Package 'readnoaa'

November 1, 2026

Title Access 'NOAA' Climate and Weather Data

Version 0.2.2

Description Provides clean, tidy access to climate and weather data from the 'National Oceanic and Atmospheric Administration' ('NOAA') via the 'National Centers for Environmental Information' ('NCEI') Data Service API <https://www.ncei.noaa.gov/support/access-data-service-api-user-documentation>. Covers daily weather observations, monthly and annual summaries, and 30-year climate normals from over 100,000 stations across 180 countries. No API key is required. Dedicated functions handle the most common datasets, while a generic fetcher provides access to all 'NCEI' datasets. Station discovery functions help users find stations by location or name. Data is downloaded on first use and cached locally for subsequent calls. This package is not endorsed or certified by 'NOAA'.

License MIT + file LICENSE

Encoding UTF-8

Language en-US

URL https://charlescoverdale.github.io/readnoaa/, https://github.com/charlescoverdale/readnoaa

BugReports https://github.com/charlescoverdale/readnoaa/issues

Depends R (>= 4.1.0)

Imports cli (>= 3.6.0), httr2 (>= 1.0.0), tools

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

Config/roxygen2/version 8.1.0

NeedsCompilation no

Author Charles Coverdale [aut, cre, cph]

Maintainer Charles Coverdale <charlesfcoverdale@gmail.com>

Repository CRAN

Date/Publication 2026-11-01 15:20:08 UTC

Contents

cache_info . . . 2
clear_cache . . . 3
list_datasets . . . 3
list_datatypes . . . 4
noaa_annual . . . 5
noaa_coverage . . . 7
noaa_daily . . . 8
noaa_get . . . 10
noaa_monthly . . . 11
noaa_nearby . . . 13
noaa_normals . . . 14
noaa_stations . . . 16

Index **18**

`cache_info` *Inspect the readnoaa cache*

Description

Lists the cached responses currently on disk, with their size and age. Useful for checking whether a result is being served from a stale copy.

Usage

```
cache_info()
```

Details

Cached responses expire automatically. Requests whose window ends within the last five weeks are treated as provisional and expire after one day, because NCEI publishes recent observations with a lag and continues to revise them. Older windows are treated as settled and expire after 30 days. Both thresholds are configurable through the options `readnoaa.cache_days_recent` and `readnoaa.cache_days`, and any single call can bypass the cache entirely with `refresh = TRUE`.

Value

A data frame with columns `file`, `size_kb`, and `age_days`, returned invisibly if the cache is empty.

See Also

Other data access: `clear_cache()`, `list_datasets()`, `list_datatypes()`, `noaa_get()`

Examples

```
op <- options(readnoaa.cache_dir = tempdir())
cache_info()
options(op)
```

`clear_cache` *Clear the readnoaa cache*

Description

Deletes all locally cached NOAA data files. The next call to any data function will re-download from the NCEI API.

Usage

```
clear_cache()
```

Value

Invisible `NULL`.

See Also

`cache_info()` to inspect the cache before clearing it.

Other data access: `cache_info()`, `list_datasets()`, `list_datatypes()`, `noaa_get()`

Examples

```
op <- options(readnoaa.cache_dir = tempdir())
clear_cache()
options(op)
```

`list_datasets` *List common NCEI datasets*

Description

Returns a curated table of the most commonly used NCEI datasets. No network call is made.

Usage

```
list_datasets()
```

Details

The `requires` column records constraints the API enforces: some datasets reject requests that carry no date window, and `global-marine` is organised by area rather than by station, so it requires a bounding box.

Value

A data frame with columns:

dataset Character. Dataset identifier for use with `noaa_get()`.

description Character. Brief description.

frequency Character. Temporal resolution.

requires Character. Arguments the API requires, or `""`.

See Also

Other data access: `cache_info()`, `clear_cache()`, `list_datatypes()`, `noaa_get()`

Examples

```
list_datasets()
```

`list_datatypes` *List available data types for a dataset*

Description

Reports the element codes a station actually records.

Usage

```
list_datatypes(
  dataset,
  station,
  start_date = NULL,
  end_date = NULL,
  cache = TRUE
)
```

Arguments

`dataset`	Character. Dataset identifier (e.g. `"daily-summaries"`).
`station`	Character. A station ID to query.
`start_date, end_date`	Optional character dates. For the daily datasets these restrict the result to elements recorded during the window; for other datasets they bound the sample request.
`cache`	Logical. Use cached data if available (default `TRUE`).

Details

For the daily datasets this reads the GHCN-Daily element inventory, which states exactly which elements a station reports and over what years. That inventory file is around 36 MB, downloaded on first use and cached thereafter.

