Package 'NEONiso'

October 12, 2022

Type Package

Title Tools to Calibrate and Work with NEON Atmospheric Isotope Data

Version 0.6.1

Maintainer Rich Fiorella <rfiorella@lanl.gov>

Depends R (>= 4.0.0)

Description Functions for downloading,
calibrating, and analyzing atmospheric isotope data bundled
into the eddy covariance data products of the National Ecological
Observatory Network (NEON) <https://www.neonscience.org>.
Calibration tools are provided for carbon and water isotope products.
Carbon isotope calibration details are found in Fiorella et al. (2021)
<doi:10.1029/2020JG005862>, and the readme
file at <https://github.com/lanl/NEONiso>. Tools for calibrating water
isotope products have been added as of 0.6.0, but have known deficiencies
and should be considered very experimental currently.

License GPL-3

BugReports https://github.com/lanl/NEONiso/issues

URL https://github.com/lanl/NEONiso

Encoding UTF-8

RoxygenNote 7.2.1

Imports dplyr, zoo, httr, lubridate, neonUtilities (>= 2.0.1),
magrittr, rhdf5 (>= 2.33.7), R.utils, tidyselect, data.table,
rlang, lifecycle, caret, ggplot2, gridExtra

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Language en-US

Config/testthat/edition 3

NeedsCompilation no

Author Rich Fiorella [aut, cre] (<https://orcid.org/0000-0002-0824-4777>)

Repository CRAN

Date/Publication 2022-09-22 19:50:02 UTC

R topics documented:

`calculate_12CO2` *calculate_12CO2*

Description

calculate_12CO2

Usage

```
calculate_12CO2(total_co2, delta13C, f = 0.00474)
```

Arguments

`total_co2`	Vector of CO2 mole fractions.
`delta13C`	Vector of d13C values.
`f`	Fraction of CO2 that is not 12CO2 or 13CO2. Assumed fixed at 0.00474

Value

Vector of 12CO2 mole fractions.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Examples

```
calculate_12CO2(total_co2 = 410, delta13C = -8.5)
```

`calculate_13CO2` *calculate_13CO2*

Description

calculate_13CO2

Usage

```
calculate_13CO2(total_co2, delta13C, f = 0.00474)
```

Arguments

`total_co2`	Vector of CO2 mole fractions.
`delta13C`	Vector of d13C values.
`f`	Fraction of CO2 that is not 12CO2 or 13CO2. Assumed fixed at 0.00474

Value

Vector of 13CO2 mole fractions.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Examples

```
calculate_13CO2(total_co2 = 410, delta13C = -8.5)
```

`calibrate_ambient_carbon_Bowling2003`
calibrate_ambient_carbon_Bowling2003

Description

calibrate_ambient_carbon_Bowling2003

Usage

```
calibrate_ambient_carbon_Bowling2003(
  amb_data_list,
  caldf,
  site,
  filter_data = TRUE,
  force_to_end = TRUE,
  force_to_beginning = TRUE,
  gap_fill_parameters = FALSE,
  r2_thres = 0.9
)
```

Arguments

amb_data_list	List containing an ambient d13C dataset. Will include all variables in 000_0x0_xxm. (character)
caldf	Calibration data frame containing gain and offset values for 12C and 13C isotopologues.
site	Four-letter NEON code corresponding to site being processed.
filter_data	Apply median absolute deviation filter from Brock 86 to remove impulse spikes? Inherited from calibrate_ambient_carbon_Bowling2003()
force_to_end	In given month, calibrate ambient data later than last calibration, using the last calibration? (default true)
force_to_beginning	In given month, calibrate ambient data before than first calibration, using the first calibration? (default true)
gap_fill_parameters	Should function attempt to 'gap-fill' across a bad calibration by carrying the last known good calibration forward? Implementation is fairly primitive currently, as it only carries the last known good calibration that's available forward rather than interpolating, etc. Default FALSE.
r2_thres	Minimum r2 value for calibration to be considered "good" and applied to ambient data.

