Package 'rcamelspe'

November 7, 2026

Title Catchment Attributes and Meteorology for Peru (CAMELS-PE)

Version 0.1.0

Description Download, load, manage, and filter the Catchment Attributes and Meteorology for Large-Sample Studies - Peru (CAMELS-PE) dataset for hydroclimatic and hydrological analysis. Offers high-performance reading and manipulation of large datasets using the 'arrow' and 'collapse' packages. Provides tools to extract catchment-scale meteorological time series, physical attributes, and spatial boundaries for 136 catchments, with 10 daily variables on the 1981-2025 calendar and 79 attributes. CAMELS-PE data are distributed separately under CC BY 4.0 <doi:10.5281/zenodo.21195425>.

License MIT + file LICENSE

Encoding UTF-8

Depends R (>= 4.1.0)

Imports arrow, cli, collapse (>= 2.0.0), dplyr, ggplot2, sf, tidyselect

Suggests testthat (>= 3.0.0), knitr, rmarkdown

VignetteBuilder knitr

Config/testthat/edition 3

URL https://github.com/PaulESantos/rcamelspe, https://paulesantos.github.io/rcamelspe/

BugReports https://github.com/PaulESantos/rcamelspe/issues

Config/roxygen2/version 8.1.1

NeedsCompilation no

Author Paul E. Santos Andrade [aut, cre, cph] (ORCID: <https://orcid.org/0000-0002-6635-0375>)

Maintainer Paul E. Santos Andrade <paulefrens@gmail.com>

Repository CRAN

Date/Publication 2026-11-07 18:10:02 UTC

Contents

download_camels_pe . . . 2
download_pe_data . . . 3
get_camels_pe_path . . . 4
load_pe_attributes . . . 4
load_pe_geospatial . . . 5
load_pe_metadata . . . 6
load_pe_timeseries . . . 7
plot_pe_attribute_map . . . 9
plot_pe_catchments . . . 10
plot_pe_timeseries . . . 11
read_attributes . . . 12
read_dictionary . . . 13
read_geospatial . . . 14
read_metadata . . . 14
read_timeseries . . . 15
set_camels_pe_path . . . 16

Index **17**

`download_camels_pe` *Download CAMELS-PE dataset from Zenodo*

Description

Compatibility alias matching `RCamelsPE::download_camels_pe()`.

Usage

```
download_camels_pe(
  path,
  version = "1.0.1",
  unzip = TRUE,
  overwrite = FALSE,
  set_path = TRUE
)
```

Arguments

`path`	Character string. Destination directory where dataset will be stored.
`version`	Character string. Dataset release: `"1.0.1"` (default) or `"1.0"`.
`unzip`	Logical. Should the downloaded file be extracted? Default is `TRUE`.
`overwrite`	Logical. If `TRUE`, existing files will be overwritten.
`set_path`	Logical. If `TRUE`, sets dataset path in session options.

Value

Character string indicating the path to the downloaded archive or directory.

Examples

```
download_camels_pe(path = tempdir(), unzip = FALSE)
```

`download_pe_data` *Download the CAMELS-PE dataset from Zenodo*

Description

Downloads the CAMELS-PE dataset zip file from the Zenodo repository (doi:10.5281/zenodo.21195425) and extracts it in the destination folder.

Usage

```
download_pe_data(
  dest_dir,
  unzip = TRUE,
  overwrite = FALSE,
  quiet = FALSE,
  version = "1.0.1",
  set_path = TRUE
)
```

Arguments

`dest_dir`	Character string. Destination directory where the dataset should be saved.
`unzip`	Logical. Should the downloaded file be unzipped? Default is `TRUE`.
`overwrite`	Logical. If destination files already exist, should they be overwritten? Default is `FALSE`.
`quiet`	Logical. Should download progress be suppressed? Default is `FALSE`.
`version`	Character string. Dataset release: `"1.0.1"` (default) or `"1.0"`.
`set_path`	Logical. Configure the session path after extraction only.

Value

Character string indicating the directory path where the dataset is located or extracted; the ZIP path when `unzip = FALSE`. Use a separate destination for each release, or `overwrite = TRUE` to replace an existing extraction.

