Package 'GSODR'

May 7, 2026

Type Package

Title Global Surface Summary of the Day ('GSOD') Weather Data Client

Version 5.0.1

Description Provides automated downloading, parsing, cleaning, unit conversion and formatting of Global Surface Summary of the Day ('GSOD') weather data from the from the USA National Centers for Environmental Information ('NCEI'). The data were retired on 2025-08-29 and are no longer updated. Units are converted from from United States Customary System ('USCS') units to International System of Units ('SI'). Stations may be individually checked for number of missing days defined by the user, where stations with too many missing observations are omitted. Only stations with valid reported latitude and longitude values are permitted in the final data. Additional useful elements, saturation vapour pressure ('es'), actual vapour pressure ('ea') and relative humidity ('RH') are calculated from the original data using the improved August-Roche-Magnus approximation (Alduchov & Eskridge 1996) and included in the final data set. The resulting metadata include station identification information, country, state, latitude, longitude, elevation, weather observations and associated flags. For information on the 'GSOD' data from 'NCEI', please see the 'GSOD' 'readme.txt' file available from, <https://www.ncei.noaa.gov/pub/data/gsod/readme.txt>.

License MIT + file LICENSE

URL https://docs.ropensci.org/GSODR/, https://codeberg.org/ropensci/GSODR

BugReports https://codeberg.org/ropensci/GSODR/issues

Depends R (>= 3.5.0)

Imports cli, curl, data.table (>= 1.15.4), R.utils, stats, utils, withr

Suggests bit64, dplyr, ggplot2, gridExtra, knitr, mapproj, maps, rmarkdown, roxyglobals, spelling, testthat, tidyr

VignetteBuilder knitr

ByteCompile TRUE

Config/Needs/build maelle/devtag2

Config/roxyglobals/filename globals.R

Config/roxyglobals/unique FALSE

Config/testthat/edition 3

Config/testthat/parallel true

Encoding UTF-8

Language en-US

NeedsCompilation no

Repository CRAN

RoxygenNote 7.3.3

X-schema.org-applicationCategory Tools

X-schema.org-isPartOf https://ropensci.org

X-schema.org-keywords US-NCEI, meteorological-data, global-weather, weather, weather-data, meteorology, station-data, surface-weather, data-access, US-NCDC

Author Adam H. Sparks [aut, cre] (ORCID: <https://orcid.org/0000-0002-0061-8359>), Tomislav Hengl [aut] (ORCID: <https://orcid.org/0000-0002-9921-5129>), Andrew Nelson [aut] (ORCID: <https://orcid.org/0000-0002-7249-3778>), Hugh Parsonage [cph, ctb] (ORCID: <https://orcid.org/0000-0003-4055-0835>), Taras Kaduk [ctb] (Suggestion for handling bulk station downloads more efficiently), Gwenael Giboire [ctb] (Several bug reports in early versions and testing feedback), Łukasz Pawlik [ctb] (Reported bug in windspeed conversion calculation), Ross Darnell [ctb] (Reported bug in 'Windows OS' versions causing 'GSOD' data untarring to fail, ORCID: <https://orcid.org/0000-0002-7973-6322>), Tyler Widdison [ctb] (Reported bug where `nearest_stations()` did not return stations in order of nearest to farthest), Wenbo Lv [ctb] (Provided suggestions for improving documentation), Curtin University [fnd, cph] (url: http://www.curtin.edu.au/, ROR: <https://ror.org/02n415q13>)

Maintainer Adam H. Sparks <adamhsparks@gmail.com>

Date/Publication 2026-02-28 23:40:15 UTC

Contents

`get_GSOD` *Download and Return a data.table Object of GSOD Weather Data*

Description

Automates downloading, cleaning, reformatting of data from the Global Surface Summary of the Day (GSOD) data provided by the US National Centers for Environmental Information (NCEI), Three additional useful elements: saturation vapour pressure (es), actual vapour pressure (ea) and relative humidity (RH) are calculated and returned in the final data frame using the improved August-Roche-Magnus approximation (Alduchov and Eskridge 1996).

