Package 'amerifluxr'

October 12, 2022

Type Package

Title Interface to 'AmeriFlux' Data Services

Version 1.0.0

Description Programmatic interface to the 'AmeriFlux' database (<https://ameriflux.lbl.gov/>). Provide query, download, and data summary tools.

License BSD_3_clause + file LICENSE

Encoding UTF-8

LazyData true

URL https://github.com/chuhousen/amerifluxr

BugReports https://github.com/chuhousen/amerifluxr/issues

Depends R (>= 3.6)

Imports heatmaply, httr, jsonlite, memoise, RCurl, readxl, tools

RoxygenNote 7.1.2

Suggests covr, data.table, knitr, pander, qpdf, remotes, rmarkdown, testthat, mockr

VignetteBuilder knitr

NeedsCompilation no

Author Housen Chu [aut, cre] (<https://orcid.org/0000-0002-8131-4938>), Koen Hufkens [ctb] (<https://orcid.org/0000-0002-5070-8109>)

Maintainer Housen Chu <hchu@lbl.gov>

Repository CRAN

Date/Publication 2022-02-08 16:30:08 UTC

R topics documented:

`amf_base` *BASE data example*

Description

Continuous flux/met data (i.e., AmeriFlux BASE data product) for the US-CRT site, as an example for demonstration. Also see AmeriFlux webpage https://ameriflux.lbl.gov/data/aboutdata/data-variables/#base for variable definitions and details.

Usage

```
amf_base
```

Format

A data frame with 336 rows and 36 variables

Details

- TIMESTAMP_START - ISO timestamp start of averaging period (YYYYMMDDHHMM)
- TIMESTAMP_END - ISO timestamp end of averaging period (YYYYMMDDHHMM)
- CO2 - Carbon Dioxide (CO2) mole fraction in wet air (µmolCO2 mol-1)
- H2O - Water (H2O) vapor in mole fraction of wet air (mmolH2O mol-1)
- FC - Carbon Dioxide (CO2) turbulent flux (µmolCO2 m-2 s-1)
- NEE_PI - Net Ecosystem Exchange (µmolCO2 m-2 s-1)
- CH4 - Methane (CH4) mole fraction in wet air (nmolCH4 mol-1)

- FCH4 - Methane (CH4) turbulent flux (nmolCH4 m-2 s-1)
- H - Sensible heat turbulent flux (W m-2)
- LE - Latent heat turbulent flux (W m-2)
- G_1_1_1 - Soil heat flux at horizontal location #1 (W m-2)
- G_2_1_1 - Soil heat flux at horizontal location #2 (W m-2)
- WD - Wind direction (Decimal degrees)
- WS - Wind speed (m s-1)
- USTAR - Friction velocity (m s-1)
- ZL - data value
- MO_LENGTH - Monin-Obukhov Stability parameter (nondimensional)
- W_SIGMA - Standard deviation of vertical velocity fluctuations (m s-1)
- V_SIGMA - Standard deviation of lateral velocity fluctuations (m s-1)
- U_SIGMA - Standard deviation of along-wind velocity fluctuations (m s-1)
- T_SONIC - Sonic temperature (deg C)
- T_SONIC_SIGMA - Standard deviation of sonic temperature (deg C)
- PA - Atmospheric pressure (kPa)
- RH - Relative humidity (%)
- TA - Air temperature (deg C)
- TS_1_1_1 - Soil temperature at horizontal location #1 (deg C)
- TS_2_1_1 - Soil temperature at horizontal location #2 (deg C)
- WTD - Water table depth (m)
- SWC - Soil water content (%)
- NETRAD - Net radiation (W m-2)
- PPFD_IN - Photosynthetic photon flux density, incoming (µmolPhoton m-2 s-1)
- SW_IN - Shortwave radiation, incoming (W m-2)
- SW_OUT - Shortwave radiation, outgoing (W m-2)
- LW_IN - Longwave radiation, incoming (W m-2)
- LW_OUT - Longwave radiation, outgoing (W m-2)
- P - Rainfall (mm)

Source

https://ameriflux.lbl.gov/

amf_bif *BADM data example*

Description

The BADM file provides a description, general background, geo-location, relevant publications and references for the site. The BADM files also list what is measured at the site (this list may include chamber, gradient or other eddy covariance measurements that are a super-set of the data available in AmeriFlux). Also see AmeriFlux web page https://ameriflux.lbl.gov/data/aboutdata/badm-data-product/ for details.

