laboratory (INL). The INL facilities have been the primary source of radioactive and chemical waste constituents in the water from the eastern Snake River Plain aquifer and in perched groundwater zones at or near the INL. This is due to the wastewater disposal practices at the INL facilities.

**Usage**

facilities

**Format**

A simple feature with fields:

- **name** Facility name.
- **id** Facility identifier. Facility abbreviations and descriptions are as follows: "TAN" Test Area North, "NRF" Naval Reactors Facility, "MRF" Materials and Fuels Complex, "ATRC" Advanced Test Reactor Complex, "INTEC" Idaho Nuclear Technology and Engineering Center, "CFA" Central Facilities Area "RWMC" Radioactive Waste Management Complex.
- **geometry** Set of polygons, where a polygon is a geometry with a positive area (two-dimensional).

**Source**

A GeoJSON file of the facilities was created by the U.S. Geological Survey Idaho National Laboratory Project Office

**Examples**

```
print(facilities)

plot(facilities["id"])
```

**gwl**

**Groundwater Levels**
Description
Groundwater levels in wells in the U.S. Geological Survey (USGS) water-level monitoring network, Idaho National Laboratory and vicinity, Idaho. The purpose of this network is to document the changes in storage and the general direction of groundwater flow within the eastern Snake River Plain (ESRP) aquifer. The data collected from this network have been used to determine changes in hydraulic-gradient that affect the rate and direction of groundwater and waste-constituent movement. It can also help identify sources of recharge and measure its effects. The groundwater measurements are taken from both the ESRP aquifer and the perched groundwater zones above the aquifer.

Usage
gwl

Format
A data frame with columns:

- **site_nm**: Local site name.
- **site_no**: USGS site number.
- **lev_dt**: Date and time the water level was measured, in "America/Denver" time zone. Missing values of time were substituted with "12:00".
- **lev_meth_cd**: Code indicating the method used to determine the water level. The codes and their meanings are as follows: "A" airline measurement, "B" analog or graphic recorder, "F" transducer, "G" pressure-gage measurement, "L" interpreted from geophysical logs, "O" observed, "S" steel-tape measurement, "T" electric-tape measurement, "V" calibrated electric tape—accuracy of instrument has been checked, "W" calibrated electric cable, and "Z" other.
- **lev_status_cd**: Code indicating the status of the site at the time the water level was measured. The codes and their meanings are as follows: "1" static, "3" true value is above the reported water level value due to local conditions, "5" groundwater level affected by surface water, "P" site was being pumped.
- **lev_age_cd**: Code indicating the water-level approval status. The codes and their meanings are as follows: "A" approved for publication—processing and review completed, and "P" provisional data subject to revision.
- **lev_va**: Depth to water level, in feet below the land surface reference point.
- **lev_acy_va**: Accuracy of depth to water-level measurement (lev_va), in feet.
- **sl_lev_acy_va**: Accuracy of groundwater level above NAVD value (sl_lev_va), in feet. Does not account for vertical datum shift.

Source
Data was obtained from the National Water Information System (NWIS) (U.S. Geological Survey, 2023). Groundwater level data was retrieved from NWIS by using the USGS R-package dataRetrieval (DeCicco and others, 2023). Data was retrieved for each site number listed in the samples dataset and removed any non-essential columns.
References

Examples
str(gwl)

poi <- as.POSIXct(c("2000-01-01", "2001-01-01")) # period of interest
site_no <- "432700112470801" # well USGS 1
is <- gwl$lev_dt >= poi[1] & gwl$lev_dt < poi[2] & gwl$site_no == site_no
d <- gwl[is, ]
plotrix::plotCI(
  x = d$lev_dt,
  y = d$sl_lev_va,
  li = d$sl_lev_va - d$sl_lev_acy_va,
  ui = d$sl_lev_va + d$sl_lev_acy_va
)

idaho

State of Idaho Boundary

Description
A simplified representation of the boundary of Idaho, a state located in the northwestern region of the United States.

Usage
idaho

Format
A simple feature with fields:

geometry Polygon geometry with a positive area (two-dimensional); sequence of points that form a closed, non-self-intersecting ring; the first ring denotes the exterior ring, zero or more subsequent rings denote holes in this exterior ring.

Source
Spatial extract was obtained from the Master Address File / Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) Database (MTDB), 2023 data collection, released November 22, 2023. Which is part of the U.S. Department of Commerce, U.S. Census Bureau, Geography Division/Cartographic Products Branch. The polygon’s geospatial features were simplified by removing the vertices and any non-essential data was removed.
Examples

```r
print(idaho)
plot(idaho, col = "grey")
```

<table>
<thead>
<tr>
<th>inl</th>
<th>Idaho National Laboratory Boundary</th>
</tr>
</thead>
</table>

Description

Geographic limits of the Idaho National Laboratory (INL). The INL is located on the west-central part of the eastern Snake River Plain and covers an area of approximately 890 square miles. It was established in 1949 to develop atomic energy, nuclear safety, defense programs, environmental research, and advanced energy concepts.

