Package 'blackmarbler'

May 7, 2026

Title Black Marble Data and Statistics

Version 0.2.5

Description Geographically referenced data and statistics of nighttime lights from NASA Black Marble <https://blackmarble.gsfc.nasa.gov/>.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.3.1

URL https://worldbank.github.io/blackmarbler/

BugReports https://github.com/worldbank/blackmarbler/issues

Imports readr, dplyr, purrr, lubridate, tidyr, terra, sf, exactextractr, stringr, httr2

Suggests geodata, ggplot2, knitr, tidyterra, testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Author Robert Marty [aut, cre] (ORCID: <https://orcid.org/0000-0002-3164-3813>), Gabriel Stefanini Vicente [aut] (ORCID: <https://orcid.org/0000-0001-6530-3780>)

Maintainer Robert Marty <rmarty@worldbank.org>

Repository CRAN

Date/Publication 2025-07-29 21:00:03 UTC

Contents

bm_extract *Extract and Aggregate Black Marble Data*

Description

Extract and aggregate nighttime lights data from NASA Black Marble data

Usage

```
bm_extract(
  roi_sf,
  product_id,
  date,
  bearer,
  aggregation_fun = c("mean"),
  add_n_pixels = TRUE,
  variable = NULL,
  quality_flag_rm = NULL,
  check_all_tiles_exist = TRUE,
  interpol_na = FALSE,
  output_location_type = "memory",
  file_dir = NULL,
  file_prefix = NULL,
  file_skip_if_exists = TRUE,
  file_return_null = FALSE,
  h5_dir = NULL,
  download_method = "httr",
  quiet = FALSE,
  ...
)
```

Arguments

`roi_sf`	Region of interest; sf polygon. Must be in the WGS 84 (epsg:4326) coordinate reference system.
`product_id`	One of the following: • `"VNP46A1"`: Daily (raw) • `"VNP46A2"`: Daily (corrected) • `"VNP46A3"`: Monthly • `"VNP46A4"`: Annual
`date`	Date of raster data. Entering one date will produce a `SpatRaster` object. Entering multiple dates will produce a `SpatRaster` object with multiple bands; one band per date. • For `product_ids` `"VNP46A1"` and `"VNP46A2"`, a date (eg, `"2021-10-03"`).

- For product_id "VNP46A3", a date or year-month (e.g., "2021-10-01", where the day will be ignored, or "2021-10").
- For product_id "VNP46A4", year or date (e.g., "2021-10-01", where the month and day will be ignored, or 2021).

bearer NASA bearer token. For instructions on how to create a token, see here.

aggregation_fun
Function used to aggregate nighttime lights data to polygons; this values is passed to the fun argument in exactextractr::exact_extract (Default: mean).

add_n_pixels Whether to add a variable indicating the number of nighttime light pixels used to compute nighttime lights statistics (eg, number of pixels used to compute average of nighttime lights). When TRUE, it adds three values: n_non_na_pixels (the number of non-NA pixels used for computing nighttime light statistics); n_pixels (the total number of pixels); and prop_non_na_pixels the proportion of the two. (Default: TRUE).

variable Variable to used to create raster (default: NULL). If NULL, uses the following default variables:

- For product_id :VNP46A1", uses DNB_At_Sensor_Radiance_500m.
- For product_id "VNP46A2", uses Gap_Filled_DNB_BRDF-Corrected_NTL.
- For product_ids "VNP46A3" and "VNP46A4", uses NearNadir_Composite_Snow_Free. To see all variable choices, set variable = "" (this will create an error message that lists all valid variables). For additional information on variable choices, see here; for VNP46A1, see Table 3; for VNP46A2 see Table 6; for VNP46A3 and VNP46A4, see Table 9.

quality_flag_rm
Quality flag values to use to set values to NA. Each pixel has a quality flag value, where low quality values can be removed. Values are set to NA for each value in the quality_flag_rm vector. Note that quality_flag_rm does not apply for VNP46A1. (Default: NULL).