Usage

```
amf_bif
```

Format

A data frame with 443 rows and 5 variables

Details

- SITE_ID - 6 digit AmeriFlux site ID (CC-Sss)
- GROUP_ID - A unique identifier for data entries belonging to the same instance of a reported variable group
- VARIABLE_GROUP - Define a set of variables that are reported together
- VARIABLE - Variable names
- DATAVALUE - Data value

Source

https://ameriflux.lbl.gov/

amf_check_site_id *Check valid AmeriFlux site ID*

Description

Check if the character is a valid AmeriFlux site ID (CC-Sss)

Usage

```
amf_check_site_id(x)
```

Arguments

x	A vector or scalar of characters

Value

logical vector or scalar

Examples

```
## Not run:
# Check if valid site ID
check_id <- amf_check_site_id(c("US-CRT", "US-crt", "USCRT"))


## End(Not run)
```

---

amf_data_coverage *Returns a list of data coverage*

Description

AmeriFlux data coverage statistics

Usage

```
amf_data_coverage(data_product = "BASE-BADM", data_policy = "CCBY4.0")
```

Arguments

data_product	A scalar of character specifying the data product Currently, only "BASE-BADM" is supported.
data_policy	A scalar of character specifying the data policy Currently, "CCBY4.0" and "LEGACY" are supported. The default is "CCBY4.0".

Value

AmeriFlux data coverage

- SITE_ID - Six character site identifier (CC-Sss)
- URL - Site page link (url)
- publish_years - List of data available years (YYYY)

Examples

```
## Not run:
# download the variable availability
data_year <- amf_data_coverage()

# download variable availability for LEGACY policy
data_year <- amf_data_coverage(data_policy = "LEGACY")

## End(Not run)
```

amf_download_base *Download AmeriFlux BASE data product*

Description

This function downloads AmeriFlux BASE and BADM data files. Note: Access to AmeriFlux data requires creating an AmeriFlux account first. Register an account through the link https://ameriflux-data.lbl.gov/Pages/RequestAccount.aspx.

For details about BASE and BADM data files, see AmeriFlux web pages https://ameriflux.lbl.gov/data/data-processing-pipelines/base-publish/ and https://ameriflux.lbl.gov/data/aboutdata/badm-data-product/.

Usage

```
amf_download_base(
  user_id,
  user_email,
  site_id,
  data_product = "BASE-BADM",
  data_policy,
  agree_policy,
  intended_use,
  intended_use_text,
  out_dir = tempdir(),
  verbose = TRUE
)
```

Arguments

user_id	AmeriFlux account username (character)
user_email	AmeriFlux account user email (character)
site_id	A scalar or vector of character specifying the AmeriFlux Site ID (CC-Sss)
data_product	AmeriFlux data product. Currently, only "BASE-BADM" is supported and used as default. (character)
data_policy	"CCBY4.0" or "LEGACY" (character). AmeriFlux data are shared under two tiers of licenses as chosen by site's PI. See https://ameriflux.lbl.gov/data/data-policy/#data-use for data use guidelines under each license. Note: Data use policy selected affects which sites' data are available for download.
agree_policy	Acknowledge you read and agree to the AmeriFlux Data use policy (TRUE/FALSE)
intended_use	The intended use category. Currently, it needs to be one of the followings: • "synthesis" (i.e., Multi-site synthesis) • "model" (i.e., Land model/Earth system model) • "remote_sensing" (i.e., Remote sensing research) • "other_research" (i.e., Other research) "

- "education" (i.e., Education (Teacher or Student))
- "other"

`intended_use_text`
: Enter a brief description of intended use. This will be recorded in the download log and emailed to site's PI (character).