Usage

```r
inl
```

Format

A simple feature with fields:

- geometry: Polygon geometry with a positive area (two-dimensional); sequence of points that form a closed, non-self-intersecting ring; the first ring denotes the exterior ring, zero or more subsequent rings denote holes in this exterior ring.

Source

The INL boundary was digitized from 7.5' quads within and in the vicinity of the Idaho National Laboratory by the U.S. Geological Survey Idaho National Laboratory Project Office. The digitized polygons were made into shapefiles and converted into GeoJSON files.

Examples

```r
print(inl)
plot(inl, col = "grey")
```
### Industrial Waste Ditch

**Description**

An industrial waste ditch located near the Naval Reactors Facility (NRF) in eastern Idaho. The NRF discharges its wastewater to this 1.65-mile-long waste ditch. The waste ditch was initially used in 1953 to dispose of non-radioactive, non-sewage industrial wastewater.

**Usage**

iwd

**Format**

A simple feature with fields:

- geometry: Sequence of points connected by straight, non-self-intersecting line pieces, one-dimensional geometry.

**Source**

A GeoJSON file of the industrial waste ditch was created by the U.S. Geological Survey Idaho National Laboratory Project Office.

**Examples**

```r
print(iwd)
plot(iwd, col = "blue")
```

### Lakes and Ponds

**Description**

Perennial lakes and ponds in the vicinity of the Idaho National Laboratory (INL) in eastern Idaho.

**Usage**

lakes
Format

A simple feature with fields:

- **gnis_nm**: GNIS name.
- **id**: Unique identifier.
- **reach_cd**: Reach code, a unique 14-digit code that identifies a continuous piece of surface water with similar hydrologic characteristics.
- **gnis_id**: Geographic Names Information System (GNIS) identifier.
- **feature_tp**: USGS National Hydrography Dataset (NHD) feature type code. "LakePond" is a standing body of water with a predominantly natural shoreline surrounded by land; "Reservoir" is a constructed basin formed to contain water or other liquids; and "SwampMarsh" is a non-cultivated, vegetated area that is inundated or saturated for a significant part of the year.
- **geometry**: Polygon geometry with a positive area (two-dimensional); sequence of points that form a closed, non-self-intersecting ring; the first ring denotes the exterior ring, zero or more subsequent rings denote holes in this exterior ring.

Source

Spatial polygon extract files obtained from the U.S. Geological Survey (USGS) National Hydrography Dataset (NHD) Medium Resolution for Idaho, released August 4, 2014. Which is part of the National Geospatial Technical Operations Center. These extracts were cropped to eastern Idaho extent and unnecessary columns were removed.

Examples

```r
print(lakes)
plot(lakes["id"], col = "lightblue", border = "blue")
```

Description

Create datasets for the **inldata** package and save each as an R-data file with the .rda extension, which is a format native to R. The **stats dataRetrieval**, and **stringi** packages must be available, and the `7z` executable must be on your path. This function is intended for use by **inldata**-package developers.

Usage

```r
make_datasets(
  path = getwd(),
  destdir = file.path(path, "data"),
  clean = FALSE,
  tz = "America/Denver",
```
make_datasets

census_yr = 2023,
buffer_dist = 1000,
warn = 1,
timeout = 10,
compress = "xz",
seed = 0L
)

Arguments

path 'character' string. Path to the package’s source directory, with tilde-expansion
performed. Defaults to the working directory. Ensure that under the path is a
folder named data-raw that contains the raw data files required for the build
process.

destdir 'character' string. Destination directory to write R-data files, with tilde-expansion
performed. Defaults to the data directory located under path.

clean 'logical' flag. Whether to delete all pre-existing R-data files in the destination
directory.

tz 'character' string. Time zone specification. Defaults to Mountain Standard Time
(North America). See OlsonNames for time zone information.

census_yr 'integer' number. United States census year.

buffer_dist 'numeric' number. Buffer distance for the study area defined by the bounding
of the sample sites dataset. The buffer distance is measured in units of the
coordinate reference system (crs$units).

warn 'integer' value. Sets the handling of warning messages. Choose value of less
than 0 to show no warnings, 1 to print warnings (default), and 2 to error on
warnings.

timeout 'integer' number. Timeout for some of the internet operations, in minutes. De-
defaults to 10 minutes.

compress 'logical' flag or 'character' string. Whether compression should be used when
saving a dataset to file. Character strings "auto", "gzip", "bzip2" and "xz" (de-
fault) are accepted. See the save function for details on compression types.

seed 'integer' count. Random number generator state, used to create reproducible
results.