Usage

```
get_GSOD(
  years,
  station = NULL,
  country = NULL,
  max_missing = NULL,
  agroclimatology = FALSE
)
```

Arguments

`years`	Year(s) of weather data to download.
`station`	Optional. Specify a station or multiple stations for which to retrieve, check and clean weather data using *STATION*. The NCEI reports years for which the data are available. This function checks against these years. However, not all cases are properly documented and in some cases files may not exist for download even though it is indicated that data was recorded for the station for a particular year. If a station is specified that does not have an existing file on the server, this function will silently fail and move on to existing files for download and cleaning.
`country`	Optional. Specify a country for which to retrieve weather data; full name, 2 or 3 letter ISO or 2 letter FIPS codes can be used. All stations within the specified country will be returned.
`max_missing`	Optional. The maximum number of days allowed to be missing from a station's data before it is excluded from final file output.
`agroclimatology`	Optional. Logical. Only clean data for stations between latitudes 60 and -60 for agroclimatology work, defaults to `FALSE`. Set to `TRUE` to include only stations within the confines of these latitudes.

Details

All units are converted to International System of Units (SI), *e.g*, Fahrenheit to Celsius and inches to millimetres.

Data summarise each year by station, which include vapour pressure and relative humidity elements calculated from existing data in GSOD.

All missing values in resulting files are represented as NA regardless of which field they occur in.

For a complete list of the fields and description of the contents and units, please refer to Appendix 1 in the **GSODR** vignette, `vignette("GSODR", package = "GSODR")`.

For more information see the description of the data provided by NCEI, https://www.ncei.noaa.gov/data/global-summary-of-the-day/doc/readme.txt.

Value

A `data.table::data.table()` object of GSOD weather data.

References

Alduchov, O.A. and Eskridge, R.E., 1996. Improved Magnus form approximation of saturation vapor pressure. Journal of Applied Meteorology and Climatology, 35(4), pp.601-609. doi:10.1175/15200450(1996)035<0601:IMFAOS>2.0.CO;2.

Note

GSODR attempts to validate year and station combination requests, however, in certain cases the start and end date may encompass years where no data is available. In these cases no data will be returned.

While **GSODR** does not distribute GSOD weather data, users of the data should note the conditions that the U.S. NCEI places upon the GSOD data. "The following data and products may have conditions placed on their international commercial use. They can be used within the U.S. or for non-commercial international activities without restriction. The non-U.S. data cannot be redistributed for commercial purposes. Re-distribution of these data by others must provide this same notification. A log of IP addresses accessing these data and products will be maintained and may be made available to data providers."

Author(s)

Adam H. Sparks, <adamhsparks@gmail.com>

See Also

`reformat_GSOD()`

Examples

```
# Download weather station data for Toowoomba, Queensland for 2010
tbar <- get_GSOD(years = 2010, station = "955510-99999")

# Download weather data for the year 1929
```

```
w_1929 <- get_GSOD(years = 1929)

# Download weather data for the year 1929 for Ireland
ie_1929 <- get_GSOD(years = 1929, country = "Ireland")
```

`get_updates` *Get updates.txt With Information on Updates to the GSOD Data Set*

Description

Gets and imports the 'updates.txt' file that has a change log of GSOD data. Changes are shown in order from most recent to oldest changes by the "DATE" field. Column names follow **GSODR** naming conventions.

Usage

```
get_updates()
```

Value

A `data.table::data.table()` object

Examples

```
get_updates()
```

`nearest_stations` *Find Nearest GSOD Stations to a Specified Latitude and Longitude*

Description

Given latitude and longitude values entered as decimal degrees (DD), this function returns a list (as an atomic vector) of station ID values, which can be used in`get_GSOD()` to query for specific stations as an argument in the `station` parameter of that function.

Usage

```
nearest_stations(LAT, LON, distance)
```

Arguments

`LAT`	Latitude expressed as decimal degrees (DD) (WGS84)
`LON`	Longitude expressed as decimal degrees (DD) (WGS84)
`distance`	Distance in kilometres from point for which stations are to be returned.

