Package 'CSLSdata'

April 1, 2021

Type Package

Title WI DNR Central Sands Lakes Study Field Data

Version 1.0.2

Author Carolyn Voter <cvoter@wisc.edu>

Maintainer Carolyn Voter <cvoter@wisc.edu>

Description The package 'CSLSdata' includes all field data used in analyses for the Wisconsin Department of Natural Resources's Central Sands Lakes Study (CSLS). Unprocessed, raw data and scripts are included in 'data-raw', with extensive notes about processing included in the documentation for each dataset.

Depends R (>= 2.10)

Suggests lubridate,
dplyr,
jsonlite,
ggsn,
ggplot2,
extrafont,
NISTunits,
CSLSevap,
bit64,
stringr,
zoo,
raster,
reshape2,
sf,
usethis,
testthat,
knitr,
rmarkdown

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Roxygen list(markdown = TRUE)

VignetteBuilder knitr

R topics documented:

`bathymetry` *CSLS bathymetry relationships*

Description

Info from bathymetry data in ArcGIS for the lake including the elevation (m), surface area (m^2), and volume (m^3) relationship. Calculated using the Storage Capacity tool in Spatial Analyst Supplemental Tools. Also includes:

- The average lake profile (horizontal distance from highest shoreline for each contour level, m), calculated from the lake elevation rasters.
- The area of each plant community at each elevation.
- The area of suitably hard substrate for centrarchid fish species at Pleasant Lake

Usage

```
data(bathymetry)
```

Format

A data frame with the following columns:

lake name of lake, i.e. "Pleasant", "Long", "Plainfield"

elev_m elevation of lake surface above mean sea level (m)

area_m2 surface area of lake when it is at this elevation (m^2)

vol_m3 total volume of the lake when it is at this elevation (m^3)

horiz_dist_m horizontal distance of this lake shoreline from highest lake shoreline (that can be calculated in the raster file) (m)

plant_area_m2 total area of lake used to calculate plant areal extents (m^2)

upland_m2 area of upland plants at this lake at this elevation (m^2)

inland_beach_m2 area of inland_beach plants at this lake at this elevation (m^2)

emergent_m2 area of emergent plants at this lake at this elevation (m^2)

floating_m2 area of floating plants at this lake at this elevation (m^2)

submergent_weed_m2 area of submergent_weed (Potamogeton) plants at this lake at this elevation (m^2)

submergent_algae_m2 area of submergent_algae (Nitella) plants at this lake at this elevation (m^2)

submergent_m2 area of submergent plants at this lake at this elevation (m^2)

upland_pcnt area of upland plants at this lake at this elevation (% of lake outline)

inland_beach_pcnt area of inland_beach plants at this lake at this elevation (% of lake outline)

emergent_pcnt area of emergent plants at this lake at this elevation (% of lake outline)

floating_pcnt area of floating plants at this lake at this elevation (% of lake outline)

submergent_weed_pcnt area of submergent_weed (Potamogeton) plants at this lake at this elevation (% of lake outline)

submergent_algae_pcnt area of submergent_algae (Nitella) plants at this lake at this elevation (% of lake outline)

submergent_pcnt area of submergent plants at this lake at this elevation (% of lake outline)

meanHrdAtNest approximate area of suitably hard substrate for centrarchid fish at this elevation (% of lake outline)

Details

Raw csv data is processed in the `data-raw/` subdirectory of this project with the function `import_bathymetry.R` which is run by `runall_cslsdata.R`.