Details

This function retrieves and parses datasets from local and remote sources. Access to the internet is
required to download data from the following remote sources:

- National Elevation Dataset (NED) on Amazon’s Cloud.
- Spatial data from the TIGER/Line Geodatabase that contains spatial extracts from the U.S.
  Census Bureau’s MAF/TIGER database.
- National Hydrography Dataset (NHD) data from the USGS NHD file geodatabase on Amaz-
  on’s Cloud.
Each of the package dataset's represents a snapshot of the data at a specified point in time. While geospatial datasets may change very little over time (such as the boundary of the Idaho National Laboratory), other datasets continue to grow as new data becomes available (such as water-quality data measured in samples collected from wells).

To ensure that the function retrieves the most recent data versions, it is recommended to periodically check the URLs of remote sources and update them within the function. It is also advisable to document any changes in the datasets and update their help documentation accordingly.

Files downloaded during intermediate stages of the build process are cached on your computer to speed up future builds. You can specify the path to the cache directory by setting an environment variable named CACHE_DIR. By default the location of the cache directory is determined by the get_cache_dir() command.

Value

Returns the paths to the newly created R Data files invisibly.

Author(s)

J.C. Fisher, U.S. Geological Survey, Idaho Water Science Center

Examples

```r
# Example requires that the 'path' argument be specified as
# the top-level directory of the inldata package repository.
## Not run:
make_datasets(destdir = tempfile())

## End(Not run)
```

Description

Create U.S. Geological Survey (USGS) data release product from R-package datasets and their documentation. Requires that the xml2 and jsonlite packages are available.

Usage

```r
make_data_release(
    metadata, package, destdir = getwd(), ...
    bounding = NULL,rngdates = NULL,
    validate = FALSE
)
```
make_data_release

Arguments

metadata `character` string or named 'list'. Either the path to a JSON formatted metadata file that contains general information for the USGS data release (see Examples section), or a named list with the equivalent information.

package `character` string. Name of a package.

destdir `character` string. Destination directory to write files, with tilde-expansion performed.

... Additional arguments to be passed to the write_datasets function. The formats argument, which is specified within the function, is the exception.

bounding 'bbox', 'sf', 'SpatRaster', or 'PackedSpatRaster' spatial feature. Object to compute spatial bounding coordinates from, see sf::st_bbox function.

rngdates 'Date' or 'POSIXct' vector. Object to compute the date range of observations from.

validate `logical` flag. Whether to perform a metadata validation and stop execution if errors are found. See validate_metadata function for system requirements.

Details

Citation entries for the R package (software release) and accompanying data release should be included in the package CITATION file, and documented in that order.

Value

Named list of metadata created for spatial and non-spatial datasets.

Author(s)

J.C. Fisher, U.S. Geological Survey, Idaho Water Science Center

See Also

parse_rd_db function for reading and parsing R-package documentation.

Examples

destdir <- tempfile("")
l <- make_data_release(
    metadata = system.file("extdata/metadata.json", package = "inldata"),
    package = "inldata",
    destdir = destdir,
    include = c("crs", "dl"),
    quiet = TRUE,
    bounding = sites,
    rngdates = c(
        samples$sample_dt,
        gwl$lev_dt
    )
)
make_shade

Description

Compute hill shade from a digital elevation model. A hill shade layer is often used as a backdrop on top of which another, semi-transparent, layer is drawn.

Usage

make_shade(
  x,
  scale = 2,
  neighbors = 8,
  angle = 40,
  direction = 270,
  normalize = TRUE
)

Arguments

x 'SpatRaster' object. Spatial raster of land-surface elevations, such as the `dem` dataset.

scale 'numeric' number. Scaling factor used to convert the elevation values. The default value is 2.

neighbors 'numeric' count. Number of neighboring cells to use to compute slope or aspect with. Either 8 (queen case) or 4 (rook case).

angle 'numeric' number. The sun's angle of elevation above the horizon, ranges from 0 to 90 degrees. A value of 0 degrees indicates that the sun is on the horizon. A value of 90 degrees indicates that the sun is directly overhead. Default is a 40 degree angle.

direction 'numeric' number. Direction (azimuth) angle of the light source (sun), in degrees. An direction of 0 degrees indicates north, east is 90 degrees, south is 180 degrees, and west is 270 degrees (default).

normalize 'logical' flag. Whether to set values below zero to zero and results normalized between 0 and 1.

Value

Spatial raster of hill shade values.
**Author(s)**

J.C. Fisher, U.S. Geological Survey, Idaho Water Science Center

**Examples**

```r
elevation <- terra::unwrap(dem)
hillshade <- make_shade(elevation)
print(hillshade)

col <- inlcolor::get_colors(n = 256, scheme = "grayC")
mar <- c(2, 2, 1, 4)
terra::plot(hillshade, col = col, mar = mar)

terra::plot(hillshade, col = col, mar = mar, legend = FALSE)
terra::plot(elevation,
    col = inlcolor::get_colors(
        n = 256,
        scheme = "dem2",
        alpha = 0.7,
        bias = 0.9
    ),
    add = TRUE
)
```

---

<table>
<thead>
<tr>
<th>parameters</th>
<th>Parameter Information</th>
</tr>
</thead>
</table>

**Description**

Parameter information for selected chemical constituents, organic compounds, and radionuclides measured for in water samples collected from monitoring sites in the U.S. Geological Survey (USGS) water-quality monitoring network, Idaho National Laboratory and vicinity, Idaho.