Value

A data.table::data.table() with full station metadata including the distance from the user specified coordinates from nearest to farthest.

Note

The GSOD data, which are downloaded and manipulated by **GSODR** stipulate that the following notice should be given. "The following data and products may have conditions placed on their international commercial use. They can be used within the U.S. or for non- commercial international activities without restriction. The non-U.S. data cannot be redistributed for commercial purposes. Re-distribution of these data by others must provide this same notification."

Author(s)

Adam H. Sparks, <adamhsparks@gmail.com>

Examples

```
# Find stations within a 100km radius of Toowoomba, QLD, AUS

n <- nearest_stations(LAT = -27.5598, LON = 151.9507, distance = 100)
n
```

---

reformat_GSOD *Tidy and Return a data.table Object of GSOD Data From Local Storage*

Description

This function automates cleaning and reformatting of GSOD station files in "YEAR.tar.gz", provided that they have been untarred or "STATION.csv" format that have been downloaded from the United States National Center for Environmental Information's (NCEI) download page. Three additional useful elements: saturation vapour pressure (es), actual vapour pressure (ea) and relative humidity (RH) are calculated and returned in the final data frame using the improved August-Roche-Magnus approximation (Alduchov and Eskridge 1996). All units are converted to International System of Units (SI), *e.g.*, Fahrenheit to Celsius and inches to millimetres.

Usage

```
reformat_GSOD(dsn = NULL, file_list = NULL)
```

Arguments

dsn	User supplied full file path to location of data files on local disk for tidying.
file_list	User supplied list of file paths to individual files of data on local disk for tidying. Ignored if dsn is set. Use if there are other files in the dsn that you do not wish to reformat.

Details

If multiple stations are given, data are summarised for each year by station, which include vapour pressure and relative humidity elements calculated from existing data in GSOD. Else, a single station is tidied and a data frame is returned.

All missing values in resulting files are represented as NA regardless of which field they occur in.

Only station files in the original "csv" file format are supported by this function. If you have downloaded the full annual ("YYYY.tar.gz") file you will need to extract the individual station files from the tar file first to use this function.

Note that `reformat_GSOD()` will attempt to reformat any ".csv" files found in the `dsn` that you provide. If there are non-GSOD files present this will lead to errors.

For a complete list of the fields and description of the contents and units, please refer to Appendix 1 in the **GSODR** vignette, `vignette("GSODR", package = "GSODR")`.

Value

A data frame as a `data.table::data.table()` object of GSOD data.

References

Alduchov, O.A. and Eskridge, R.E., 1996. Improved Magnus form approximation of saturation vapor pressure. Journal of Applied Meteorology and Climatology, 35(4), pp.601-609. DOI: <10.1175%2F1520-0450%281996%29035%3C0601%3AIMFAOS%3E2.0.CO%3B2>.

Note

While **GSODR** does not distribute GSOD weather data, users of the data should note the conditions that the U.S. NCEI places upon the GSOD data. "The following data and products may have conditions placed on their international commercial use. They can be used within the U.S. or for non-commercial international activities without restriction. The non-U.S. data cannot be redistributed for commercial purposes. Re-distribution of these data by others must provide this same notification. A log of IP addresses accessing these data and products will be maintained and may be made available to data providers."

Author(s)

Adam H. Sparks, <adamhsparks@gmail.com>

See Also

For automated downloading and tidying see the `get_GSOD()` function, which provides expanded functionality for automatically downloading and expanding annual GSOD files and cleaning station files.

`get_GSOD()`

Examples

```
# Download data to 'tempdir()'
download.file(
  url =
   "https://www.ncei.noaa.gov/data/global-summary-of-the-day/access/2010/95551099999.csv",
  destfile = file.path(tempdir(), "95551099999.csv"),
  mode = "wb"
)

# Reformat station data files in R's tempdir() directory
tbar <- reformat_GSOD(dsn = tempdir())

tbar
```

Index