

# Package ‘climateBR’

August 19, 2026

**Type** Package

**Title** Download Rainfall, Temperature, and Wind Data from Brazil

**Version** 0.2.0

**Description** Provides functions to download and import meteorological data from Brazil's National Institute of Meteorology (INMET)  
<<https://portal.inmet.gov.br>>.

**License** MIT + file LICENSE

**Encoding** UTF-8

**LazyData** true

**LazyDataCompression** xz

**RoxygenNote** 7.3.3

**Imports** glue, arrow, dplyr, tidyr, janitor, data.table, pracma, purrr,  
tibble

**Suggests** knitr, rmarkdown, ggplot2, patchwork, geobr

**Depends** R (>= 4.1.0), rlang, sf, gstat

**VignetteBuilder** knitr

**URL** <https://github.com/kaiorb52/climateBR>

**BugReports** <https://github.com/kaiorb52/climateBR/issues>

**NeedsCompilation** no

**Author** Kaio Bárbara [aut, cre, cph] (ORCID:  
<<https://orcid.org/0009-0003-4884-7980>>)

**Maintainer** Kaio Bárbara <[kaio.rbarbara@ufpe.br](mailto:kaio.rbarbara@ufpe.br)>

**Repository** CRAN

**Date/Publication** 2026-08-19 13:40:09 UTC

## Contents

|                                 |    |
|---------------------------------|----|
| build_inmet_dataset . . . . .   | 2  |
| download_inmet . . . . .        | 3  |
| floods_rs . . . . .             | 4  |
| inmet_stations . . . . .        | 5  |
| kriging_inmet . . . . .         | 6  |
| municipality . . . . .          | 7  |
| mun_stations . . . . .          | 8  |
| mun_stations_distance . . . . . | 9  |
| nearest_stations . . . . .      | 10 |
| rain_stations . . . . .         | 11 |
| read_inmet . . . . .            | 12 |

|              |           |
|--------------|-----------|
| <b>Index</b> | <b>14</b> |
|--------------|-----------|

---

|                     |                                                               |
|---------------------|---------------------------------------------------------------|
| build_inmet_dataset | <i>Build a partitioned Arrow dataset from INMET CSV files</i> |
|---------------------|---------------------------------------------------------------|

---

## Description

Converts the raw CSV files downloaded from INMET into a partitioned Arrow/Parquet dataset optimized for fast querying with [read\_inmet()].

## Usage

```
build_inmet_dataset(
  input,
  output,
  years = 2000:2026,
  partitioning_by = c("ano", "codigo_wmo"),
  progress = TRUE
)
```

## Arguments

|                 |                                                                                                                                                        |
|-----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| input           | Character. Directory containing the raw CSV files downloaded with [download_inmet()].                                                                  |
| output          | Character. Directory where the partitioned Arrow/Parquet dataset will be written.                                                                      |
| years           | Vector. Group of years of the INMET database located in 'input' to be transformed into an Arrow/Parquet dataset.                                       |
| partitioning_by | Vector. Variable(s) in the INMET database used to create the parquet folders.                                                                          |
| progress        | Logical. Should a progress bar be displayed while the INMET files are being processed? Defaults to 'TRUE'. Set to 'FALSE' to disable the progress bar. |

**Details**

During the conversion, metadata are extracted from each file, column names are standardized, numeric variables are converted to numeric format, and the resulting dataset is partitioned by year ('ano') and WMO station code ('codigo\_wmo').

This function only needs to be executed once for a collection of downloaded INMET files. After the dataset has been created, it can be accessed efficiently using [read\_inmet()] without repeatedly parsing the original CSV files.

The resulting dataset is partitioned by year ('ano') and weather station ('codigo\_wmo'), allowing Arrow to read only the files required by a query.

**Value**

Invisibly returns the output directory.

**See Also**

[download\_inmet()], [read\_inmet()]

**Examples**

```
build_inmet_dataset(
  input = file.path(tempdir(), "inmet_raw"),
  output = file.path(tempdir(), "inmet_arrow"),
  progress = FALSE
)
```

---

download\_inmet

*Download historical meteorological data from INMET*


---

**Description**

Downloads historical weather station data from the Brazilian National Institute of Meteorology (INMET) and extracts the downloaded ZIP files into a user-specified directory.

**Usage**

```
download_inmet(years = 2008, unzip_to = tempdir(), progress = TRUE)
```

**Arguments**

|          |                                                                                                                                                        |
|----------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| years    | Integer vector specifying the years to download. Historical data are available from 2000 onwards. The default is '2008'.                               |
| unzip_to | Character. Directory where the downloaded files will be extracted.                                                                                     |
| progress | Logical. Should a progress bar be displayed while the INMET files are being processed? Defaults to 'TRUE'. Set to 'FALSE' to disable the progress bar. |

### Details

INMET provides historical observations dating back to 2000. However, only a small number of weather stations were operating in the early years of the dataset. For most applications, we recommend using data from **\*\*2008 onwards\*\***, when the monitoring network became substantially more comprehensive.