`out_dir`
: Output directory for downloaded data, default tempdir()

`verbose`
: Show feedback on download progress (TRUE/FALSE)

Value

A vector of download file names on the local drive

Examples

```
## Not run:
## Download a single site, under CCBY4.0 policy
amf_download_base(user_id = "test",
 user_email = "test _at_ mail.com",
 site_id = "US-CRT",
 data_product = "BASE-BADM",
 data_policy = "CCBY4.0",
 agree_policy = TRUE,
 intended_use = "other",
 intended_use_text = "testing download",
 out_dir = tempdir())

## Download several sites, under LEGACY data policy
#  When finished, return a list of downloaded files
#  in your local drive.
file.ls <- amf_download_base(user_id = "test",
 user_email = "test _at_ mail.com",
 site_id = c("US-CRT", "US-WPT", "US-Oho"),
 data_product = "BASE-BADM",
 data_policy = "LEGACY",
 agree_policy = TRUE,
 intended_use = "other",
 intended_use_text = "testing download",
 out_dir = tempdir())

## End(Not run)
```

`amf_download_bif` *Download AmeriFlux BADM data product*

Description

This function downloads all AmeriFlux sites' BADM data files in a single file. Note: Access to AmeriFlux data requires creating an AmeriFlux account first. Register an account through the link https://ameriflux-data.lbl.gov/Pages/RequestAccount.aspx.

For details about BADM data files, see AmeriFlux web page https://ameriflux.lbl.gov/data/aboutdata/badm-data-product/.

Usage

```
amf_download_bif(
  user_id,
  user_email,
  data_policy,
  agree_policy,
  intended_use,
  intended_use_text,
  out_dir = tempdir(),
  verbose = TRUE,
  site_w_data = FALSE
)
```

Arguments

user_id	AmeriFlux account username (character)
user_email	AmeriFlux account user email (character)
data_policy	"CCBY4.0" or "LEGACY". AmeriFlux data are shared under two tiers of licenses as chosen by site's PI. See https://ameriflux.lbl.gov/data/data-policy/#data-use for data use guidelines under each license. Note: Data use policy selected affects which sites' data are available for download.
agree_policy	Acknowledge you read and agree to the AmeriFlux Data use policy (TRUE/FALSE)
intended_use	The intended use category. Currently, it needs to be one of the followings: • "synthesis" (i.e., Multi-site synthesis) • "model" (i.e., Land model/Earth system model) • "remote_sensing" (i.e., Remote sensing research) • "other_research" (i.e., Other research) " • "education" (i.e., Education (Teacher or Student)) • "other"
intended_use_text	Enter a brief description of intended use. This will be recorded in the data download log and emailed to site's PI (character).
out_dir	Output directory for downloaded data, default tempdir()
verbose	Show feedback on download progress (TRUE/FALSE)
site_w_data	Logical, download all registered sites (FALSE) or only sites with available BASE data (TRUE)

Value

A vector of download file names on the local drive

See Also

amf_download_base

Examples

```
## Not run:
## Download all sites with BASE data, under CCBY4.0 policy
amf_download_bif(user_id = "test",
  user_email = "test@mail.com",
  data_policy = "CCBY4.0",
  agree_policy = TRUE,
  intended_use = "other",
  intended_use_text = "testing download",
  out_dir = tempdir(),
  site_w_data = TRUE)

## Download all registered sites, under LEGACY policy
amf_download_bif(user_id = "test",
  user_email = "test@mail.com",
  data_policy = "LEGACY",
  agree_policy = TRUE,
  intended_use = "other",
  intended_use_text = "testing download",
  out_dir = tempdir(),
  site_w_data = FALSE)


## End(Not run)
```

---

amf_extract_badm *Extract BADM data of a specific BADM group*

Description

This function extracts BADM data of a specific BADM group from the imported BADM (BIF) file. Use function amf_read_bif to import BADM (BIF) file.

Usage

```
amf_extract_badm(bif_data, select_group)
```

Arguments

`bif_data`	A data frame consists of 5 columns: SITE_ID, GROUP_ID, VARIABLE_GROUP, VARIABLE, DATAVALUE, imported from function `amf_read_bif`.
`select_group`	A string (character), selected from VARIABLE_GROUP in the `bif_data`

Value

A data frame of re-structured BADM data with the following columns:

- GROUP_ID - A unique identifier for data belonging to the same instance of a reported variable group
- SITE_ID - Six character site identifier (CC-Sss)
- VALUE - Values for all available VARIABLES in the selected group
- ...