**Usage**

`parameters`

**Format**

A data frame with columns:

- `pcode` USGS 5-digit parameter code.
- `parm_group_nm` Parameter group name. Abbreviations and descriptions are as follows: "Information" includes information about the water sample such as the date and time of collection, the location of the sample, and the method of analysis used; "Inorganics, Major, Metals" includes major inorganic ions and metals; "Inorganics, Major, Non-metals" includes major inorganic non-metallic ions; "Inorganics, Minor, Metals" includes minor inorganic metallic ions; "Inorganics, Minor, Non-metals" includes minor inorganic non-metallic ions; "Nutrient" includes
nutrients such as nitrogen and phosphorus that can be present in water samples; "Organics, Other" includes organic compounds that are not classified as pesticides or PCBs; "Organics, Pesticide" includes organic compounds that are used as pesticides, such as insecticides, herbicides, and fungicides; "Physical" includes physical characteristics of water; "Radiochemical" includes radioactive isotopes that can be present in water samples; and "Stable Isotopes" includes non-radioactive isotopes of elements that can be used to trace the movement of water through the hydrologic cycle.

`parm_nm` Long parameter name, such as "Strontium-90, water, unfiltered, picocuries per liter".
`casrn` Chemical Abstracts Service (CAS) registry number, such as "10098-97-2" for Strontium-90.
`srsname` Substance Registry Services (SRS) name, such as "Strontium-90".
`parm_unit` Parameter units of measure, see units dataset for unit descriptions.
`min_dt` Collection date of first sample analyzed for the parameter.
`max_dt` Collection date of last sample analyzed for the parameter.
`nrecords` Number of records associated with the parameter.
`nsites` Number of sampling sites where the parameter was observed.

**Source**

USGS water data acquired from the National Water Information System (U.S. Geological Survey, 2023). The SRS name (`srsname`) for "Trihalomethanes (four), total, from SDWA NPDWR" was shorten to its preferred acronym "TTHM4".

**References**


**Examples**

```r
str(parameters)
```

<table>
<thead>
<tr>
<th><code>percponds</code></th>
<th>Percolation Ponds</th>
</tr>
</thead>
</table>

**Description**

Percolation ponds in the vicinity of Idaho National Laboratory (INL), eastern Idaho. Wastewater from facilities at the INL has been and is currently disposed of in percolation (infiltration), evaporation (lined and unlined), and infiltration ponds.

**Usage**

`percponds`
Format

A simple feature with fields:

name Name of the percolation ponds.

facility_id INL facility the percolation pond is located at. Abbreviations and descriptions are as follows: "ATRC" is the Advanced Test Reactor Complex, "INTEC" is the Idaho Nuclear Technology and Engineering Center, "MFC" is the Materials and Fuels Complex, "NRF" is the Nuclear Reactors Facility, and "TAN" is the Test Area North facility.

geometry Polygon geometry with a positive area (two-dimensional); sequence of points that form a closed, non-self-intersecting ring; the first ring denotes the exterior ring, zero or more subsequent rings denote holes in this exterior ring.

Source

U.S. Geological Survey Idaho National Laboratory Project Office

Examples

print(percponds)

plot(percponds["name"], col = "lightblue", border = "blue")

roads

Road Network

Description

Road network in the vicinity of Idaho National Laboratory, eastern Idaho.

Usage

roads

Format

A simple feature with fields:

name Street or road name.

id Unique identifier.

prisec_f1 Whether a road is classified as primary or secondary.

geometry Sequence of points connected by straight, non-self-intersecting line pieces, one-dimensional geometry.
Source

Spatial line extracts were obtained from the Master Address File / Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) Database (MTDB), 2023 data collection, released November 22, 2023. Which is a part of the U.S. Department of Commerce, U.S. Census Bureau, Geography Division/Cartographic Products Branch. These lines were cropped to the study area, and any non-essential data was removed.

Examples

```r
print(roads)
plot(roads["prisec_fl"])
```

<table>
<thead>
<tr>
<th>samples</th>
<th>Discrete Sample Data</th>
</tr>
</thead>
</table>

Description

Water-quality information for both groundwater and surface water collected from monitoring stations in and around the Idaho National Laboratory, Idaho. The water samples were collected in the field and analyzed in a laboratory to obtain water-quality data. The dataset was obtained from the National Water Information System (NWIS), which is maintained by the U.S. Geological Survey. The NWIS is a comprehensive and distributed application that supports the acquisition, processing, and long-term storage of water data.