Examples

```
# Download to a temporary folder
download_pe_data(dest_dir = tempdir())
```

get_camels_pe_path *Get the path to the CAMELS-PE dataset*

Description

Retrieves the path to the CAMELS-PE dataset directory. It looks up paths in the following order of precedence:

1. The package option `rcamelspe.path`.
2. The environment variable `CAMELS_PE_PATH`.
3. The persistent user data directory (`tools::R_user_dir("rcamelspe", "data")`).
4. Default search paths in the current working directory (`raw-data/CAMELS-PE`, `data-raw/CAMELS-PE`, etc.).
5. The bundled minimal sample dataset in `inst/extdata/sample_camels_pe`.

Usage

```
get_camels_pe_path()

get_camels_path()
```

Value

Character string containing the directory path where the dataset is located, or `NULL` if no valid path is found.

Examples

```
camels_path <- get_camels_pe_path()
is.character(camels_path)
```

load_pe_attributes *Load CAMELS-PE catchment attributes*

Description

Loads one or more catchment attribute files and merges them by `gauge_id`. High-performance implementation using 'arrow' and 'collapse'.

Usage

```
load_pe_attributes(
  attributes = "all",
  gauge_ids = NULL,
  variables = NULL,
  path = get_camels_pe_path()
)
```

Arguments

attributes	Character vector. The attribute groups to load. Can be any combination of "topographic", "climatic", "geologic", "soil", "landcover", "intervention", and "signatures", or "all" (default) to load and merge all attribute tables. Aliases "hydrological" (for "signatures") and "human_intervention" (for "intervention") are also supported.
gauge_ids	Character vector. Optional gauge identifiers to filter the returned attributes. If NULL, attributes for all catchments are returned.
variables	Character vector or NULL. Optional variable names to retain in the returned table. If NULL, all variables in the selected attribute groups are returned.
path	Character string. Path to the CAMELS-PE dataset directory. If NULL, retrieved automatically via get_camels_pe_path().

Details

CAMELS-PE v1.0.1 documents 79 attributes in seven thematic tables. Hydrological signatures are derived from simulated streamflow, not observations.

Value

A data.frame containing the merged catchment attributes with gauge_id as primary identifier.

Examples

```
# Load all attributes merged
attrs_all <- load_pe_attributes()
head(attrs_all)

# Load only topographic and climatic attributes
attrs_sub <- load_pe_attributes(c("topographic", "climatic"))
head(attrs_sub)

# Load attributes for specific stations
attrs_sel <- load_pe_attributes(gauge_ids = c("PE_110139", "PE_111151"))
attrs_sel
```

---

load_pe_geospatial *Load CAMELS-PE geospatial data*

Description

Loads geospatial catchment boundary polygons or gauging station point locations from the CAMELS-PE dataset as sf spatial objects.

Usage

```
load_pe_geospatial(
  type = c("catchments", "gauges"),
  gauge_ids = NULL,
  path = get_camels_pe_path()
)
```

Arguments

type	Character string. Either "catchments" to load catchment boundary polygons or "gauges" to load gauge point locations.
gauge_ids	Character vector. Station identifiers to filter the spatial features. If NULL (default), loads all spatial features.
path	Character string. Path to the CAMELS-PE dataset directory. If NULL, retrieved automatically via get_camels_pe_path().

Value

An sf spatial data frame object containing the requested geometries and station attributes.

Examples

```
# Load all catchment boundaries
catchments <- load_pe_geospatial(type = "catchments")
catchments

# Load specific gauging stations
gauges_sub <- load_pe_geospatial(type = "gauges", gauge_ids = "PE_110139")
gauges_sub
```

---

load_pe_metadata *Load CAMELS-PE metadata*

Description

Loads gauging station metadata or the data dictionary from the CAMELS-PE dataset.

Usage

```
load_pe_metadata(
  type = c("stations", "dictionary"),
  category = NULL,
  variable = NULL,
  file = NULL,
  path = get_camels_pe_path()
)
```

Arguments

type	Character string. Either "stations" to load gauging station metadata or "dictionary" to load the data dictionary.
category	Character vector or NULL. Optional category filter for the dictionary (e.g. "climatic").
variable	Character vector or NULL. Optional variable filter for the dictionary (e.g. "flow_obs").
file	Character vector or NULL. Optional file filter for the dictionary (e.g. "stations.csv").
path	Character string. Path to the CAMELS-PE dataset directory. If NULL, retrieved automatically via get_camels_pe_path().