For other datasets, where no such inventory exists, a short sample request is made and the columns that came back with data are reported. The sample window is taken from the end of the requested range, or from the recent past when no range is given.

A station's element list spans its entire history, and stations routinely stop recording some variables. Supply `start_date` and `end_date` to see only the elements whose record overlaps the period you care about.

Value

A character vector of available data type codes.

See Also

`noaa_coverage()` for the years each element spans.

Other data access: `cache_info()`, `clear_cache()`, `list_datasets()`, `noaa_get()`

Examples

```
op <- options(readnoaa.cache_dir = tempdir())
try({
  # Everything Central Park has ever recorded
  list_datatypes("daily-summaries", "USW00094728")

  # Only what it still records
  list_datatypes("daily-summaries", "USW00094728", start_date = "2025-01-01")
})
options(op)
```

---

noaa_annual *Annual weather summaries*

Description

Returns annual summary data from the NCEI Global Summary of the Year dataset.

Usage

```
noaa_annual(
  station,
  start_date,
  end_date,
  datatypes = NULL,
```

```
  units = "metric",
  include_flags = FALSE,
  include_location = FALSE,
  drop_empty = is.null(datatypes),
  cache = TRUE,
  refresh = FALSE
)
```

Arguments

station	Character. One or more station IDs.
start_date	Character. Start date in "YYYY-MM-DD" or "YYYY-MM" format.
end_date	Character. End date in the same format.
datatypes	Optional character vector of data type codes.
units	Character. "metric" (default) or "standard".
include_flags	Logical. Include data quality flags from NCEI (default FALSE).
include_location	Logical. Include station latitude, longitude, and elevation columns (default FALSE).
drop_empty	Logical. Drop columns that contain no data at all. Defaults to TRUE when datatypes is NULL.
cache	Logical. Use cached data if available (default TRUE).
refresh	Logical. Ignore any cached copy and refetch (default FALSE).

Value

A data frame with columns including:

station Character. Station identifier.

date Date. First day of the year.

name Character. Station name.

... Numeric. Data columns vary by station and request.

See Also

Other weather data: noaa_daily(), noaa_monthly(), noaa_normals()

Examples

```
op <- options(readnoaa.cache_dir = tempdir())
try({
  noaa_annual("USW00094728", "2020-01-01", "2024-01-01")
})
options(op)
```

noaa_coverage *Check what a station records, and for how long*

Description

Returns the first and last year of data for each element a station reports, taken from the GHCN-Daily element inventory. This is the authoritative answer to two questions the data functions cannot answer on their own: which variables a station actually measures, and how current its record is.

Usage

```
noaa_coverage(station, element = NULL, cache = TRUE, refresh = FALSE)
```

Arguments

station	Character. One or more station IDs.
element	Optional character vector of element codes (e.g. `c("TMAX", "PRCP")`) to restrict the result to.
cache	Logical. Use the cached inventory if available (default `TRUE`).
refresh	Logical. Ignore any cached copy and refetch (default `FALSE`).

Details

Currency varies widely. United States stations are typically complete to within a few days, while many international stations lag by months, and others stopped reporting years ago while remaining in the station list. Checking coverage first avoids requesting a window a station never covered and receiving an empty result.

The inventory file is around 36 MB. It is downloaded on first use and cached locally thereafter.

Value

A data frame with columns:

station Character. Station identifier.

element Character. Element code, e.g. `"TMAX"`.

first_year Integer. First year with data.

last_year Integer. Most recent year with data.

years Integer. Length of the record in years.

See Also

Other station discovery: `noaa_nearby()`, `noaa_stations()`

Examples

```
op <- options(readnoaa.cache_dir = tempdir())
try({
  # What does Central Park record, and through when?
  noaa_coverage("USW00094728", element = c("TMAX", "TMIN", "PRCP"))
})
options(op)
```

---

noaa_daily *Daily weather observations*

Description

Returns daily weather data from the NCEI Daily Summaries dataset (GHCN-Daily). Common data types include TMAX (maximum temperature), TMIN (minimum temperature), PRCP (precipitation), SNOW (snowfall), and SNWD (snow depth).