`import_bathymetry.R` loads bathymetry data from ArcGIS for all lakes including the elevation (m), surface area (m^2), and volume (m^3) relationship. No cleaning is needed. It then calculates the average lake profile from the lake rasters. Note that due to limitation in the raster extent, Long Lake calculations begin ~0.5 ft below maximum elevation available from ArcGIS elev/area/vol calculations. Next, it calculates the estimated area of each plant community at each elevation given rules on plant water depth limits (`import_plant_limits.R`). Lastly, it combines all of this with additional information on the approximate area of suitably hard substrate for centrarchid fish species at Pleasant Lake (`import_fish_substrate.R`)

References

https://www.arcgis.com/home/item.html?id=3528bd72847c439f88190a137a1d0e67

Examples

```
bathymetry <- CSLSdata::bathymetry
```

---

dictionary *CSLS dictionary of sites*

Description

Dictionary for looking up alternate names and identification numbers for sample sites in the CSLS.

Usage

```
data(dictionary)
```

Format

A list of three data frames, one for each lake:

Pleasant site information for sites associated with Pleasant Lake, including precipitation

Long site information for sites associated with Pleasant Lake, including precipitation

Plainfield site information for sites associated with Pleasant Lake, including precipitation

Details

Raw csv data is processed in the `data-raw/` subdirectory of this project with the functions `import_dictionary.R` and `subset_dictionary.R` which are run by `runall_cslsdata.R`. No cleaning or transformations are performed on raw csv data beyond converting to .Rda.

Final data frame for each lake includes columns for:

- **lake:** lake name, "Pleasant", "Long", or "Plainfield", factor.
- **obs_type:** observation type, "LK" (lake), "GW" (groundwater), or "P" (precipitation), factor.
- **site_id:** unique site id for measurement site, e.g., LL-01, factor.
- **SWIMS_station_name:** SWIMS station name, if exists for this site, factor.
- **SWIMS_station_id:** SWIMS station id, if exists for this site, integer.
- **USGS_id:** USGS site number, corresponds to site_no in "water_levels" data frame, numeric.
- **WBIC:** water body identification code, for lake sites only, integer.
- **static_iso_class:** upgradient, downgradient, precipitation, lake, deep, invalid, or NA, factor.
- **static_iso_class_detailed:** upgradient, downgradient, precipitation, lake, deep, invalid, inconsistent, typically_downradient, or NA, factor.
- **lat_deg:** latitude of the lake, decimal degrees, numeric.
- **long_deg:** longitude of the lake, decimal degrees, numeric.
- **elev_m:** approximate elevation of the lake, meters above mean sea level, integer.
- **bouy_bottom_elev_m:** elevation of the bottom of the lake at the bouy location, meters above mean sea level, numeric.

Examples

```
# Load dictionary for Pleasant Lake directly
lake       <- "Pleasant"
dictionary <- CSLSdata::dictionary[[lake]]

# Load dictionary, then subset for specific lake
data(dictionary)
psnt_dictionary <- dictionary$Pleasant
```

gw_levels *CSLS daily groundwater levels*

Description

Daily groundwater level observations for groundwater monitoring wells at the CSLS lakes. Groundwater monitoring wells and sensors are maintained by the Wisconsin State Geologic and Natural History Survey.

Usage

```
data(gw_levels)
```

Format

A data frame with the following columns:

site_id unique CSLS site id

lake associated lake (Pleasant, Long, Plainfield)

date date of measurement

level_m gw elevation, meters above mean sea level

lake_level_m lake elevation, meters above mean sea level

diff_m difference between lake and gw elevation on this day (m)

window_diff_m median difference over past 30 days

site_type gradient classification based on median 30-day difference (upgradient, downgradient, or no gradient)

Details

Raw data are retrieved from DNR Water Use section ArcGIS feature services for the Central Sands. Water quantity information is summarized at a daily time step for these measurements. The feature layer of interest is https://uadnrmaps.wi.gov/arcgis/rest/services/DG_HiCap/DG_CSLS_QUANT_MON_WTM_EXT/FeatureServer/2.

Raw data is processed in with the function "import_water_levels.R" in the "data-raw" subdirectory of this project, then subsetted with the "subset_gw_levels.R" function.

Raw csv data is processed in the `data-raw/` subdirectory of this project with the functions `import_water_levels.R` and `subset_gw_levels.R` which are run by `runall_cslsdata.R`.