For VNP46A2 (daily data):

- 0: High-quality, Persistent nighttime lights
- 1: High-quality, Ephemeral nighttime Lights
- 2: Poor-quality, Outlier, potential cloud contamination, or other issues

For VNP46A3 and VNP46A4 (monthly and annual data):

- 0: Good-quality, The number of observations used for the composite is larger than 3
- 1: Poor-quality, The number of observations used for the composite is less than or equal to 3
- 2: Gap filled NTL based on historical data

check_all_tiles_exist
Check whether all Black Marble nighttime light tiles exist for the region of interest. Sometimes not all tiles are available, so the full region of interest may not be covered. If TRUE, skips cases where not all tiles are available. (Default: TRUE).

`interpol_na` When data for more than one date is downloaded, whether to interpolate `NA` values in rasters using the `terra::approximate` function. Additional arguments for the `terra::approximate` function can also be passed into `bm_extract` (eg, `method`, `rule`, `f`, `ties`, `z`, `NA_rule`). (Default: `FALSE`).

`output_location_type` Where to produce output; either `memory` or `file`. If `memory`, functions returns a dataframe in R. If `file`, function exports a `.csv` file and returns `NULL`.

`file_dir` (If `output_location_type = file`). The directory where data should be exported (default: `NULL`, so the working directory will be used)

`file_prefix` (If `output_location_type = file`). Prefix to add to the file to be saved. The file will be saved as the following: `[file_prefix][product_id]_t[date].csv`

`file_skip_if_exists` (If `output_location_type = file`). Whether the function should first check wither the file already exists, and to skip downloading or extracting data if the data for that date if the file already exists (default: `TRUE`).

`file_return_null` Whether to return `NULL` instead of a `dataframe`. When `output_location_type = 'file'`, the function will export data to the `file_dir` directory. When `file_return_null = FALSE`, the function will also return a `dataframe` of the queried data—so the data is available in R memory. Setting `file_return_null = TRUE`, data will be saved to `file_dir` but no data will be returned by the function to R memory (default: `FALSE`).

`h5_dir` Black Marble data are originally downloaded as `h5` files. If `h5_dir = NULL`, the function downloads to a temporary directory then deletes the directory. If `h5_dir` is set to a path, `h5` files are saved to that directory and not deleted. The function will then check if the needed `h5` file already exists in the directory; if it exists, the function will not re-download the `h5` file.

`download_method` Method to download data from NASA LAADS Archive: `"httr"` or `"wget"`. If `httr`, uses the `httr2` R package to download data. If `wget`, uses the `wget` command line tool. `httr` is fully integrated in R, while `wget` requires the `wget` system command. `wget` can be more efficient and can help avoid network issues. (Default: `"httr"`).

`quiet` Suppress output that show downloading progress and other messages. (Default: `FALSE`).

`...` Additional arguments for `terra::approximate`, if `interpol_na = TRUE`

Value

Raster

Author(s)

Robert Marty rmarty@worldbank.org

Examples

```
## Not run:
# Define bearer token
bearer <- "BEARER-TOKEN-HERE"

# sf polygon of Ghana
library(geodata)
roi_sf <- gadm(country = "GHA", level=1, path = tempdir()) %>% st_as_sf()

# Daily data: raster for October 3, 2021
ken_20210205_r <- bm_extract(roi_sf = roi_sf,
                             product_id = "VNP46A2",
                             date = "2021-10-03",
                             bearer = bearer)

# Monthly data: raster for March 2021
ken_202103_r <- bm_extract(roi_sf = roi_sf,
                           product_id = "VNP46A3",
                           date = "2021-03-01",
                           bearer = bearer)

# Annual data: raster for 2021
ken_2021_r <- bm_extract(roi_sf = roi_sf,
                         product_id = "VNP46A4",
                         date = 2021,
                         bearer = bearer)

## End(Not run)
```

---

bm_raster *Make Black Marble Raster*

Description

Make a raster of nighttime lights from NASA Black Marble data

Usage

```
bm_raster(
  roi_sf,
  product_id,
  date,
  bearer,
  variable = NULL,
  quality_flag_rm = NULL,
  check_all_tiles_exist = TRUE,
  interpol_na = FALSE,
  output_location_type = "memory",
```

```
  file_dir = NULL,
  file_prefix = NULL,
  file_skip_if_exists = TRUE,
  file_return_null = FALSE,
  h5_dir = NULL,
  download_method = "httr",
  quiet = FALSE,
  ...
)
```