Value

A data.frame containing the requested metadata:

- When type = "stations": station metadata including gauge_id, gauge_name, geographic coordinates, elevation, and hydrologic regions.
- When type = "dictionary": dictionary entries with columns folder, file, category, variable, description, unit, and source.

Examples

```
# Load stations metadata
stations <- load_pe_metadata(type = "stations")
head(stations)

# Load data dictionary
data_dict <- load_pe_metadata(type = "dictionary")
head(data_dict)

# Load data dictionary filtered by category
clim_dict <- load_pe_metadata(type = "dictionary", category = "climatic")
```

---

load_pe_timeseries *Load CAMELS-PE daily timeseries data*

Description

Efficiently loads daily hydroclimatic timeseries for Peruvian catchments. Features an optimized dual-pathway execution engine using 'arrow' and 'collapse':

- For selective requests, reads individual catchment files with column projection.
- For global or multi-catchment requests, uses arrow::open_dataset() with predicate push-down (filtering stations and date ranges directly at scan time) to minimize memory footprint and maximize throughput.

Usage

```
load_pe_timeseries(
  gauge_ids = NULL,
  variables = NULL,
  start_date = NULL,
  end_date = NULL,
  path = get_camels_pe_path(),
  parse_dates = TRUE,
  use_arrow = TRUE,
  global = FALSE
)
```

Arguments

gauge_ids	Character vector. Station/catchment identifiers (e.g. "PE_110139"). If NULL (default), loads timeseries for all available catchments.
variables	Character vector. Names of hydroclimatic variables to select. If NULL (default), all variables are loaded. Allowed variables: "prec", "prec_var", "flow_obs", "flow_sim", "pet", "tmin", "tmean", "tmax", "srad", "vprp".
start_date	Optional character string or Date object ("YYYY-MM-DD"). Start date for temporal filtering.
end_date	Optional character string or Date object ("YYYY-MM-DD"). End date for temporal filtering.
path	Character string. Path to the CAMELS-PE dataset directory. If NULL, retrieved automatically via get_camels_pe_path().
parse_dates	Logical. Should the date column be parsed into R Date objects? Default is TRUE.
use_arrow	Logical. Should the 'arrow' package be used for reading? Default is TRUE (recommended for high performance).
global	Logical. If TRUE, read the master CSV even for a small selection.

Details

The common calendar is 1981-2025; unavailable observations remain NA. Streamflow is expressed in mm/day, precipitation variance (prec_var) in mm^2/day^2, solar radiation in MJ/m^2/day, and vapor pressure in hPa. Identifier columns date and gauge_id may also be requested in variables. Automatic routing compares selected CSV bytes plus a 4 MiB per-file opening cost estimate against the master CSV size. This heuristic avoids scanning the national file for small subsets while retaining it for large requests. Missing individual files fall back to the master when it is available. global = TRUE always forces the master. No data cache is created.

Value

A data.frame containing daily hydroclimatic timeseries data with at least date and gauge_id columns, plus the requested hydrometeorological variables.

Examples

```
# Load timeseries for a specific station
ts_station <- load_pe_timeseries(gauge_ids = "PE_110139")
head(ts_station)

# Load only precipitation and observed streamflow with date range filter
ts_sub <- load_pe_timeseries(
  gauge_ids = "PE_110139",
  variables = c("prec", "flow_obs"),
  start_date = "2000-01-01",
  end_date = "2000-03-31"
)
head(ts_sub)
```

`plot_pe_attribute_map` *Plot CAMELS-PE Attribute Map*

Description

Creates a thematic map of CAMELS-PE catchments using a selected attribute. The function joins attribute data with catchment geometries by `gauge_id` using `collapse::join` and returns a `ggplot` object.

Usage

```
plot_pe_attribute_map(
  catchments,
  attributes,
  variable,
  gauges = NULL,
  na_color = "grey80",
  ...
)

plot_attribute_map(
  catchments,
  attributes,
  variable,
  gauges = NULL,
  na_color = "grey80",
  ...
)
```

Arguments

`catchments`	An `sf` object with catchment polygons.
`attributes`	A data frame containing CAMELS-PE attributes.

variable	Character string. Name of the attribute to visualize.
gauges	Optional sf object with gauge point locations.
na_color	Character string. Color for missing values. Default is "grey80".
...	Additional arguments passed to ggplot2::geom_sf() for catchments.