`import_water_levels.R` downloads the current feature layer from DNR feature services with water quantity measurements in the Central Sands (summarized at a daily time step). Feature layer of interest is DG_HiCap > DG_CSLS_QUANT_MON_WTM_EXT > Layer 2 (W13101.WU_CSLS_QUANT_DATA). The only cleaning performed is subsetting to desired parameters and adjusting datetime formats.

`subset_gw_levels.R` subsets the larger water level data frame by USGS_id to retrieve only groundwater records for a given lake, then attaches the CSLS site_id to the record. Also calculates the median 30-day differences and uses this to define gradient of well on each day.

Source

https://uadnrmaps.wi.gov/arcgis/rest/services/DG_HiCap/DG_CSLS_QUANT_MON_WTM_EXT/FeatureServer/2

See Also

lake_levels

Examples

```
gw_levels <- CSLSdata::gw_levels
```

---

lake_levels *CSLS daily lake levels*

Description

Daily lake level observations for the CSLS lakes. Lake gages are maintained by the United States Geological Survey.

Usage

```
data(lake_levels)
```

Format

A data frame with the following columns:

lake associated lake (Pleasant, Long, Plainfield)

site_id unique CSLS site id

date date of measurement

level_m lake elevation, meters above mean sea level

area_m2 lake area at this elevation, m^2

vol_m3 lake volume at this elevation, m^3

Details

Raw data are retrieved from DNR Water Use section ArcGIS feature services for the Central Sands. Water quantity information is summarized at a daily time step for these measurements. The feature layer of interest is https://uadnrmaps.wi.gov/arcgis/rest/services/DG_HiCap/DG_CSLS_QUANT_MON_WTM_EXT/FeatureServer/2.

Raw csv data is processed in the data-raw/ subdirectory of this project with the functions import_water_levels.R and subset_lake_levels.R which are run by runall_cslsdata.R.

import_water_levels.R downloads the current feature layer from DNR feature services with water quantity measurements in the Central Sands (summarized at a daily time step). Feature layer of interest is DG_HiCap > DG_CSLS_QUANT_MON_WTM_EXT > Layer 2 (W13101.WU_CSLS_QUANT_DATA). The only cleaning performed is subsetting to desired parameters and adjusting datetime formats.

subset_lake_levels.R subsets larger water level data frame by USGS_id to retrieve only lake level records.

Source

https://uadnrmaps.wi.gov/arcgis/rest/services/DG_HiCap/DG_CSLS_QUANT_MON_WTM_EXT/FeatureServer/2

See Also

gw_levels

Examples

```
lake_levels <- CSLSdata::lake_levels
```

---

lake_raster *CSLS lake rasters*

Description

Lake elevation rasters for each lake as exported from ArcGIS.

Usage

```
data(lake_raster)
```

Format

A list with the following columns:

Pleasant elevation raster for Pleasant Lake (m)

Long elevation raster for Long Lake (m)

Plainfield elevation raster for Plainfield Lake (m)

Details

Raw raster data is processed in the `data-raw/` subdirectory of this project with the function `import_lake_rasters.R` which is run by `runall_cslsdata.R`.

`import_lake_rasters.R` loads the lake rasters for CSLS lakes (as exported from ArcGIS) into a list, with no transformations of the data. These rasters are clipped to approximately the highest observed lake level (in 2019) plus a spatial buffer of 20m.

Examples

```
lake_raster <- CSLSdata::lake_raster
```

MODFLOW *CSLS MODFLOW model results*

Description

Lake information from MODFLOW models.