See Also

`amf_read_bif`

Examples

```
# read the BADM BIF file, using an example data file
bif <- amf_read_bif(file = system.file("extdata",
                                       "AMF_AA-Flx_BIF_CCBY4_20201218.xlsx",
                                        package = "amerifluxr"))

# get a list of valid VARIALBE_GROUP
unique(bif$VARIABLE_GROUP)

# extract the selected VARIALBE_GROUP
amf_extract_badm(bif_data = bif, select_group = "GRP_FLUX_MEASUREMENTS")
amf_extract_badm(bif_data = bif, select_group = "GRP_IGBP")
```

---

amf_filter_base *Filter AmeriFlux BASE data based on plausible range*

Description

The function filters BASE data based on the expected plausible ranges specified for each variable. See AmeriFlux web site https://ameriflux.lbl.gov/data/data-processing-pipelines/data-qaqc/ for description of the plausible ranges.

Usage

```
amf_filter_base(
  data_in,
  limit_ls = NULL,
  basename_decode = NULL,
  loose_filter = 0.05
)
```

Arguments

data_in	A data frame containing BASE data, e.g., import from amf_read_base.
limit_ls	A data frame with at least three columns: • Name: variable base name • Min: expected lower bound • Max: expected upper bound If not specified, use amf_variables by default.
basename_decode	A data frame with at least two columns: • variable_name: actual variable name • basename: variable base name If not specified, use amf_parse_basename by default.
loose_filter	A number in ratio (0-1) used to adjust the physical range for filtering. Set it to 0 if not used. The default is 0.05.

Value

A data frame similar to data_in filtered out off-range points

Examples

```
## Not run:
# read the BASE from a csv file
base <- amf_read_base(file = system.file("extdata",
                                         "AMF_US-CRT_BASE_HH_2-5.csv",
                                          package = "amerifluxr"),
                      unzip = FALSE,
                      parse_timestamp = FALSE)

# filter data, using default physical range +/- 5% buffer
base_f <- amf_filter_base(data_in = base)

# filter data, using default physical range without buffer
base_f <- amf_filter_base(data_in = base, loose_filter = 0)

## End(Not run)
```

`amf_list_data` *Get BASE data variable availability*

Description

This function obtains the BASE data availability for all or selected AmeriFlux sites. See AmeriFlux page https://ameriflux.lbl.gov/data/aboutdata/data-variables/ for details about the variable naming.

Usage

```
amf_list_data(site_set = NULL, var_set = NULL)
```

Arguments

`site_set`	A scalar or vector of character specifying the target AmeriFlux Site ID (CC-Sss). If not specified, it returns all sites.
`var_set`	A scalar or vector of character specifying the target variables as in basename. See AmeriFlux pagehttps://ameriflux.lbl.gov/data/aboutdata/data-variables/#base for a list of variable names. If not specified, it returns all variables.

Value

A data frame of variable-specific data availability (per year) for selected AmeriFlux sites.

- Site_ID - Six character site identifier (CC-Sss)
- VARIABLE - Variable name of the data included in the BASE file
- BASENAME - Variable base name of the data included in the BASE file.
- GAP_FILLED - Whether a variable is a gap-filled variable (TRUE/FALSE)
- Y1990 - Percentage of data availability in the year 1990 (0-1).
- Y1991 - Percentage of data availability in the year 1991 (0-1).
- Y1992 - Percentage of data availability in the year 1992 (0-1).
- ...

Examples

```
## Not run:
# obtain the data variable availability for all sites
data_aval <- amf_list_data()

# obtain the data variable availability for selected sites
data_aval <- amf_list_data(site_set = c("US-CRT","US-WPT"))

# obtain the data variable availability for selected variables
data_aval <- amf_list_data(var_set = c("FCH4", "WTD"))


## End(Not run)
```

`amf_list_metadata` *Get metadata availability*

Description

This function obtains the metadata (i.e., BADM) availability for all or selected AmeriFlux sites. See AmeriFlux page https://ameriflux.lbl.gov/data/badm/ for details about the BADM.

Usage

```
amf_list_metadata(site_set = NULL, group_only = TRUE)
```

Arguments

`site_set`	A scalar or vector of character specifying the target AmeriFlux Site ID (CC-Sss). If not specified, it returns all sites.
`group_only`	Logical (TRUE/FALSE). Should it return availability for BADM variable groups or variables? BADM Groups contain Variables that describe related metadata or an observation with related metadata.