Usage

samples

Format

A data frame with columns:

- **site_nm** Local site name.
- **sample_dt** Date and time the sample was collected, in "America/Denver" time zone. Missing values of time were substituted with "12:00".
- **parm_shortNm** Parameter short name assigned by the USGS, such as "pH".
- **parm_unit** Parameter units of measure, see units dataset for unit descriptions.
- **remark_cd** Remark code (result level) used to qualify the parameter value. The codes and their meanings are as follows: "" is a quantified value; "<" is where the actual value is known to be less than the value reported, that is, the measured concentration is below the reporting limit (RL) and represented as a censored (or nondetection) value. For censored values, the value reported is the RL; "E" is an estimated value, that is, the actual value is greater than the minimum detection limit (MDL) and less than the laboratory reporting level (LRL); "R" is a nondetect, result less than sample-specific critical level; and "U" is a material specifically analyzed for but not detected.
result_va  Parameter value.
lab_sd_va  Laboratory standard deviation (SD). For radiochemical data, SD is typically determined from the counting error. Prior to January 1, 2008, counting error was reported as two SD (Bartholomay and others, 2020, p. 27), therefore, these values were divided by 2.
lab_li_va  Lower confidence interval of the result value based on laboratory analysis.
lab_ui_va  Upper confidence interval of the result value based on laboratory analysis. In cases where the upper and lower limits are identical, the parameter is expressed as an exact value.
rpt_lev_va  Laboratory reporting limit in effect for the parameter and analytical method at the time the measurement was made.
rpt_lev_cd  Reporting level code that identifies the analytical reporting level appropriate for the analytical method. The codes and their meanings are as follows: "DLBLK" detection limit by blank data; "DLDQC" detection limit by DQCALC, lowest concentration that with 90 percent confidence will be exceeded no more than 1 percent of the time when a blank sample is measured; "IRL" interim reporting level, a temporary reporting level; "LRL" laboratory reporting level, equal to twice the yearly-determined LT-MDL; "LT-MDL" long-term method detection limit, a detection level derived by determining the standard deviation of a minimum of 24 MDL spike sample measurements over an extended period of time; "MDL" method detection limit, minimum concentration of a substance that can be measured and reported with a 99 percent confidence that the analyte concentration is greater than zero; "PQL" practical quantitation limits; "MRL" minimum reporting level, smallest measured concentration that can be reliably measured using a given analytical method; "RLDQC" reporting limit by DQCALC, is greater than or equal to two times the DLDQC; "SSLC" sample-specific critical level, the calculated and reported value is below which the radiochemistry result is considered a non-detect; and "SSMDC" sample-specific minimum detectable concentration, a reporting level that varies for each sample and is primarily used in radiochemical analyses.
medium_cd  Medium code that identifies the material type and quality assurance type of the sample. The codes and their meanings are as follows: "OAQ" is a blank sample collected for QC purposes; "WG" is water below land surface contained in the saturated zone (groundwater); "WGQ" is a groundwater quality-control (QC) sample; "WS" is water on the surface of the Earth (surface water); and "WSQ" surface water QC sample.
anl_ent_cd  Analyzing entity code of the organizational unit that performed the sample analysis used to obtain the result.
dqi_cd  Data quality indicator code that indicates the review status of a result. The codes and their meanings are as follows: "A" historical data, "R" reviewed and accepted, and "S" provisional (presumed satisfactory).
meth_cd  Method code, the codes are documented in the NWIS Method Code Dictionary.
sample_type_cd  Sample type code that identifies the quality-assurance (QA) type of a sample. The codes and their meanings are as follows: "2" is a blank sample; "6" is a reference material sample; "7" is a replicate sample taken from the environment; "9" is a regular sample taken from the environment; "B" is a unspecified QA sample; and "H" is a composite (time) sample.
db_no  2-digit NWIS database number. The codes and their meanings are as follows: "01" is the environmental database, and "10" is the QA database.
sample_id  Unique identifier for the water sample. The sample code is a concatenation of the site number, medium code, and date-time the sample was collected.
site_no  USGS site identification number.
pcod  USGS 5-digit parameter code. For example, the parameter code for Tritium is "07000".

rep_pair_id  Unique identifier used for matching pairs of replicate samples for a specific parameter. Replicate pairs are identified by matching a replicate sample (sample_type_cd equal to 7) with its corresponding regular environmental sample (sample_type_cd equal to 9).

result_tx  Remark about the water quality result

remark  Remarks pertaining to changes applied after the records were obtained from NWIS.

anl_dt  Result analysis date.