Value

Depends on write_to_file argument. If true, returns nothing to environment; but returns calibrated ambient observations to the output file. If false, returns modified version of amb_data_list that include calibrated ambient data.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Function called by calibrate_carbon_bymonth() to apply gain and offset parameters to the ambient datasets (000_0x0_09m and 000_0x0_30m). This function should generally not be used independently, but should be used in coordination with calibrate_carbon_bymonth().

`calibrate_ambient_carbon_linreg` *calibrate_ambient_carbon_linreg*

Description

calibrate_ambient_carbon_linreg

Usage

```
calibrate_ambient_carbon_linreg(
  amb_data_list,
  caldf,
  outname,
  site,
  file,
  filter_data = TRUE,
  force_to_end = TRUE,
  force_to_beginning = TRUE,
  gap_fill_parameters = FALSE,
  r2_thres = 0.9
)
```

Arguments

`amb_data_list`	List containing an ambient d13C dataset. Will include all variables in 000_0x0_xxm. (character)
`caldf`	Calibration data frame containing gain and offset values for 12C and 13C isotopologues.
`outname`	Output variable name. Inherited from `calibrate_ambient_carbon_linreg`
`site`	Four-letter NEON code corresponding to site being processed.
`file`	Output file name. Inherited from `calibrate_ambient_carbon_linreg`
`filter_data`	Apply median absolute deviation filter from Brock 86 to remove impulse spikes? Inherited from `calibrate_ambient_carbon_linreg`
`force_to_end`	In given month, calibrate ambient data later than last calibration, using the last calibration? (default true)
`force_to_beginning`	In given month, calibrate ambient data before than first calibration, using the first calibration? (default true)
`gap_fill_parameters`	Should function attempt to 'gap-fill' across a bad calibration by carrying the last known good calibration forward? Implementation is fairly primitive currently, as it only carries the last known good calibration that's available forward rather than interpolating, etc. Default FALSE.
`r2_thres`	Minimum r2 value for calibration to be considered "good" and applied to ambient data.

Value

Nothing to environment; returns calibrated ambient observations to the output file. This function is not designed to be called on its own, and is not exported to the namespace.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Function called by `calibrate_ambient_carbon_linreg` to apply gain and offset parameters to the ambient datasets (000_0x0_09m and 000_0x0_30m). This function should generally not be used independently, but should be used with `calibrate_ambient_carbon_linreg`.

---

calibrate_ambient_water_linreg
calibrate_ambient_water_isotopes

Description

calibrate_ambient_water_isotopes

Usage

```
calibrate_ambient_water_linreg(
  amb_data_list,
  caldf,
  outname,
  site,
  file,
  filter_data,
  force_to_end,
  force_to_beginning,
  r2_thres
)
```

Arguments

`amb_data_list`	List containing ambient d18O/d2H datasets. Will include all variables in 000_0x0_xxm. (character)
`caldf`	Calibration data frame containing slope and intercept values for d18O and d2H values.
`outname`	Output variable name. Inherited from `calibrate_ambient_water_linreg`
`site`	Four-letter NEON code corresponding to site being processed.
`file`	Output file name. Inherited from `calibrate_ambient_water_linreg`
`filter_data`	Apply a median filter to output ambient data? inherited.
`force_to_end`	In given month, calibrate ambient data later than last calibration, using the last calibration? (default true)

force_to_beginning
: In given month, calibrate ambient data before than first calibration, using the first calibration? (default true)

r2_thres
: Minimum r2 value for calibration to be considered "good" and applied to ambient data.

Value

Nothing to environment; returns calibrated ambient observations to the output file. This function is not designed to be called on its own.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Function called by `calibrate_ambient_water_linreg` to apply slope and intercept parameters to the ambient datasets (000_0x0_09m and 000_0x0_30m) to correct to the VSMOW scale. This function should generally not be used independently, but should be used with `calibrate_ambient_water_linreg`. Note that in this version *NO CORRECTION FOR HUMIDITY* is performed. Use with caution.