Value

A data frame of data variable availability (per year) for selected AmeriFlux sites. The first column contains the SITE ID. The remaining columns contains the number of entries for a variable or a variable group, with column names specifying the BADM variable or group names.

- Site_ID - Six character site identifier (CC-Sss)
- ...

Examples

```
## Not run:
# obtain the metadata availability for all sites, at variable group levels
metadata_aval <- amf_list_metadata()

# obtain the metadata availability for selected sites, at variable levels
metadata_aval <- amf_list_metadata(site_set = c("US-CRT","US-WPT"),
group_only = FALSE)

## End(Not run)
```

`amf_parse_basename` *Parse BASE data variable name and qualifier*

Description

This function parse variable names and qualifiers of AmeriFlux BASE data product. See AmeriFlux web page https://ameriflux.lbl.gov/data/aboutdata/data-variables/ about the details of variable naming and qualifiers.

Usage

```
amf_parse_basename(var_name, FP_ls = NULL, gapfill_postfix = "_PI_F")
```

Arguments

`var_name`	A vector of variable names (character) to be parsed
`FP_ls`	A vector of standard variable names. If not specified, use `amf_variables` by default to get the latest list.
`gapfill_postfix`	A scalar of expected suffix (character) appended to a variable that is gap-filled. The default is "_PI_F".

Value

A data frame containing the parsed results for all variables in `var_name`:

- variable_name - original variable name
- basename - associated basename, w/o qualifier
- qualifier_gf - qualifier associated with gap-filling
- qualifier_pi - qualifier associated with PI version, excluding gap-filling
- qualifier_pos - qualifier associated with position
- qualifier_ag - qualifier associated with layer-aggregation, e.g., _N, _SD
- layer_index - layer index provided, if any
- H_index - H index provided, if any
- V_index - V index provided, if any
- R_index - R index provided, if any
- is_correct_basename - is the parsed basename recognized in FP-Standard
- is_pi_provide - is this a PI provided variable e.g., _PI
- is_gapfill - is this a gap-filled variable, _PF_F or _F
- is_fetch - is this a fetch quantile variable, e.g., FETCH_70...
- is_layer_aggregated - is this a layer-integrated variable, i.e., _#
- is_layer_SD - is this a standard deviation of layer-integrated variable, i.e., spatial variability

- is_layer_number - is this a number of samples of layer-integrated variable, i.e., spatial variability
- is_replicate_aggregated - is this a replicate-averaged variable, e.g., _1_1_A
- is_replicate_SD - is this a standard deviation of replicate-averaged variable, e.g., _1_1_A_SD
- is_replicate_number - is this a number of samples of replicate-averaged variable, e.g., _1_1_A_N
- is_quadruplet - is this a quadruplet, e.g., _1_1_1

See Also

amf_variables

Examples

```
## Not run:
# read the BASE from a csv file
base <- amf_read_base(file = system.file("extdata",
                                         "AMF_US-CRT_BASE_HH_2-5.csv",
                                          package = "amerifluxr"),
                      unzip = FALSE,
                      parse_timestamp = FALSE)

# parse variable names/qualifiers
basename_decode <- amf_parse_basename(var_name = colnames(base))

## End(Not run)
```

---

amf_plot_datasummary *Plot data summary*

Description

This function visualizes the BASE data summary for selected AmeriFlux sites and variables. This is a wrapper around amf_summarize_data. However, it is strongly advised to subset either sites or variables for faster processing and better visualization.

Usage

```
amf_plot_datasummary(
  data_sum = NULL,
  site_set = NULL,
  var_set = NULL,
  nonfilled_only = TRUE,
  show_cluster = FALSE,
  scale = FALSE
)
```

Arguments

data_sum	A data frame with following columns: • Site_ID - Six character site identifier (CC-Sss) • VARIABLE - Variable name of the data included in the BASE file • BASENAME - Variable base name of the data included in the BASE file. • GAP_FILLED - Whether a variable is a gap-filled variable (TRUE/FALSE) • Any statistics, e.g., P01, P05... output from amf_summarize_data If not specified, use amf_summarize_data by default.
site_set	A scalar or vector of character specifying the target AmeriFlux Site ID (CC-Sss). If not specified, it returns all sites.
var_set	A scalar or vector of character specifying the target variables as in basename. See AmeriFlux pagehttps://ameriflux.lbl.gov/data/aboutdata/data-variables/#base for a list of variable names. If not specified, it returns all variables.
nonfilled_only	Logical, whether only showing non-filled variables, or both non- and gap-filled variables. The default is TRUE.
show_cluster	Logical, whether showing clustering (dendrogram) of site-variables. The default is FALSE.
scale	Logical, whether the values should be centered and scaled among site-variables. The default is FALSE.