Source

Data were obtained from the NWIS-QWDATA database on January 22, 2024, in tab-delimited output-format using the QWDATA system (U.S. Geological Survey, 2024). The following steps were taken to process the data:

- **Column Name Translation**: Column names were switched from NWIS alpha codes to NWIS parameter codes.
- **Class Conversion**: Classes for each column were converted to the appropriate type (numeric, POSIX, factor).
- **Column Removal**: Unnecessary columns were removed.
- **Duplicate Removal**: Duplicate records in the NWIS and QWDATA databases were removed.
- **Data Cleaning**: Corrupted results were removed. A column was added for data processing remarks. Zero and negative results were reported as nondetects.
- **Radiochemical Parameter Identification**: Radiochemical parameter codes were identified. Nondetects were reported as less than the reporting level.
- **Sample Type Reporting**: In the absence of a sample type code, results were classified as environmental samples. Data from blank, spiked, and reference samples, as well as composite samples (collected over a period of time), were excluded.
- **Sample Identifier Creation**: Site number and date/time were combined to create a unique sample identifier.
- **Error Conversion**: Counting error was converted to laboratory standard deviation. Data entered into NWIS with an uncertainty of twice the standard deviation (data entered before 2008), were converted to one standard deviation.
- **Record Removal**: Records associated with counting error parameter codes were removed.
- **Replicate Pairing**: Replicate samples were paired based on the parameter code, site number, and time difference between sample collection.
- **Data Adjustment**: The detection limit, left-censored values, and non-positive lower limit were accounted for. Interval censored data were represented using upper and lower limits of the 95-percent confidence interval of the parameter value.

References


Examples

```r
str(samples)

poi <- as.POSIXct(c("1989-01-01", "2019-01-01")) # period of interest
is_poi <- samples$sample_dt >= poi[1] & samples$sample_dt < poi[2]
is_stc <- samples$sample_type_cd %in% c("7", "9")

site_no <- "433253112545901" # well USGS 20
pcode <- "07000" # tritium, water, unfiltered, picocuries per liter
is <- is_poi & is_stc & samples$site_no == site_no & samples$pcode == pcode
d <- samples[is, ]
plotrix::plotCI(
  x = d$sample_dt,
  y = d$result_va,
  li = d$lab_li_va,
  ui = d$lab_ui_va
)

site_no <- "433322112564301" # well USGS 38
pcode <- "01030" # chromium, water, filtered, micrograms per liter
is <- is_poi & is_stc & samples$site_no == site_no & samples$pcode == pcode
d <- samples[is, ]
plotrix::plotCI(
  x = d$sample_dt,
  y = d$result_va,
  li = d$lab_li_va,
  ui = d$lab_ui_va
)
```

<table>
<thead>
<tr>
<th>Site Information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>site_no</strong></td>
</tr>
<tr>
<td><strong>pcode</strong></td>
</tr>
<tr>
<td><strong>is</strong></td>
</tr>
<tr>
<td><strong>d</strong></td>
</tr>
</tbody>
</table>

Description

Information for sites in the U.S. Geological Survey (USGS) monitoring networks, Idaho National Laboratory (INL) and vicinity, Idaho. The monitoring networks at the INL have evolved to their current 2024 schedules of monitoring because groundwater and surface-water sites were installed for various reasons; for example, the need to estimate the volume of the eastern Snake River Plain aquifer, or to document and understand changes in water quality. The USGS compiled the list of sites monitored by the USGS using publications by Knobel and others (2005, table 1) and Bartholomay (2022, table 1). One well site (site no. 425812113271201) was omitted from the dataset because its geographical location falls outside the designated study area.
Usage

sites

Format

A simple feature with fields:

- site_nm: Local site name.
- station_nm: USGS station name with all extra spaces removed from text. Extra spaces include leading, trailing and in-between spaces except for a single space character between words.
- site_no: USGS site identification number.
- coord_meth_cd: Latitude/longitude coordinate method code. The codes and their meanings are as follows: "C" is surveyed; "D" is a differentially corrected Global Positioning System (GPS); "G" mapping grade GPS unit (handheld accuracy range 3.7 to 12.2 meters); "L" long range navigation system; "M" interpolated from topographic map; "N" is not surveyed; "S" is a transit, theodolite, or other surveying method; and "X" is not applicable.
- coord_acy_va: Accuracy of latitude/longitude value in seconds.
- alt_meth_cd: Method code for measuring elevation. The codes and their meanings are as follows: "D" is a differentially corrected global positioning system, "L" is a level or other surveyed method, and "M" is interpolated from a topographic map; "N" is interpolated from a Digital Elevation Model; "X" is a GNSS2 - Level 2 Quality Survey Grade Global Navigation Satellite System; and "Y" is a GNSS3 - Level 3 Quality Survey Grade Global Navigation Satellite System.
- alt_acy_va: Accuracy of the elevation value (alt_va), in feet. Does not account for vertical datum shift.
- construction_dt: Date the well was completed.
- huc_cd: Hydrologic unit code (HUC). Hydrologic units are geographic areas representing part or all of a surface drainage basin or distinct hydrologic feature and are delineated on the Hydrologic Unit Map. The codes and their meanings are as follows: "17040201" is the Idaho Falls watershed; "17040206" is the Portneuf River; "17040209" is the Blackfoot River; "17040214" is Beaver-Camas; "17040215" is the Medicine Lodge Creek; "17040216" is the Big Lost River; "17040217" is the Little Lost River; and "17040218" is Birch Creek.
- reliability_cd: Reliability code for data available for the site. The codes and their meanings are as follows: "C" is data that have been checked by the reporting agency; "M" is minimal data; and "U" is unchecked data.
- nat_aqfr_cd: National aquifer code where "N600SKRVPB" is the Snake River Plain basin-fill aquifers, and "N600SKRVPV" is the Snake River Plain basaltic-rock aquifers.
- aqfr_cd: Aquifer code defined by the catalog of aquifer names and geologic unit codes used by the Water Mission Area. The codes and their meanings are as follows: "100CNZC" is the Cenozoic Era, and "110SKRV" is the Snake River Group.
- aqfr_type_cd: Aquifer type code. The codes and their meanings are as follows: "C" is a confined single aquifer, "M" is confined multiple aquifers, "U" is an unconfined single aquifer, and "X" is mixed (confined and unconfined) multiple aquifers.
sites 29