Usage

```
data(MODFLOW)
```

Format

A data frame with the following columns:

scenario type of model simulation, e.g. "no_irr", or "cur_irr"

sim simulation number from Monte Carlo runs

lake name of lake, i.e. "Pleasant", "Long", "Plainfield"

date date of output, POSIXct

level_m mean lake stage for the month, mamsl

vol_m3 mean lake volume for the month, (m^3)

P_m3 precipitation volume this month (m^3)

E_m3 evaporation volume this month (m^3)

GWin_m3 groundwater inflow volume this month (m^3)

GWout_m3 groundwater outflow volume this month (m^3)

dV_m3 change in lake volume this month (m^3)

P_m3_d precipitation flow rate (m^3/d)

E_m3_d evaporation flow rate (m^3/d)

GWin_m3_d groundwater inflow rate into lake (m^3/d)

GWout_m3_d groundwater outflow rate into lake (m^3/d)

dV_m3_d change in lake volume as a flow rate (m^3/d)

Details

Raw data is processed in with the function "import_model_results.R" in the "data-raw" subdirectory of this project. This function loads model outputs from csv files and saves as combined data frame with new columns for the lake and simulation type.

Examples

```
MODFLOW <- CSLSdata::MODFLOW
```

plant_limits *CSLS plant water depth limits*

Description

Water depth limits for all plant communities present in the CSLS lakes.

Usage

```
data(plant_limits)
```

Format

A data frame with the following columns:

lake lake, e.g., "Pleasant", "Long", or "Plainfield"

variable plant community, e.g. "upland" or "inland_beach"

shallow the shallowest water depth this plant community can tolerate (m)

deep the deepest water depth this plant community can tolerate (m)

Details

Raw csv data is processed in the `data-raw/` subdirectory of this project with the function `import_plant_limits.R` which is run by `runall_cslsdata.R`.

`import_plant_limits` loads the rules for plant community water depth limits for the CSLS lakes and transforms depths from feet to meters.

See Also

`bathymetry`

Examples

```
plant_limits <- CSLSdata::plant_limits
```

sushi *CSLS raw fish data*

Description

Species, dates, techniques, some age/length/weight data

Usage

```
data(sushi)
```

Format

A list of four data frames:

pleas data for Pleasant Lake

pleasALW age/length/weight data for Pleasant Lake

long data for Long Lake

longALW age/length/weight for Long Lake

Details

Downloaded from the fisheries database, searching by WBIC for each lake.

Raw data is processed in with the script "format_sushi" in the data-raw folder

Examples

```
# Load fish data for Long Lake directly
lake  <- "pleas"
sushi <- CSLSdata::sushi[[lake]]

# Load fish data, then subset for specific lake
data(sushi)
sushi <- sushi$pleas
```

---

water_chem *CSLS Water Chemistry data*

Description

All water chemistry data associated with the Central Sands Lakes Study from:

- The Wisconsin Department of Natural Resources (DNR) Surface Water Integrated Management System (SWIMS), project "Central Sands Lake Study" or "csls".
- The National Atmospheric Deposition Program (NADP) precipitation chemistry data from Devil's Lake.
- Isotope measurements
- HOBO continuous logger measurements

Usage

```
data(water_chem)
```

Format

A data frame with the following columns:

lake name of "lake", either "Pleasant", "Long", "Plainfield", or "Precip

site_type type of site, e.g. "lake", "precipitation", "upgradient", "nogradient", "downgradient", "deep"

site_id unique id for CSLS site

date date of measurement

dnr_parameter DNR parameter code for analyte measured

description Description of analyte measured

result value of sample result

units units of sample result

depth1_m sample depth (m)

depth2_m if sample depth is a range, deepest end of that range (m)

lod limit of detection for lab method

loq limit of quality for lab method

flag notes if data is flagged for being under the LOD ("LOD"), under the LOQ ("LOQ"), analyzed past the holding date ("AGE"), duplicate sample ("DUPLICATE"), blank sample ("BLANK"), other reason ("COMMENT"), bad well ("BAD_WELL") or not flagged ("NONE")

flag_reason longer comment about reason for flag

Details

Raw csv data is processed in the `data-raw/` subdirectory of this project with the functions `import_isotopes`, `import_NADP`, `import_SWIMS.R`, `import_HOBO`, and `combine_chem` which are run by `runall_cslsdata.R`.