`calibrate_carbon` *calibrate_carbon*

Description

[Experimental] This function drives a workflow that reads in NEON carbon isotope data of atmospheric CO2, calibrates it to the VPDB scale, and (optionally) writes the calibrated data to a new HDF5 file. Two different approaches are possible: a) a calibration on 12CO2 and 13CO2 isotopologues independently, after Bowling et al. 2003 (Agr. For. Met.), or b) a direct calibration of d13C and CO2 values using linear regression. The vast majority of the time the results generated from either method are extremely similar to each other. Wen et al. 2013 compared several different carbon isotope calibration techniques and found this to be the superior method under most circumstances. We also found this to be the case for NEON data (Fiorella et al. 2021; JGR-Biogeosciences).

Usage

```
calibrate_carbon(
  inname,
  outname,
  site,
  method = "Bowling_2003",
  calibration_half_width = 0.5,
  force_cal_to_beginning = TRUE,
  force_cal_to_end = TRUE,
  gap_fill_parameters = FALSE,
  filter_ambient = TRUE,
  r2_thres = 0.95,
  correct_refData = TRUE,
```

```
  write_to_file = TRUE,
  remove_known_bad_months = TRUE,
  plot_regression_data = FALSE,
  plot_directory = NULL
)
```

Arguments

inname	Input file(s) that are to be calibrated. If a single file is given, output will be a single file per site per month. If a list of files corresponding to a timeseries at a given site is provided, will calibrate the whole time series.
outname	Name of the output file. (character)
site	Four letter NEON site code for site being processed. (character)
method	Are we using the Bowling et al. 2003 method ("Bowling_2003") or direct linear regression of d13C and CO2 mole fractions ("linreg")?
calibration_half_width	Determines the period (in days) from which reference data are selected (period is 2*calibration_half_width).
force_cal_to_beginning	Extend first calibration to the beginning of the file? (default true)
force_cal_to_end	Extend last calibration to the end of the file? (default true)
gap_fill_parameters	Should function attempt to 'gap-fill' across a bad calibration by carrying the last known good calibration forward? Implementation is fairly primitive currently, as it only carries the last known good calibration that's available forward rather than interpolating, etc. Default FALSE.
filter_ambient	Apply the median absolute deviation filter (Brock 86) to remove impulse spikes in output ambient data? (logical; default true)
r2_thres	Minimum r2 threshold of an "acceptable" calibration. Acts to remove calibration periods where a measurement error makes relationship nonlinear. Default = 0.95
correct_refData	NEON has indicated there are a few instances where reported d13C or CO2 reference values are wrong. If set to true, correct known incorrect values. This argument will (hopefully, eventually) go away after NEON has fixed the reference database. Users will be warned prior to removal of this argument.
write_to_file	Write calibrated ambient data to file? (Mostly used for testing)
remove_known_bad_months	There are a few site months with known spectral issues where the isotope ratios are likely unrecoverable. This parameter allows removal of these files, but allows them to remain in archive.
plot_regression_data	Default false; this is useful for diagnostics.
plot_directory	Only used if plot_regression_data is TRUE, but specify where to write out diagnostic plot of regression data.

Details

The 'linreg' method simply takes measured and reference d13C and CO2 values and generates a transfer function between them using `lm()`. For the gain-and-offset method, d13C and CO2 values are converted to 12CO2 and 13CO2 mole fractions. Gain and offset parameters are calculated for each isotopologue independently, and are analogous to regression slope and intercepts, but jointly correct for CO2 concentration dependence and place d13C values on the VPDB scale. The gain and offset parameters are defined by:

$$G = (X_{2,ref} - X_{1,ref})/(X_{2,meas} - X_{1,meas})$$

$$O = X_{2,ref} - GX_{2,meas}$$

Calibrated ambient isotopologues are then given as:

$$X_{cal} = X_{meas}G + O$$

Measurements of reference materials were considered "good" if the following conditions were met:

- Measured CO2 concentrations were within 10 ppm of known "reference" concentrations.
- Variance of the CO2 concentration in standard peak was < 5 ppm.
- Measured d13C value must be within 5 per mil of known "reference" d13C value.