Existing directories containing extracted files are skipped to avoid downloading the same data multiple times.

The downloaded files can subsequently be processed with `[build_inmet_dataset()]`.

### Value

This function is called for its side effects. ZIP files are downloaded, extracted into ‘unzip\_to’, and the extracted files are stored on disk.

### See Also

`[build_inmet_dataset()]`, `[read_inmet()]`

### Examples

```
## Download a single year
download_inmet(
  years = 2000,
  unzip_to = tempdir()
)

## Download multiple years
download_inmet(
  years = 2004:2006,
  unzip_to = tempdir()
)
```

---

floods\_rs

*Rainfall during the 2024 Rio Grande do Sul floods*


---

### Description

Dataset containing accumulated rainfall observed at meteorological stations operated by the Brazilian National Institute of Meteorology (INMET) during the extreme flooding event that affected Rio Grande do Sul, Brazil, in 2024.

### Usage

```
floods_rs
```

Format

A data frame with 545 rows and 4 variables:

- code\_wmo** Character. World Meteorological Organization (WMO) identifier of the INMET weather station.
- lat** Numeric. Latitude of the station in decimal degrees (WGS84).
- lon** Numeric. Longitude of the station in decimal degrees (WGS84).
- total\_rainfall** Numeric. Total accumulated rainfall (mm) during the study period.

Details

The dataset contains 545 monitoring stations distributed across Brazil. Each row corresponds to a single INMET weather station and includes its identification code, geographic coordinates, and the total accumulated rainfall (in millimeters) recorded between April 27 and May 5, 2024.

Source

Brazilian National Institute of Meteorology (INMET).

Examples

```
head(floods_rs)
summary(floods_rs$total_rainfall)
```

---

|                |                                           |
|----------------|-------------------------------------------|
| inmet_stations | <i>INMET rainfall monitoring stations</i> |
|----------------|-------------------------------------------|

---

Description

Dataset containing metadata for rainfall monitoring stations operated by the Brazilian National Institute of Meteorology (INMET). Each row represents a specific meteorological station collect based on data from 2000 to 2026.

Usage

```
inmet_stations
```

Format

A data frame with 700 rows and 6 variables:

- state\_station** Brazilian state abbreviation.
- code\_wmo** WMO station identifier.
- lat** Station Latitude in decimal degrees (WGS84).
- lon** Station Longitude in decimal degrees (WGS84).
- creation\_year** First year with available observations for the station.
- last\_used\_year** Last year with available observations for the station.

**Details**

The dataset includes station identifiers, location information, state, municipality codes, and the first and last years in which data are available for each station.

**Source**

Instituto Nacional de Meteorologia (INMET).

**Examples**

```
data(inmet_stations)
head(inmet_stations)
```

---

|               |                                                                     |
|---------------|---------------------------------------------------------------------|
| kriging_inmet | <i>Perform ordinary kriging interpolation of INMET observations</i> |
|---------------|---------------------------------------------------------------------|

---

**Description**

Interpolates meteorological observations from INMET weather stations using ordinary kriging and predicts values for a set of target geometries, such as Brazilian municipalities.

**Usage**

```
kriging_inmet(stations_df, mun_geo, var = "total_rainfall")
```

**Arguments**

|             |                                                                                                                                                                         |
|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| stations_df | An ‘sf’ object containing weather station observations. The object must include point geometries and a numeric column corresponding to the variable specified in ‘var’. |
| mun_geo     | An ‘sf’ object containing the target geometries where predictions will be generated.                                                                                    |
| var         | Character. Name of the numeric variable to interpolate. Defaults to “total_rainfall”.                                                                                   |

**Details**

The empirical variogram is estimated with [gstat::variogram()] and a spherical variogram model is fitted using [gstat::fit.variogram()]. Ordinary kriging is then performed with [gstat::krige()].

Both ‘stations\_df’ and ‘mun\_geo’ must use the same projected coordinate reference system (CRS). Using geographic coordinates (longitude/latitude) is not recommended for kriging because distance calculations are performed in map units.

**Value**

An ‘sf’ object containing the geometries from ‘mun\_geo’ together with the kriging predictions:

\* ‘var1.pred’ - Predicted values. \* ‘var1.var’ - Prediction variance.