Arguments

`roi_sf`	Region of interest; sf polygon. Must be in the WGS 84 (epsg:4326) coordinate reference system.
`product_id`	One of the following: • `"VNP46A1"`: Daily (raw) • `"VNP46A2"`: Daily (corrected) • `"VNP46A3"`: Monthly • `"VNP46A4"`: Annual
`date`	Date of raster data. Entering one date will produce a `SpatRaster` object. Entering multiple dates will produce a `SpatRaster` object with multiple bands; one band per date. • For `product_ids` `"VNP46A1"` and `"VNP46A2"`, a date (eg, `"2021-10-03"`). • For `product_id` `"VNP46A3"`, a date or year-month (e.g., `"2021-10-01"`, where the day will be ignored, or `"2021-10"`). • For `product_id` `"VNP46A4"`, year or date (e.g., `"2021-10-01"`, where the month and day will be ignored, or `2021`).
`bearer`	NASA bearer token. For instructions on how to create a token, see here.
`variable`	Variable to used to create raster (default: `NULL`). If `NULL`, uses the following default variables: • For `product_id` `:VNP46A1"`, uses `DNB_At_Sensor_Radiance_500m`. • For `product_id` `"VNP46A2"`, uses `Gap_Filled_DNB_BRDF-Corrected_NTL`. • For `product_ids` `"VNP46A3"` and `"VNP46A4"`, uses `NearNadir_Composite_Snow_Free`. To see all variable choices, set `variable = ""` (this will create an error message that lists all valid variables). For additional information on variable choices, see here; for `VNP46A1`, see Table 3; for `VNP46A2` see Table 6; for `VNP46A3` and `VNP46A4`, see Table 9.
`quality_flag_rm`	Quality flag values to use to set values to `NA`. Each pixel has a quality flag value, where low quality values can be removed. Values are set to `NA` for each value in the `quality_flag_rm` vector. Note that `quality_flag_rm` does not apply for `VNP46A1`. (Default: `NULL`). For `VNP46A2` (daily data): • `0`: High-quality

- 1: Poor-quality - Main Algorithm (Outlier, Potential cloud contamination or other issues)
- 2: Poor-quality - Main Algorithm (high solar zenith angle 102-108 degrees)
- 3: Poor-quality - Main Algorithm (Lunar eclipse)
- 4: Poor-quality - Main Algorithm (Aurora)
- 5: Poor-quality - Main Algorithm (Glint)

For VNP46A3 and VNP46A4 (monthly and annual data):

- 0: Good-quality, The number of observations used for the composite is larger than 3
- 1: Poor-quality, The number of observations used for the composite is less than or equal to 3
- 2: Gap filled NTL based on historical data

check_all_tiles_exist
: Check whether all Black Marble nighttime light tiles exist for the region of interest. Sometimes not all tiles are available, so the full region of interest may not be covered. If TRUE, skips cases where not all tiles are available. (Default: TRUE).

interpol_na
: When data for more than one date is downloaded, whether to interpolate NA values using the terra::approximate function. Additional arguments for the terra::approximate function can also be passed into bm_raster (eg, method, rule, f, ties, z, NA_rule). (Default: FALSE).

output_location_type
: Where to produce output; either memory or file. If memory, functions returns a raster in R. If file, function exports a .tif file and returns NULL. For output_location_type = file:

file_dir
: The directory where data should be exported (default: NULL, so the working directory will be used)

file_prefix
: Prefix to add to the file to be saved. The file will be saved as the following: [file_prefix][product_id]_t[date].tif

file_skip_if_exists
: Whether the function should first check wither the file already exists, and to skip downloading or extracting data if the data for that date if the file already exists (default: TRUE).

file_return_null
: Whether to return NULL instead of a SpatRaster. When output_location_type = 'file', the function will export data to the file_dir directory. When file_return_null = FALSE, the function will also return a SpatRaster of the queried data—so the data is available in R memory. Setting file_return_null = TRUE, data will be saved to file_dir but no data will be returned by the function to R memory (default: FALSE).