Usage

```
noaa_daily(
  station,
  start_date,
  end_date,
  datatypes = NULL,
  units = "metric",
  include_flags = FALSE,
  include_location = FALSE,
  drop_empty = is.null(datatypes),
  cache = TRUE,
  refresh = FALSE
)
```

Arguments

station	Character. One or more station IDs (e.g. "USW00094728" for Central Park, NYC).
start_date	Character. Start date in "YYYY-MM-DD" or "YYYY-MM" format.
end_date	Character. End date in the same format.
datatypes	Optional character vector of data type codes to retrieve (e.g. c("TMAX", "TMIN")). If NULL, all available types are returned.
units	Character. "metric" (default, Celsius/mm) or "standard" (Fahrenheit/inches).
include_flags	Logical. Include data quality flags from NCEI (default FALSE).
include_location	Logical. Include station latitude, longitude, and elevation columns (default FALSE).

drop_empty	Logical. Drop columns that contain no data at all. Defaults to TRUE when datatypes is NULL, because an unfiltered request returns the whole GHCN-Daily element set as columns and few stations report more than a handful of them. When you name datatypes explicitly, every requested column is kept.
cache	Logical. Use cached data if available (default TRUE).
refresh	Logical. Ignore any cached copy and refetch (default FALSE).

Details

Requests spanning more than one year are automatically split into yearly chunks to avoid API timeouts.

Station records vary enormously in how current they are. United States stations are typically complete to within a few days, while many international stations lag by months or have stopped reporting altogether. Use noaa_coverage() to check a station's record before relying on it.

Value

A data frame with columns including:

station Character. Station identifier.

date Date. Observation date.

name Character. Station name.

... Numeric. Data columns vary by station and request (e.g. tmax, tmin, prcp).

See Also

noaa_coverage() to check which elements a station reports and how current its record is.

Other weather data: noaa_annual(), noaa_monthly(), noaa_normals()

Examples

```
op <- options(readnoaa.cache_dir = tempdir())
try({
  # Daily temperatures for Central Park, NYC
  noaa_daily("USW00094728", "2024-01-01", "2024-01-31",
  datatypes = c("TMAX", "TMIN"))
})
options(op)
```

noaa_get *Fetch any NCEI dataset*

Description

A generic fetcher for direct access to any NCEI dataset. Use `list_datasets()` to see common dataset identifiers and the arguments each one requires.

Usage

```
noaa_get(
  dataset,
  station = NULL,
  start_date = NULL,
  end_date = NULL,
  datatypes = NULL,
  bbox = NULL,
  units = "metric",
  include_flags = FALSE,
  include_location = FALSE,
  drop_empty = is.null(datatypes),
  cache = TRUE,
  refresh = FALSE
)
```

Arguments

`dataset`	Character. The dataset identifier (e.g. `"daily-summaries"`, `"global-summary-of-the-month"`).
`station`	Optional character vector of station IDs. Note that station identifiers are dataset-specific: the daily datasets use GHCN-Daily IDs such as `"USW00094728"`, while `global-hourly` and `global-summary-of-the-day` use ISD IDs such as `"72505394728"`.
`start_date`	Optional start date in `"YYYY-MM-DD"` or `"YYYY-MM"` format.
`end_date`	Optional end date in the same format.
`datatypes`	Optional character vector of data type codes.
`bbox`	Optional numeric vector of length 4 defining a bounding box: `c(south_lat, west_lon, north_lat, east_lon)`. Required by `global-marine`.
`units`	Character. `"metric"` (default) or `"standard"`. The normals datasets are always requested in their published US customary units: asked for metric, the API converts only some of their variables, and converts temperature ranges as if they were temperatures.
`include_flags`	Logical. Include data quality flags from NCEI (default `FALSE`).
`include_location`	Logical. Include station latitude, longitude, and elevation columns (default `FALSE`).

drop_empty	Logical. Drop columns that contain no data at all. Defaults to TRUE when datatypes is NULL.
cache	Logical. Use cached data if available (default TRUE).
refresh	Logical. Ignore any cached copy and refetch (default FALSE).

Value

A data frame. Columns vary by dataset. The date column is a Date for daily, monthly, and annual datasets, and a POSIXct in UTC for the hourly datasets, which publish ISO 8601 timestamps.

See Also

Other data access: cache_info(), clear_cache(), list_datasets(), list_datatypes()

Examples

```
op <- options(readnoaa.cache_dir = tempdir())
try({
  # Fetch daily data using the generic function
  noaa_get("daily-summaries", station = "USW00094728",
  start_date = "2024-01-01", end_date = "2024-01-31")
})
options(op)
```

---

noaa_monthly *Monthly weather summaries*

Description

Returns monthly summary data from the NCEI Global Summary of the Month dataset.