Value

A ggplot object visualizing the attribute map.

Examples

```
catchments <- load_pe_geospatial(type = "catchments")
attrs <- load_pe_attributes(attributes = "topographic")
plot_pe_attribute_map(catchments, attrs, variable = "area")
```

---

plot_pe_catchments *Plot CAMELS-PE Catchments*

Description

Creates a map of CAMELS-PE catchments and optionally overlays gauge locations. If gauge_id is provided, only one catchment and its outlet are plotted, with the gauge ID shown as a panel header.

Usage

```
plot_pe_catchments(
  catchments,
  gauges = NULL,
  gauge_id = NULL,
  fill = NULL,
  ...
)

plot_catchments(catchments, gauges = NULL, gauge_id = NULL, fill = NULL, ...)
```

Arguments

catchments	An sf object with catchment polygons.
gauges	Optional sf object with gauge point locations.
gauge_id	Optional character string. Gauge ID used to filter one catchment and its gauge before plotting. If NULL, all catchments are plotted.
fill	Optional character string. Name of a catchment column used to fill polygons. If NULL, a constant fill is used.
...	Additional arguments passed to ggplot2::geom_sf().

Value

A ggplot object with the spatial catchment geometries.

Examples

```
catchments <- load_pe_geospatial(type = "catchments")
gauges <- load_pe_geospatial(type = "gauges")
plot_pe_catchments(catchments, gauges)
```

---

plot_pe_timeseries *Plot CAMELS-PE Time Series*

Description

Creates a time series plot for one CAMELS-PE variable, such as precipitation, observed streamflow, simulated streamflow, or temperature. The function returns a ggplot object, styled with the project color palette.

Usage

```
plot_pe_timeseries(
  data,
  variable = "flow_obs",
  gauge_id = NULL,
  date_col = "date",
  facet = TRUE,
  scales = "free_y",
  line_color = "#2F4156",
  ...
)

plot_timeseries(
  data,
  variable = "flow_obs",
  gauge_id = NULL,
  date_col = "date",
  facet = TRUE,
  scales = "free_y",
  line_color = "#2F4156",
  ...
)
```

Arguments

data	A data frame containing CAMELS-PE time series.
variable	Character string. Name of the variable to plot. Default is "flow_obs".

gauge_id	Optional character vector. Gauge IDs used to filter the data before plotting. If NULL, all available gauges in data are plotted.
date_col	Character string. Name of the date column. Default is "date".
facet	Logical value. If TRUE (default), creates one panel per gauge ID.
scales	Character string. Scales passed to ggplot2::facet_wrap(). One of "fixed", "free", "free_x", or "free_y" (default).
line_color	Character string. Line color for the time series. Default is "#2F4156" (Navy).
...	Additional arguments passed to ggplot2::geom_line().

Value

A ggplot visualization object representing the time series.

Examples

```
ts <- load_pe_timeseries(
  gauge_ids = "PE_110139",
  variables = c("prec", "flow_sim")
)
plot_pe_timeseries(ts, variable = "prec")
```

---

read_attributes *Read CAMELS-PE catchment attributes*

Description

Compatibility alias matching the RCamelsPE::read_attributes() interface, powered by the high-performance 'arrow' and 'collapse' backend of rcamelspe.

Usage

```
read_attributes(
  type = "all",
  gauge_id = NULL,
  vars = NULL,
  path = get_camels_pe_path()
)
```

Arguments

type	Character string. Attribute group to read. One of "topographic", "climatic", "hydrological", "landcover", "geologic", "soil", "human_intervention", or "all". Default is "all".
gauge_id	Character vector or NULL. Optional gauge identifiers.
vars	Character vector or NULL. Optional variable names to retain.
path	Character string. Optional path to the CAMELS-PE root directory.

Value

A `data.frame` with CAMELS-PE catchment attributes.

Examples

```
# Read all attributes using RCamelsPE compatible syntax
attrs <- read_attributes(type = "topographic")
head(attrs)
```

`read_dictionary` *Read CAMELS-PE data dictionary*

Description

Reads the CAMELS-PE data dictionary with optional filters by category, variable, or source file. Compatible with the `RCamelsPE` interface.