h5_dir
: Black Marble data are originally downloaded as h5 files. If h5_dir = NULL, the function downloads to a temporary directory then deletes the directory. If h5_dir is set to a path, h5 files are saved to that directory and not deleted. The function will then check if the needed h5 file already exists in the directory; if it exists, the function will not re-download the h5 file.

download_method
: Method to download data from NASA LAADS Archive: "httr" or "wget". If httr, uses the httr2 R package to download data. If wget, uses the wget command line tool. httr is fully integrated in R, while wget requires the wget system command. wget can be more efficient and can help avoid network issues. (Default: "httr").

quiet
: Suppress output that show downloading progress and other messages. (Default: FALSE).

...
: Additional arguments for terra::approximate, if interpol_na = TRUE

Value

Raster

Author(s)

Robert Marty rmarty@worldbank.org

Examples

```
## Not run:
# Define bearer token
bearer <- "BEARER-TOKEN-HERE"

# sf polygon of Ghana
library(geodata)
roi_sf <- gadm(country = "GHA", level=0, path = tempdir()) %>% st_as_sf()

# Daily data: raster for October 3, 2021
ken_20210205_r <- bm_raster(roi_sf = roi_sf,
                            product_id = "VNP46A2",
                            date = "2021-10-03",
                            bearer = bearer)

# Monthly data: raster for March 2021
ken_202103_r <- bm_raster(roi_sf = roi_sf,
                          product_id = "VNP46A3",
                          date = "2021-03-01",
                          bearer = bearer)

# Annual data: raster for 2021
ken_2021_r <- bm_raster(roi_sf = roi_sf,
                        product_id = "VNP46A4",
                        date = 2021,
                        bearer = bearer)

## End(Not run)
```

get_nasa_token *Get a NASA Earthdata bearer token*

Description

Fetch a NASA Earthdata bearer token from using the Earthdata API. If none exist, this will create one, or if one already exists it will fetch that one instead.

Usage

```
get_nasa_token(username, password)
```

Arguments

username	character. NASA Earthdata username
password	character. NASA Earthdata password

Value

character

Author(s)

Simon E. H. Smart simon.smart@cantab.net

wget_h5_files *Download h5 files using wget*

Description

Download h5 files from from NASA Black Marble data using `wget`. The wget_h5_files() function requires the wget command line tool to be installed on your system. If you do not have wget installed, please install it from https://www.gnu.org/software/wget/.

Usage

```
wget_h5_files(roi_sf = NULL, product_id, date, h5_dir, bearer)
```

Arguments

roi_sf	Region of interest; sf polygon. Must be in the WGS 84 (epsg:4326) coordinate reference system. If NULL, all h5 files for the inputted date(s) are downloaded.
product_id	One of the following: • "VNP46A1": Daily (raw) • "VNP46A2": Daily (corrected) • "VNP46A3": Monthly • "VNP46A4": Annual
date	Date(s) to download h5 files. • For product_ids "VNP46A1" and "VNP46A2", a date (eg, "2021-10-03"). • For product_id "VNP46A3", a date or year-month (e.g., "2021-10-01", where the day will be ignored, or "2021-10"). • For product_id "VNP46A4", year or date (e.g., "2021-10-01", where the month and day will be ignored, or 2021).
h5_dir	Path to download h5 files to.
bearer	NASA bearer token. For instructions on how to create a token, see here.

Value

NULL

Author(s)

Robert Marty rmarty@worldbank.org

Examples

```
## Not run:
# Define bearer token
bearer <- "BEARER-TOKEN-HERE"

# sf polygon of Ghana
library(geodata)
roi_sf <- gadm(country = "GHA", level=0, path = tempdir()) %>% st_as_sf()

# h5 files for Ghana for October 3, 2021
download_h5_files(roi_sf = roi_sf,
                  product_id = "VNP46A2",
                  date = "2021-10-03",
                  h5_dir = getwd(),
                  bearer = bearer)

# Make raster using h5_files
ken_202103_r <- bm_raster(roi_sf = roi_sf,
                          product_id = "VNP46A3",
                          date = "2021-03-01",
                          bearer = bearer,
                          h5_dir = getwd())
```

```
## End(Not run)
```

Index