Value

An object of class 'plotly' from heatmaply

See Also

amf_summarize_data, heatmaply

Examples

```
## Not run:
# plot data summary for selected variables at two sites
amf_plot_datasummary(site_set = c("US-CRT", "US-WPT"),
                     var_set = c("H", "LE", "NETRAD"))

# plot data summary for FCH4 at all sites, and show
#  clustering among sites
amf_plot_datasummary(var_set = "FCH4",
                     show_cluster = TRUE)

# plot data summary for TA at all grassland sites,
#  and show clustering among sites
sites <- amf_site_info()
sites <- subset(sites, IGBP == "CRO")
amf_plot_datasummary(site_set = sites$SITE_ID,
                     var_set = "TA",
                     show_cluster = TRUE)
```

```
# normalize TA among sites
amf_plot_datasummary(site_set = sites$SITE_ID,
                     var_set = "TA",
                     show_cluster = TRUE,
                     scale = TRUE)

## End(Not run)
```

`amf_plot_datayear` *Plot data availability*

Description

This function visualizes the BASE data availability for selected AmeriFlux sites, variables, and years. This is a wrapper around `amf_list_data`. However, it is strongly advised to subset the sites, variables, and/or years for faster processing and better visualization.

Usage

```
amf_plot_datayear(
  data_aval = NULL,
  site_set = NULL,
  var_set = NULL,
  nonfilled_only = TRUE,
  year_set = NULL
)
```

Arguments

`data_aval`	A data frame with at least five columns: • SITE_ID: • VARIABLE: • BASENAME: variable basename • GAP_FILLED • Y1990: Percentage of data availability in the year 1990 (0-1). • ... If not specified, use `amf_list_data` by default.
`site_set`	A scalar or vector of character specifying the target AmeriFlux Site ID (CC-Sss). If not specified, it returns all sites.
`var_set`	A scalar or vector of character specifying the target variables as in basename. See AmeriFlux page`https://ameriflux.lbl.gov/data/aboutdata/data-variables/#base` for a list of variable names. If not specified, it returns all variables.
`nonfilled_only`	Logical, whether only showing non-filled variables, or both non- and gap-filled variables. The default is TRUE.
`year_set`	A scalar or vector of integers. If not specified, it plots only years with any available data in selected sites and variables

Value

An object of class 'plotly' from heatmaply

See Also

amf_list_data, heatmaply

Examples

```
## Not run:
# plot data availability for all variables at a single site
#  in all years
amf_plot_datayear(site_set = "US-CRT",
                  nonfilled_only = FALSE)

# plot data availability for non-filled FCH4 and WTD at all
#  sites in all years
amf_plot_datayear(var_set = c("FCH4", "WTD"),
                  nonfilled_only = TRUE)

# plot data availability for non-filled FCH4 at all sites
#  in 2018-2020
amf_plot_datayear(var_set = "FCH4",
                  year_set = c(2018:2020),
                  nonfilled_only = TRUE)

## End(Not run)
```

---

amf_read_base *Read AmeriFlux BASE data product*

Description

This function read in the BASE data file downloaded from AmeriFlux. See AmeriFlux web page https://ameriflux.lbl.gov/data/data-processing-pipelines/base-publish/ for details about BASE data product. Use amf_variables to get a list of standard variable names and units.

Usage

```
amf_read_base(file, unzip = TRUE, parse_timestamp = FALSE)
```

Arguments

file	A BASE data file, either in a zipped file or a comma-separate value (csv) file
unzip	Logical, whether to unzip. The default is TRUE. Set FALSE if reading from a previously unzipped csv file.
parse_timestamp	Logical, whether to parse the time stamp. Set TRUE to parse and add timekeeping columns.