The first two criteria are intended to filter out periods where there is a clear issue with the gas delivery system (i.e., nearly empty gas tank, problem with a valve in the manifold, etc.); the third criterion was adopted after visual inspection of data timeseries revealed that often the first standard measurement following an instrument issue had higher-than-expected error. This criterion clips clearly poor values. Selection of these criteria will become a function argument, and therefore customizable, in a future release.

The behavior of this function will be a bit different depending on what is supplied as `inname`. If a single file is provided, the output will be monthly. However, a list of files corresponding to a site can also be provided, and then a single output file per site will be generated.

Value

Returns nothing to the environment, but creates a new output HDF5 file containing calibrated carbon isotope values.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Examples

```
## Not run: fin <- system.file('extdata',
         'NEON.D15.ONAQ.DP4.00200.001.nsae.2019-05.basic.20201020T211037Z.packed.h5',
         package = 'NEONiso', mustWork = TRUE)
calibrate_carbon_bymonth(inname = fin, outname = 'out.h5',
         site = 'ONAQ', write_to_file = FALSE)
calibrate_carbon_bymonth(inname = fin, outname = 'out.h5',
         site = 'ONAQ', method = 'linreg', write_to_file = FALSE)
## End(Not run)
```

calibrate_carbon_bymonth *calibrate_carbon_bymonth*

Description

[Deprecated] This function drives a workflow that reads in NEON carbon isotope data of atmospheric CO2, calibrates it to the VPDB scale, and (optionally) writes the calibrated data to a new HDF5 file. Two different approaches are possible: a) a calibration on 12CO2 and 13CO2 isotopologues independently, after Bowling et al. 2003 (Agr. For. Met.), or b) a direct calibration of d13C and CO2 values using linear regression. The vast majority of the time the results generated from either method are extremely similar to each other. Wen et al. 2013 compared several different carbon isotope calibration techniques and found this to be the superior method under most circumstances. We also found this to be the case for NEON data (Fiorella et al. 2021; JGR-Biogeosciences).

Usage

```
calibrate_carbon_bymonth(
  inname,
  outname,
  site,
  method = "Bowling_2003",
  calibration_half_width = 0.5,
  force_cal_to_beginning = TRUE,
  force_cal_to_end = TRUE,
  gap_fill_parameters = FALSE,
  filter_ambient = TRUE,
  r2_thres = 0.95,
  correct_refData = TRUE,
  write_to_file = TRUE
)
```

Arguments

inname	Name of the input file. (character)
outname	Name of the output file. (character)
site	Four letter NEON site code for site being processed. (character)
method	Are we using the Bowling et al. 2003 method ("Bowling_2003") or direct linear regression of d13C and CO2 mole fractions ("linreg")?
calibration_half_width	Determines the period (in days) from which reference data are selected (period is 2*calibration_half_width).
force_cal_to_beginning	Extend first calibration to the beginning of the file? (default true)
force_cal_to_end	Extend last calibration to the end of the file? (default true)

`gap_fill_parameters`
: Should function attempt to 'gap-fill' across a bad calibration by carrying the last known good calibration forward? Implementation is fairly primitive currently, as it only carries the last known good calibration that's available forward rather than interpolating, etc. Default FALSE.

`filter_ambient`
: Apply the median absolute deviation filter (Brock 86) to remove impulse spikes in output ambient data? (logical; default true)

`r2_thres`
: Minimum r2 threshold of an "acceptable" calibration. Acts to remove calibration periods where a measurement error makes relationship nonlinear. Default = 0.95

`correct_refData`
: NEON has indicated there are a few instances where reported d13C or CO2 reference values are wrong. If set to true, correct known incorrect values. This argument will (hopefully, eventually) go away after NEON has fixed the reference database. Users will be warned prior to removal of this argument.