well_depth_va  Depth of the finished well, in feet below the land surface datum.
hole_depth_va  Total depth of the borehole, in feet below the land surface datum.
depth_src_cd  Source code for depth measurements. The codes and their meanings are as follows: "A" is reported by another government agency, "D" is from driller’s log or report, "G" is reported by a private geologist-consultant or university associate, "L" is interpreted from geophysical logs by personnel of source agency, "O" is reported by owner of well, "R" is reported by person other than the owner, driller, or another government agency, "S" is measured by personnel of reporting agency; and "Z" is other source.
completion_cd  Borehole completion code. The codes and their meanings are as follows: "M" is a multilevel completion, "O" is an open hole completion, and "P" is an open hole completion prior to multilevel completion.
network_cd  Monitoring network code. The codes and their meanings are as follows: "A" is the aquifer-monitoring network, "P" is the perched-groundwater monitoring network, "S" is the surface-water monitoring network.
pos  Position specifier for site-labels on a map. Values of 1, 2, 3 and 4, respectively indicate positions below (south), to the left (west) of, above (north), and to the right (east) of the site coordinates.
min_dt  Date of first record.
max_dt  Date of last record.
nmeas  Number of measurements recorded at the site includes either groundwater-level or streamflow measurements.
nsamples  Number of water-quality samples collected from the site.
nreps  Number of replicate samples collected from the site.
geometry  Zero-dimensional geometry containing a single point.

Source

USGS site data acquired from the NWIS (U.S. Geological Survey, 2023). The station name (station_nm) for the data was parsed for a more readable name. A borehole completion code (completion_cd), as well as an aquifer monitoring network code (network_cd) were created. A position specifier for site-labels on a map (pos) was also created. The classes of the columns were specified, and any unnecessary columns were removed.

References


Examples

```
print(sites)
plot(sites["network_cd"])
```

---

**Description**

Stream segments in the vicinity of Idaho National Laboratory (INL), eastern Idaho. Surface water infiltrated to the eastern Snake River Plain aquifer through river and streams.

**Usage**

```
streams
```

**Format**

A simple feature with fields:

- **gnis_nm** GNIS name.
- **id** Unique identifier.
- **reach_cd** Reach code, a unique 14-digit code that identifies a continuous piece of surface water with similar hydrologic characteristics.
- **gnis_id** Geographic Names Information System (GNIS) identifier.
- **feature_tp** USGS National Hydrography Dataset (NHD) feature type code. The codes and their meanings are as follows: "ArtificialPath" is a surrogate for general flow direction; "CanalDitch" is an artificial open waterway constructed to transport water, to irrigate or drain land, to connect two or more bodies of water, or to serve as a waterway for watercraft; "Connector" is a linear water feature that connects two or more waterbodies or other linear water features; and "StreamRiver" is a linear water feature that is a natural or man-made flowing body of water.
- **geometry** Sequence of points connected by straight, non-self-intersecting line pieces, one-dimensional geometry.

**Source**

Spatial line extract files obtained from the U.S. Geological Survey (USGS) National Hydrography Dataset (NHD) Medium Resolution for Idaho, released August 4, 2014. Which is part of the National Geospatial Technical Operations Center. These extracts were cropped to eastern Idaho extent and unnecessary columns were removed.

Examples

```
print(streams)
plot(streams["id"], col = "blue")
```
Surface-Water Measurements

Description

Stage measurements and calculated discharge at gauging stations in the U.S. Geological Survey (USGS) surface-water monitoring network, Idaho National Laboratory and vicinity, Idaho. The purpose of this network is to provide information on the availability and distribution of surface water resources.

Usage

Format

A data frame with columns:

- site_nm Local site name.
- site_no USGS site number.
- stage_dt Date and time the water level was measured, in "America/Denver" time zone. Missing values of time were substituted with "12:00".
- stage_va Gage height (stage) as shown on the inside staff gage at the site or read off the recorder inside the gage house in feet.
- disch_va Computed discharge in cubic feet per second (cfs).
- stage_acy_va Accuracy of stage measurement in feet.
- disch_acy_va Accuracy of computed discharge in cfs.