`import_isotopes` loads isotope measurements for the CSLS lakes and extracts isotopic measurements for 18O and 2H (duterium) flagged as VALID = TRUE.

`import_NADP` loads in NADP data and cleans with the following steps:

1. Defines SWIMS equivalent of NADP parameters of interest
2. Limits only valid data to the desired site (default: Devil's Lake).
3. Rearranges data frame for parameters of interest
4. Retains any LOD flags, eliminates clearly invalid results (< 0)
5. Converts NADP param names to SWIMS equivalents
6. Add units, date, and site_id

`import_SWIMS.R` loads in SWIMS data associated with the Central Sands Lakes Study (project: "csls") and cleans with the following steps:

1. Transforms column names to characters that are easier to work with in R.
2. Imports and appends LDES data, if using.
3. Uses station_name and WBIC to match samples to CSLS site ids. Makes sure all samples have a CSLS site_id match.
4. Extracts and parses sample depth or sample depth range, where exists.
5. Updates column classes (to datetime, numeric, etc.)
6. Performs QC on results and makes note of any flagged samples
7. Subsets to useful columns and returns SWIMS data frame.

`import_HOBO.R` loads in HOBO data for a lake, cleans up formatting and handling of missing data, and rearranges so that information about the depth of measurements and the parameter measured is treated as data, rather than embedded in column names. It also merges depth information (either depth from bottom or depth from top) with lake level information to get the elevation above mean sea level of each sensor measurement. There is currently no cleaning done to remove measurements when the sensor was out of the water for downloads.

`combine_chem.R` combines all water chemistry data frames, using the dictionary and gw_levels to define site types.

Source

https://dnr.wi.gov/topic/surfacewater/swims/

http://nadp.slh.wisc.edu/data/sites/list/?net=NTN

Examples

```
water_chem <- CSLSdata::water_chem
```

`weather` *CSLS hourly weather data*

Description

Hourly weather data from the Hancock Agricultural Research Station website (https://enviroweather.msu.edu/weather.php?stn=hck) (station id: hck). The Hancock Agricultural Research Station is located in Hancock, WI (location: 44.1188, -89.533, elevation: 241m), approximately 8 miles from Plainfield Lake, 8.5 miles from Long Lake, and 14.5 miles from Pleasant Lake.

Usage

```
data(weather)
```

Format

A data frame with hourly records for:

date date and time of weather observation

atmp air temperature (deg C)

P precipitation (mm)

RH relative humidity (percent)

Rs incoming solar radiation (MJ/m^2)

wind wind speed (m/s)

Details

To download raw data, go to the Hancock Station website and click on "Weather Station at Hancock" or "More weather for this station" to be routed to a page with options for "Custom Reports (Data-on-Demand)". Select Data Type "Hourly Data", then check Air Temperature, Precipitation, Relative Humidity, Total Solar Flux, and Wind Speed. Select starting date and end date, display units (metric), and output (csv), then "Generate Report". Note that the earliest starting date allowed (as of January and September 2019) is February 15, 2018.

Raw data is processed in with the function "import_weather.R" in the "data-raw" subdirectory of this project. This function removes extraneous lines at the start and end of the raw data files, converts the class of data, and linearly interpolates any missing values.

Raw csv data is processed in the `data-raw/` subdirectory of this project with the function `import_weather.R` which is run by `runall_cslsdata.R`.

`import_weather.R` reads in hourly weather from the Hancock, WI weather station. Required weather fields include air temperature (deg C), relative humidity (%), precipitation (mm), solar radiation (MJ/m^2/hr), and wind speed (m/s). This funcition cuts lines known to contain uneeded metadata, updates the format of values, and fills in NA values via linear interpolation.

Source

https://enviroweather.msu.edu/weather.php?stn=hck

Examples

```
# Load weather data (two different options):
weather <- CSLSdata::weather
data(weather)
```

Index