Usage

```
noaa_monthly(
  station,
  start_date,
  end_date,
  datatypes = NULL,
  units = "metric",
  include_flags = FALSE,
  include_location = FALSE,
  drop_empty = is.null(datatypes),
  cache = TRUE,
  refresh = FALSE
)
```

Arguments

station	Character. One or more station IDs.
start_date	Character. Start date in "YYYY-MM-DD" or "YYYY-MM" format.
end_date	Character. End date in the same format.
datatypes	Optional character vector of data type codes.
units	Character. "metric" (default) or "standard".
include_flags	Logical. Include data quality flags from NCEI (default FALSE).
include_location	Logical. Include station latitude, longitude, and elevation columns (default FALSE).
drop_empty	Logical. Drop columns that contain no data at all. Defaults to TRUE when datatypes is NULL.
cache	Logical. Use cached data if available (default TRUE).
refresh	Logical. Ignore any cached copy and refetch (default FALSE).

Value

A data frame with columns including:

station Character. Station identifier.

date Date. First day of the month.

name Character. Station name.

... Numeric. Data columns vary by station and request.

See Also

Other weather data: noaa_annual(), noaa_daily(), noaa_normals()

Examples

```
op <- options(readnoaa.cache_dir = tempdir())
try({
  noaa_monthly("USW00094728", "2024-01", "2024-12")
})
options(op)
```

noaa_nearby *Find stations near a location*

Description

Searches for weather stations within a given radius of a point, sorted by distance. Uses the GHCN-Daily station inventory.

Usage

```
noaa_nearby(
  lat,
  lon,
  radius_km = 50,
  element = NULL,
  active_since = NULL,
  limit = 25L,
  cache = TRUE,
  refresh = FALSE
)
```

Arguments

lat	Numeric. Latitude of the target location.
lon	Numeric. Longitude of the target location.
radius_km	Numeric. Search radius in kilometres (default 50).
element	Optional character vector of element codes (e.g. "TMAX"). Only stations reporting all of them are returned.
active_since	Optional integer year. Only stations whose record extends to that year or later are returned.
limit	Integer. Maximum number of results (default 25). Use Inf for no limit.
cache	Logical. Use cached station list if available (default TRUE).
refresh	Logical. Ignore any cached copy and refetch (default FALSE).

Value

A data frame with the same columns as noaa_stations() plus:

distance_km Numeric. Distance from the target point in kilometres.

See Also

Other station discovery: noaa_coverage(), noaa_stations()

Examples

```
op <- options(readnoaa.cache_dir = tempdir())
try({
  # Stations within 25 km of central London
  noaa_nearby(51.5, -0.1, radius_km = 25)

  # Only those still reporting maximum temperature recently
  noaa_nearby(51.5, -0.1, radius_km = 25,
  element = "TMAX", active_since = 2024)
})
options(op)
```

---

noaa_normals *Climate normals (1991-2020)*

Description

Returns 30-year climate normals from the NCEI Normals datasets. Normals are the average conditions over the 1991-2020 reference period, and are used as the baseline against which current weather is compared.

Usage

```
noaa_normals(
  station,
  period = "monthly",
  datatypes = NULL,
  start_date = NULL,
  end_date = NULL,
  include_flags = FALSE,
  include_location = FALSE,
  drop_empty = is.null(datatypes),
  cache = TRUE,
  refresh = FALSE
)
```

Arguments

station	Character. One or more station IDs.
period	Character. One of "monthly", "daily", "hourly", or "annual".
datatypes	Optional character vector of data type codes.
start_date, end_date	Optional character dates bounding the window for the "daily" and "hourly" periods. Ignored for "monthly" and "annual", which cover the whole year by construction.
include_flags	Logical. Include data quality flags from NCEI (default FALSE).

include_location
: Logical. Include station latitude, longitude, and elevation columns (default FALSE).

drop_empty
: Logical. Drop columns that contain no data at all. Defaults to TRUE when datatypes is NULL.

cache
: Logical. Use cached data if available (default TRUE).

refresh
: Logical. Ignore any cached copy and refetch (default FALSE).

Details

Four periods are available, each backed by a different NCEI dataset:

"monthly" normals-monthly-1991-2020. Twelve rows per station.

"daily" normals-daily-1991-2020. One row per day of the year.

"hourly" normals-hourly-1991-2020. One row per hour of the year.

"annual" normals-annualseasonal-1991-2020. One row per station, covering annual and seasonal statistics.

The daily and hourly datasets require a date window, which is supplied automatically as a full calendar year unless you narrow it with start_date and end_date. Because normals are climatological rather than tied to a particular year, only the month and day of those arguments are meaningful.

Normals carry a climatological pseudo-date rather than a calendar date: "01" for a month, "01-31" for a day of the year. These are returned verbatim in date, with integer month, day, and hour columns added alongside for filtering and joining. The annual and seasonal dataset has no date column at all.