Usage

```
read_dictionary(
  category = NULL,
  variable = NULL,
  file = NULL,
  path = get_camels_pe_path()
)
```

Arguments

`category`	Character vector or `NULL`. Optional category filter.
`variable`	Character vector or `NULL`. Optional variable filter.
`file`	Character vector or `NULL`. Optional file filter.
`path`	Character string. Optional path to the CAMELS-PE root directory. If not provided, retrieved automatically via `get_camels_pe_path()`.

Value

A `data.frame` containing the CAMELS-PE data dictionary entries.

Examples

```
# Read full dictionary
dict <- read_dictionary()
head(dict)

# Read dictionary for topographic category
topo_dict <- read_dictionary(category = "topographic")
head(topo_dict)
```

read_geospatial *Read CAMELS-PE geospatial data*

Description

Compatibility alias matching `RCamelsPE::read_geospatial()`.

Usage

```
read_geospatial(type = c("gauges", "catchments"), path = get_camels_pe_path())
```

Arguments

type	Character string. Either `"gauges"` or `"catchments"`.
path	Character string. Optional path to the CAMELS-PE root directory.

Value

An `sf` spatial object.

Examples

```
catchments <- read_geospatial(type = "catchments")
head(catchments)
```

read_metadata *Read CAMELS-PE station metadata*

Description

Convenience alias for `load_pe_metadata(type = "stations")`, compatible with the `RCamelsPE` interface.

Usage

```
read_metadata(path = get_camels_pe_path())
```

Arguments

path	Character string. Optional path to the CAMELS-PE root directory. If not provided, retrieved automatically via `get_camels_pe_path()`.

Value

A `data.frame` containing gauging station metadata.

Examples

```
stations <- read_metadata()
head(stations)
```

`read_timeseries` *Read CAMELS-PE time series*

Description

Compatibility alias matching the `RCamelsPE::read_timeseries()` interface, powered by the high-performance 'arrow' dataset and 'collapse' query engine.

Usage

```
read_timeseries(
  gauge_id = NULL,
  global = FALSE,
  vars = NULL,
  start_date = NULL,
  end_date = NULL,
  path = get_camels_pe_path()
)
```

Arguments

`gauge_id`	Character vector or `NULL`. Gauge identifiers to read.
`global`	Logical value. If `TRUE`, forces reading from the global file. If `FALSE`, reads individual catchment files when available.
`vars`	Character vector or `NULL`. Optional variable names to retain.
`start_date`	Optional character string or Date (`"YYYY-MM-DD"`).
`end_date`	Optional character string or Date (`"YYYY-MM-DD"`).
`path`	Character string. Optional path to the CAMELS-PE root directory.

Value

A `data.frame` containing CAMELS-PE time series data.

Examples

```
# Read timeseries using RCamelsPE compatible syntax
ts <- read_timeseries(gauge_id = "PE_110139", vars = c("date", "prec", "flow_obs"))
head(ts)
```

`set_camels_pe_path` *Set the path to the CAMELS-PE dataset*

Description

Sets the path to the directory containing the CAMELS-PE dataset. The path is stored in the package session options.

Usage

```
set_camels_pe_path(path)

set_camels_path(path)
```

Arguments

`path`	Character string. Path to the CAMELS-PE dataset directory.

Value

Invisible `NULL`. Called for its side effect of setting the dataset path.

Examples

```
old_path <- getOption("rcamelspe.path")
sample_path <- system.file("extdata", "sample_camels_pe", package = "rcamelspe")
set_camels_pe_path(sample_path)
# Reset options
options(rcamelspe.path = old_path)
```

Index

download_camels_pe, 2
download_pe_data, 3

get_camels_path *(*get_camels_pe_path*)*, 4
get_camels_pe_path, 4
get_camels_pe_path(), *5–8*, *13, 14*

load_pe_attributes, 4
load_pe_geospatial, 5
load_pe_metadata, 6
load_pe_timeseries, 7

plot_attribute_map
        *(*plot_pe_attribute_map*)*, 9
plot_catchments *(*plot_pe_catchments*)*, 10
plot_pe_attribute_map, 9
plot_pe_catchments, 10
plot_pe_timeseries, 11
plot_timeseries *(*plot_pe_timeseries*)*, 11

read_attributes, 12
read_dictionary, 13
read_geospatial, 14
read_metadata, 14
read_timeseries, 15

set_camels_path *(*set_camels_pe_path*)*, 16
set_camels_pe_path, 16