Source

Data was obtained from the National Water Information System (NWIS) (U.S. Geological Survey, 2023). Groundwater level data was retrieved from NWIS by using the USGS R-package dataRetrieval (DeCicco and others, 2023). Data was retrieved for each site number listed in the samples dataset and removed any non-essential columns.

References


Examples

```r
str(swm)
poi <- as.POSIXct(c("2000-01-01", "2024-01-01")) # period of interest
site_no <- "13132500" # station BIG LOST RIVER NR ARCO ID
is <- swm$stage_dt >= poi[1] & swm$stage_dt < poi[2] & swm$site_no == site_no
d <- swm[is, ]
plotrix::plotCI(
  x = d$stage_dt,
  y = d$stage_va,
  li = d$stage_va - d$stage_acy_va,
  ui = d$stage_va + d$stage_acy_va
)
```

---

### units

<table>
<thead>
<tr>
<th>Parameter Units</th>
</tr>
</thead>
</table>

### Description

Information describing the parameter units of measure.

### Usage

`units`

### Format

A data frame with columns:

- `parm_unit` Parameter units abbreviation.
- `desc` Parameter units description.
- `siunitx` Parameter units formatted for LaTeX using the `siunitx` package notation.

### Source

Idaho National Laboratory Project Office

### Examples

```r
str(units)
```
write_datasets

Write Package Datasets

Description

Exports the content of package datasets into non-proprietary, open, and well-documented standard file formats, enhancing their accessibility for future use. It supports multiple formats: CSV, JSON (requires jsonlite package), Arrow Parquet (requires arrow package), Excel XLMS (requires writexl package), GeoJSON, Shapefile, and GeoTIFF.

Usage

```r
write_datasets(
  package,
  destdir = getwd(),
  formats = NULL,
  overwrite = TRUE,
  include = NULL,
  exclude = NULL,
  pretty = TRUE,
  quiet = FALSE
)
```

Arguments

- `package` 'character' string. Name of a package.
- `destdir` 'character' string. Destination directory to write files, with tilde-expansion performed.
- `formats` 'character' vector. Formats for saving datasets. Choose from one or more of the following formats: csv, geojson, json, parquet, shp, tiff, txt, and xlsx. Please refer to the details section for a description of each format. All file formats are saved to the disk by default.
- `overwrite` 'logical' flag. Whether to overwrite an existing file.
- `include` 'character' vector. Names(s) of datasets to include. By default, a file is created for each package dataset.
- `exclude` 'character' vector. Name(s) of datasets to exclude. By default, none are excluded.
- `pretty` 'logical' flag. Whether to add indentation, whitespace, and newlines to JSON output (default is TRUE). See prettify function for details. The tradeoff for human-readable output is a much larger file size.
- `quiet` 'logical' flag. Whether to suppress printing of debugging information.
Details

Different types of datasets are written to various file formats. For instance, a dataset in the form of a data table (data.frame class) is exported to a CSV (csv), JSON (json), Parquet (parquet), and Spreadsheet (xlsx) file. A dataset representing simple features (sf or sfc class) is exported to a GeoJSON (geojson) and Shapefile (shp) file. A spatial-raster dataset (SpatRaster or PackedSpatRaster class) is exported to a GeoTIFF (tiff) file. Lastly, a coordinate reference system (crs class) is exported to a text (txt) file.

Value

Invisibly returns the output file path(s).

Author(s)

J.C. Fisher, U.S. Geological Survey, Idaho Water Science Center

Examples

dir <- tempfile(""
write_datasets(
    package = "inldata",
    destdir = dir,
    include = c("crs", "dl", "inl"),
    pretty = FALSE,
    quiet = TRUE
)
)

unlink(dir, recursive = TRUE)
Index

* datasets
  background, 3
  benchmarks, 4
  cities, 6
  counties, 6
  crs, 7
  dem, 8
  dl, 9
  esrp, 10
  facilities, 11
  gwl, 11
  idaho, 13
  inl, 14
  iwd, 15
  lakes, 15
  parameters, 21
  percponds, 22
  roads, 23
  samples, 24
  sites, 27
  streams, 30
  swm, 31
  units, 32

background, 3
benchmarks, 4

cities, 6
counties, 6
crs, 7

idaho, 13
inl, 14
iwd, 15

lakes, 15

make_data_release, 18
make_datasets, 16
make_shade, 8, 20

OlsonNames, 17

parameters, 3, 5, 9, 21
parse_rd_db, 19
percponds, 22
prettify, 33

roads, 23

samples, 18, 24
save, 17

sf::st_bbox, 19
sites, 17, 27
SpatRaster, 8

streams, 30
swm, 31

Time zone, 17

units, 3, 5, 9, 22, 24, 32

validate_metadata, 19

write_datasets, 19, 33

35