Value

A data frame containing data. See AmeriFlux website https://ameriflux.lbl.gov/data/aboutdata/data-variables/ for details about file format, variable definition, units, and convention. If parse_timestamp = TRUE, the following six time-keeping columns are added in the returned data frame:

- YEAR - Year (YYYY)
- MONTH - Month (MM)
- DAY - Day of the month (DD)
- DOY - Day of the year (DDD)
- HOUR - Hour of the day (HH), based on the middle time of the interval
- MINUTE - Minute of the hour (mm), based on the middle time of the interval
- TIMESTAMP - An object of class "POSIXlt" in the UTC time zone, based on the middle time of the interval

See Also

amf_parse_basename, amf_filter_base

Examples

```
# read the BASE from a zip file, using the example data file
base <- amf_read_base(file = system.file("extdata",
                                         "AMF_US-CRT_BASE-BADM_2-5.zip",
                                          package = "amerifluxr"),
                      unzip = TRUE,
                      parse_timestamp = TRUE)

# read the BASE from a csv file
base <- amf_read_base(file = system.file("extdata",
                                         "AMF_US-CRT_BASE_HH_2-5.csv",
                                          package = "amerifluxr"),
                      unzip = FALSE,
                      parse_timestamp = FALSE)
```

---

amf_read_bif *Read AmeriFlux BADM data product*

Description

This function read in the BADM data file formatted in BADM Interchange Format (BIF).

Usage

```
amf_read_bif(file)
```

Arguments

file	A BADM data file

Value

A data frame containing the following 5 columns. See AmeriFlux website https://ameriflux.lbl.gov/data/aboutdata/badm-data-product/ for details.

- SITE_ID - Six character site identifier (CC-Sss)
- GROUP_ID - A unique identifier for data belonging to the same instance of a reported variable group
- VARIABLE_GROUP - A set of variables that are reported together
- VARIABLE - The variable name
- DATAVALUE - The reported value of a variable

Examples

```
# read the BADM BIF file, using an example data file
bif <- amf_read_bif(file = system.file("extdata",
                                       "AMF_AA-Flx_BIF_CCBY4_20201218.xlsx",
                                        package = "amerifluxr"))
```

---

amf_sites *Lists AmeriFlux sites*

Description

Lists available site (names) and basic meta-data

Usage

```
amf_sites()
```

Value

A data frame containing the following columns. See AmeriFlux BADM standard https://ameriflux.lbl.gov/data/badm/ for detailed explanation.

- SITE_ID - Six character site identifier (CC-Sss)
- SITE_NAME - Site name (free text)
- COUNTRY - Country (free text)
- STATE - State (free text)
- IGBP - Vegetation type based on the IGBP definition (character)
- URL_AMERIFLUX - Site web site URL, maintained by AmeriFlux (URL)
- TOWER_BEGAN - The starting year of flux measurement (YYYY)

- TOWER_END - The ending year of flux measurement (YYYY), NA if still active or unspecified
- LOCATION_LAT - Latitude of the site (decimal deg ref WGS84)
- LOCATION_LONG - Longitude of the site (decimal deg ref WGS84)
- LOCATION_ELEV - Elevation of the site above sea level (m)
- CLIMATE_KOEPPEN - Koppen climate classification (character)
- MAT - Long-term mean annual average air temperature (degree C)
- MAP - Long-term mean annual average precipitation (mm)
- DATA_POLICY - LEGACY / CCBY4.0 (character)

Examples

```
## Not run:
# download a list of sites and basic info
sites <- amf_sites()


## End(Not run)
```

`amf_site_info` *Get AmeriFlux site general info*

Description

This function obtains the latest AmeriFlux site list and sites' general info through the AmeriFlux web service.

Usage

```
amf_site_info()
```

Details

This combines the information of various other functions.

Value

A data frame containing the following columns. See AmeriFlux BADM standard https://ameriflux.lbl.gov/data/badm/ for detailed explanation.