Value

A data frame. Columns vary by period, but typically include station, a climatological date with derived month/day/hour columns, and normal values for temperature, precipitation, and other variables.

Units

The NCEI normals datasets are published in United States customary units (degrees Fahrenheit, inches). Asked for metric units, the API converts them inconsistently: precipitation and monthly maximum and minimum temperature are converted, monthly average temperature is not, and temperature ranges are converted as if they were temperatures (giving negative diurnal ranges). This function therefore always requests the published units and, unlike the observational functions, has no metric option. Convert after the fact if you need Celsius or millimetres.

See Also

Other weather data: noaa_annual(), noaa_daily(), noaa_monthly()

Examples

```
op <- options(readnoaa.cache_dir = tempdir())
try({
  # Monthly normals: twelve rows, values in Fahrenheit and inches
  noaa_normals("USW00094728", "monthly")

  # Daily normals for January only
  noaa_normals("USW00094728", "daily",
  start_date = "2020-01-01", end_date = "2020-01-31")
})
options(op)
```

`noaa_stations` *Search for weather stations*

Description

Searches the GHCN-Daily station inventory by bounding box or text query. The station list (~130,000 stations worldwide) is downloaded once and cached locally.

Usage

```
noaa_stations(
  bbox = NULL,
  text = NULL,
  element = NULL,
  active_since = NULL,
  regex = FALSE,
  limit = 25L,
  cache = TRUE,
  refresh = FALSE
)
```

Arguments

`bbox`	Optional numeric vector of length 4 defining a bounding box: `c(south_lat, west_lon, north_lat, east_lon)`.
`text`	Optional character string to search station names (case-insensitive). Matched literally unless `regex = TRUE`.
`element`	Optional character vector of element codes (e.g. `"TMAX"`). Only stations reporting all of them are returned.
`active_since`	Optional integer year. Only stations whose record extends to that year or later are returned. When `element` is also given, the test applies to those elements.
`regex`	Logical. Treat `text` as a regular expression rather than a literal string (default `FALSE`).

limit	Integer. Maximum number of results (default 25). Use Inf for no limit.
cache	Logical. Use cached station list if available (default TRUE).
refresh	Logical. Ignore any cached copy and refetch (default FALSE).

Details

Set element or active_since to restrict results to stations that actually report a given variable, or that were still reporting recently. Both draw on the GHCN-Daily element inventory, an additional file of around 36 MB that is downloaded on first use and cached thereafter.

Value

A data frame with columns:

station Character. Station identifier.

name Character. Station name.

latitude Numeric. Latitude in decimal degrees.

longitude Numeric. Longitude in decimal degrees.

elevation Numeric. Elevation in metres, or NA where GHCN-Daily records none.

state Character. US state or Canadian province, where applicable.

gsn_flag Character. "GSN" for GCOS Surface Network stations.

hcn_crn_flag Character. "HCN" or "CRN" for US Historical Climatology Network and Climate Reference Network stations.

wmo_id Character. Five-digit WMO identifier, where assigned.

See Also

noaa_coverage() for the full element record of a station.

Other station discovery: noaa_coverage(), noaa_nearby()

Examples

```
op <- options(readnoaa.cache_dir = tempdir())
try({
  # Search for stations in the London area
  noaa_stations(bbox = c(51.3, -0.5, 51.7, 0.3))

  # Search by name
  noaa_stations(text = "Heathrow")
})
options(op)
```

Index

∗ **data access**
  cache_info, 2
  clear_cache, 3
  list_datasets, 3
  list_datatypes, 4
  noaa_get, 10
∗ **station discovery**
  noaa_coverage, 7
  noaa_nearby, 13
  noaa_stations, 16
∗ **weather data**
  noaa_annual, 5
  noaa_daily, 8
  noaa_monthly, 11
  noaa_normals, 14

cache_info, 2
cache_info(), *3–5*, *11*
clear_cache, 3
clear_cache(), *2*, *4*, *5*, *11*

list_datasets, 3
list_datasets(), *2*, *3*, *5*, *10*, *11*
list_datatypes, 4
list_datatypes(), *2–4*, *11*

noaa_annual, 5
noaa_annual(), *9*, *12*, *15*
noaa_coverage, 7
noaa_coverage(), *5*, *9*, *13*, *17*
noaa_daily, 8
noaa_daily(), *6*, *12*, *15*
noaa_get, 10
noaa_get(), *2–5*
noaa_monthly, 11
noaa_monthly(), *6*, *9*, *15*
noaa_nearby, 13
noaa_nearby(), *7*, *17*
noaa_normals, 14
noaa_normals(), *6*, *9*, *12*
noaa_stations, 16
noaa_stations(), *7*, *13*