**See Also**

[gstat::krige()], [gstat::variogram()], [gstat::fit.variogram()]

**Examples**

```
# Requires spatial data (e.g., municipal boundaries) together with
# INMET stations observations. The example dataset `floods_rs`
# illustrates the required input format for the `stations_df` parameter.
# See the vignette "Spatial Interpolation Using Ordinary Kriging"
# for the complete workflow of this function.

## Not run:

krig_df <- kriging_inmet(
  stations_df = inmet_data,
  mun_geo = municipalities_sf,
  var = "total_rainfall"
)

head(krig_df)

## End(Not run)
```

---

municipality

*Municipality Database*


---

**Description**

Municipality Database

**Usage**

```
municipality
```

**Format**

A data frame with the following variables:

**state\_muni** Brazilian state abbreviation.

**code\_ibge7** Seven-digit IBGE municipality code.

**code\_ibge6** Six-digit IBGE municipality code.

**code\_tse** Six-digit IBGE municipality code.

**lat** Municipality Centroid Latitude in decimal degrees (WGS84).

**lon** Municipality Centroid Longitude in decimal degrees (WGS84).

## Examples

```
data(municipality)
head(municipality)
```

---

mun\_stations

*Nearest INMET Weather Stations for Brazilian Municipalities*


---

## Description

A dataset containing the five nearest INMET weather stations for each Brazilian municipality based on the distance between the municipality centroid and the station location.

## Usage

```
mun_stations
```

## Format

A tibble with 27,850 rows and 4 variables:

**state\_muni** Brazilian state abbreviation.

**code\_ibge7** Seven-digit IBGE municipality code.

**code\_wmo** World Meteorological Organization (WMO) identifier of the INMET weather station.

**distance** Distance between the municipality centroid and the station, in kilometers.

**station\_order** Rank of the station by proximity, where 1 indicates the nearest station.

## Details

Distances were calculated using the Haversine formula. Each municipality is associated with its five closest INMET stations, ordered by increasing distance.

The dataset was generated using the ‘nearest\_stations()’ function from the climateBR package, matching municipality centroids to the five closest INMET weather stations available in the package’s ‘inmet\_stations’ dataset.

This dataset is based on the most recent version of the INMET station network available in the package. It is intended for analyses using recent observations. Because the INMET network changes over time (stations may be added or removed), this dataset should not be used to match historical data from earlier years. For historical analyses, users should create a year-specific municipality-to-station mapping using the corresponding historical INMET station

## Source

Distances were computed from municipality centroid coordinates and INMET station coordinates using the Haversine formula.

### Examples

```
data(mun_stations)
head(mun_stations)
```

---

|                       |                                                                                   |
|-----------------------|-----------------------------------------------------------------------------------|
| mun_stations_distance | <i>Distance Between Brazilian Municipal Centroids and INMET Rainfall Stations</i> |
|-----------------------|-----------------------------------------------------------------------------------|

---

### Description

A dataset containing the distances between the centroids of Brazilian municipalities and rainfall stations operated by the Brazilian National Institute of Meteorology (INMET). Distances were computed using the Haversine formula.

### Usage

```
mun_stations_distance
```

### Format

A data frame with the following variables:

**code\_muni** Seven-digit IBGE municipality code.

**codigo\_wmo** WMO identifier of the INMET rainfall station.

**distance** Distance between the municipality centroid and the station, in kilometers.

**i** Rank of the station by distance, where 1 indicates the nearest station.

**ano** Reference year (2008, 2010, ..., 2024).

### Details

The dataset covers the period from 2008 to 2024. Because the INMET station network changes over time, distances were calculated for snapshots taken every two years (2008, 2010, 2012, 2014, 2016, 2018, 2020, 2022, 2024).

Each municipality is associated with all available INMET stations for the corresponding year, ordered by increasing distance. The variable 'i' indicates the rank of the station according to its proximity to the municipality centroid.

### Source

Distances computed from municipality centroids and INMET rainfall station coordinates using the Haversine formula.

### Examples

```
data(mun_stations_distance)
head(mun_stations_distance)
```

---

|                  |                                                              |
|------------------|--------------------------------------------------------------|
| nearest_stations | <i>Find the nearest INMET stations for each municipality</i> |
|------------------|--------------------------------------------------------------|

---

### Description

Computes the great-circle distance between each municipality and all INMET stations, returning the ‘n’ nearest stations for every municipality.

### Usage

```
nearest_stations(municipality, inmet_stations, n = 5)
```

### Arguments

|                |                                                                                                                |
|----------------|----------------------------------------------------------------------------------------------------------------|
| municipality   | A data frame containing the municipality coordinates. Must include the columns ‘code_ibge7’, ‘lat’, and ‘lon’. |
| inmet_stations | A data frame containing the INMET station coordinates. Must include the columns ‘code_wmo’, ‘lat’, and ‘lon’.  |
| n              | Number of nearest stations to return for each municipality.                                                    |

### Value

A tibble with the columns:

**state\_muni** Brazilian state abbreviation.

**code\_ibge7** Municipality code.

**code\_wmo** INMET station code.

**distance** Distance (in meters) between the municipality and station.

**station\_order** Rank of the station by distance, where 1 indicates the nearest station.

### Examples

```
data("municipality")
data("inmet_stations")

mun_rs <- municipality[municipality$state_muni == "RS", ]

mun_stations1 <- nearest_stations(
  municipality = mun_rs,
  inmet_stations = inmet_stations,
  n = 1
)

# It is fine to include stations from other states. This is expected,
# as stations from neighboring states may be closer to municipalities
# near state borders.
```

```
mun_stations2 <- nearest_stations(  
  municipality = mun_rs,  
  inmet_stations = inmet_stations,  
  n = 5  
)
```

---

|               |                                                   |
|---------------|---------------------------------------------------|
| rain_stations | <i>INMET rainfall monitoring stations by year</i> |
|---------------|---------------------------------------------------|

---

**Description**

Dataset containing metadata for rainfall monitoring stations operated by the Brazilian National Institute of Meteorology (INMET). Each row represents a meteorological station in a specific year between 2000 and 2024.

**Usage**

rain\_stations

**Format**

A data frame with 9,459 rows and 11 variables:

- uf** Brazilian state abbreviation.
- estacao** Name of the meteorological station.
- codigo\_wmo** WMO station identifier.
- ano** Reference year.
- nome\_formatado** Standardized station name.
- frist\_year** First year with available observations for the station.
- last\_year** Last year with available observations for the station.
- id\_ibge7** Seven-digit IBGE municipality code.
- id\_tse** Municipality code used by the Brazilian Electoral Court (TSE).
- latitude** Latitude in decimal degrees (WGS84).
- longitude** Longitude in decimal degrees (WGS84).

**Details**

The dataset includes station identifiers, location information, state, municipality codes, and the first and last years in which data are available for each station.

**Source**

Instituto Nacional de Meteorologia (INMET).

**Examples**

```
head(rain_stations)
unique(rain_stations$ano)
```

---

read\_inmet

---

*Read INMET meteorological observations*


---

**Description**

Reads an INMET dataset previously created with [build\_inmet\_dataset()]. The dataset is accessed through the Arrow Dataset interface, allowing efficient filtering without loading all observations into memory.

**Usage**

```
read_inmet(
  path = NULL,
  years = NULL,
  stations = NULL,
  variables = NULL,
  collect = FALSE
)
```

**Arguments**

|           |                                                                                                                                              |
|-----------|----------------------------------------------------------------------------------------------------------------------------------------------|
| path      | Character. Path to the directory containing the processed INMET dataset.                                                                     |
| years     | Integer vector of years to read. If 'NULL', all available years are returned.                                                                |
| stations  | Character vector of WMO station codes. If 'NULL', all stations are returned.                                                                 |
| variables | Character vector of variables (columns) to return. If 'NULL', all variables are returned.                                                    |
| collect   | Logical. If 'TRUE', the filtered dataset is collected into memory as a data frame. If 'FALSE' (default), an Arrow Dataset query is returned. |

**Details**

The function performs filtering directly on disk whenever possible, making it suitable for working with large datasets.

Setting 'collect = TRUE' loads the selected observations into memory. This may require a large amount of RAM when reading many years or stations simultaneously. Consider filtering by year, station, or variables before collecting the data.

**Value**

If 'collect = FALSE', returns an Arrow Dataset query. If 'collect = TRUE', returns a data frame containing the selected observations.

**See Also**

[download\_inmet()], [build\_inmet\_dataset()]

**Examples**

```
# Requires INMET data downloaded with download_inmet() and
# processed into an Arrow dataset with build_inmet_dataset().
# See the vignette "climateBR: An R package to download meteorological data from Brazil".
# for the complete workflow of this function.

## Not run:

## Read a single year without loading the data into memory
rainfall_df1 <- read_inmet(
  path = dataset_dir,
  years = 2000,
  collect = FALSE
)

## Read multiple years and collect the results into memory
rainfall_df2 <- read_inmet(
  path = dataset_dir,
  years = 2000:2005,
  collect = TRUE
)

## For large datasets, keeping collect = FALSE is generally
## recommended to avoid excessive memory usage.

## End(Not run)
```

# Index

- \* **datasets**
  - floods\_rs, [4](#)
  - inmet\_stations, [5](#)
  - mun\_stations, [8](#)
  - mun\_stations\_distance, [9](#)
  - municipality, [7](#)
  - rain\_stations, [11](#)
- build\_inmet\_dataset, [2](#)
- download\_inmet, [3](#)
- floods\_rs, [4](#)
- inmet\_stations, [5](#)
- kriging\_inmet, [6](#)
- mun\_stations, [8](#)
- mun\_stations\_distance, [9](#)
- municipality, [7](#)
- nearest\_stations, [10](#)
- rain\_stations, [11](#)
- read\_inmet, [12](#)