- SITE_ID - Six character site identifier (CC-Sss)
- SITE_NAME - Site name (free text)
- COUNTRY - Country (free text)
- STATE - State (free text)

- IGBP - Vegetation type based on the IGBP definition (character)
- URL_AMERIFLUX - Site web site URL, maintained by AmeriFlux (URL)
- TOWER_BEGAN - The starting year of flux measurement (YYYY)
- TOWER_END - The ending year of flux measurement (YYYY), NA if still active or unspecified
- LOCATION_LAT - Latitude of the site (decimal deg ref WGS84)
- LOCATION_LONG - Longitude of the site (decimal deg ref WGS84)
- LOCATION_ELEV - Elevation of the site above sea level (m)
- CLIMATE_KOEPPEN - Koppen climate classification (character)
- MAT - Long-term mean annual average air temperature (degree C)
- MAP - Long-term mean annual average precipitation (mm)
- DATA_POLICY - LEGACY / CCBY4.0 (character)
- DATA_START - The starting year with published AmeriFlux BASE data (YYYY)
- DATA_END - The ending year with published AmeriFlux BASE data (YYYY)

Examples

```
## Not run:
# obtain the basic general info for all sites
site <- amf_site_info()

## End(Not run)

## End(Not run)
```

`amf_summarize_data` *Get BASE data summary*

Description

This function obtains the BASE data summary for all or selected AmeriFlux sites. See AmeriFlux page https://ameriflux.lbl.gov/data/aboutdata/data-variables/ for details about the variable naming.

Usage

```
amf_summarize_data(site_set = NULL, var_set = NULL)
```

Arguments

`site_set`	A scalar or vector of character specifying the target AmeriFlux Site ID (CC-Sss). If not specified, it returns all sites.
`var_set`	A scalar or vector of character specifying the target variables as in basename. See AmeriFlux pagehttps://ameriflux.lbl.gov/data/aboutdata/data-variables/#base for a list of variable names. If not specified, it returns all variables.

Value

A data frame of site-specific variable summary statistics (selected percentiles) for selected AmeriFlux sites.

- Site_ID - Six character site identifier (CC-Sss)
- VARIABLE - Variable name of the data included in the BASE file
- BASENAME - Variable base name of the data included in the BASE file.
- GAP_FILLED - Whether a variable is a gap-filled variable (TRUE/FALSE)
- DATA_RECORD - Number of supposed data record (counts)
- DATA_MISSING - Number of missing data record (counts)
- Q01 - 1th percentile of the data
- Q05 - 5th percentile of the data
- ...
- Q95 - 95th percentile of the data
- Q99 - 99th percentile of the data

Examples

```
## Not run:
# obtain the data variable availability for all sites & variables
data_sum <- amf_summarize_data()

# obtain the data variable availability for selected sites, all variables
data_sum <- amf_summarize_data(site_set = c("US-CRT","US-WPT"))

# obtain the data variable availability for selected sites & variables
data_sum <- amf_summarize_data(site_set = c("US-CRT","US-WPT"),
                               var_set = c("FC", "LE", "H"))

## End(Not run)
```

`amf_variables` *Get FP (Flux-Processing) Standard Variable List*

Description

This function obtains the latest AmeriFlux FP (Flux-Processing) standard variable list. FP standard defines the variable names and units used for continuously sampled data within the AmeriFlux. Also see AmeriFlux Data Variables page https://ameriflux.lbl.gov/data/aboutdata/data-variables/ for details.

Usage

```
amf_variables()
```

Value

A data frame containing the following columns:

- Name - Standard variable name
- Description - Description of the variable
- Units - Standard variable unit
- Min - Expected minimal value
- Max - Expected maximal value

Examples

```
## Not run:
# download the list of standard variable names and units
FP_ls <- amf_variables()

## End(Not run)
```

---

amf_var_info *Get variable information*

Description

This function obtains the measurement height metadata for the AmeriFlux BASE data product. See AmeriFlux page https://ameriflux.lbl.gov/data/measurement-height/ for details.

Usage

```
amf_var_info(out_dir = tempdir(), verbose = TRUE)
```

Arguments

out_dir	The output directory (default = tempdir())
verbose	Logical, whether to show download progress (TRUE/FALSE)

Value

A data frame of measurement height data for all AmeriFlux sites #'

- Site_ID - Six character site identifier (CC-Sss)
- Variable - Variable name of the data included in the BASE file
- Start_Date - Date when the information first applies
- Height - Distance above the ground surface in meters
- Instrument_Model - Instrument model used to collect the data variable
- Instrument_Model2 - A second instrument model used to collect the data variable
- Comment - Additional information provided by the site team
- BASE_Version - The most recent BASE data product version number for which the information applies

Examples

```
## Not run:
# download the measurement height data for all sites
var_info <- amf_var_info()

## End(Not run)
```

Index