`write_to_file`
: Write calibrated ambient data to file? (Mostly used for testing)

Details

The 'linreg' method simply takes measured and reference d13C and CO2 values and generates a transfer function between them using `lm()`. For the gain-and-offset method, d13C and CO2 values are converted to 12CO2 and 13CO2 mole fractions. Gain and offset parameters are calculated for each isotopologue independently, and are analogous to regression slope and intercepts, but jointly correct for CO2 concentration dependence and place d13C values on the VPDB scale. The gain and offset parameters are defined by:

$$G = (X_{2,ref} - X_{1,ref})/(X_{2,meas} - X_{1,meas})$$

$$O = X_{2,ref} - GX_{2,meas}$$

Calibrated ambient isotopologues are then given as:

$$X_{cal} = X_{meas}G + O$$

Measurements of reference materials were considered "good" if the following conditions were met:

- Measured CO2 concentrations were within 10 ppm of known "reference" concentrations.
- Variance of the CO2 concentration in standard peak was < 5 ppm.
- Measured d13C value must be within 5 per mil of known "reference" d13C value.

The first two criteria are intended to filter out periods where there is a clear issue with the gas delivery system (i.e., nearly empty gas tank, problem with a valve in the manifold, etc.); the third criterion was adopted after visual inspection of data timeseries revealed that often the first standard measurement following an instrument issue had higher-than-expected error. This criterion clips clearly poor values. Selection of these criteria will become a function argument, and therefore customizable, in a future release.

Value

Returns nothing to the environment, but creates a new output HDF5 file containing calibrated carbon isotope values.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Examples

```
## Not run: fin <- system.file('extdata',
         'NEON.D15.ONAQ.DP4.00200.001.nsae.2019-05.basic.20201020T211037Z.packed.h5',
         package = 'NEONiso', mustWork = TRUE)
calibrate_carbon_bymonth(inname = fin, outname = 'out.h5',
         site = 'ONAQ', write_to_file = FALSE)
calibrate_carbon_bymonth(inname = fin, outname = 'out.h5',
         site = 'ONAQ', method = 'linreg', write_to_file = FALSE)
## End(Not run)
```

---

calibrate_carbon_reference_data
calibrate_carbon_reference_data

Description

calibrate_carbon_reference_data

Usage

```
calibrate_carbon_reference_data(inname, outname, standard, site, calDf)
```

Arguments

inname	Input file name.
outname	Output file name.
standard	Which standard are we working on? Must be "Low", "Med", or "High"
site	NEON 4-letter site code.
calDf	Calibration data frame - this is the output from fit_carbon_regression

Value

Nothing to the environment.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

calibrate_standards_carbon *calibrate_standards_carbon*

Description

Not sure this is used anymore.

Usage

```
calibrate_standards_carbon(
  cal_df,
  ref_df,
  f = 0.00474,
  r2_thres = 0.95,
  correct_bad_refvals = FALSE,
  site,
  refGas
)
```

Arguments

cal_df	Data.frame containing calibration parameters
ref_df	Data.frame containing reference gas measurements
f	Fraction of CO2 isotopologues that are not 12CO2 or 13CO2. Inherited from script calling this function.
r2_thres	Threshold for calibration regression to be used to calibrate standards data. Default is 0.95. Calibrated reference gas measurements occurring during calibration periods with r2 values less than r2_thres will be marked NA.
correct_bad_refvals	Should we correct known/suspected incorrect reference values in the NEON HDF5 files? (Default = FALSE).
site	Four letter NEON site code. Only used if correct_bad_refvals = TRUE.
refGas	One of "low", "med", or "high." Only used if correct_bad_refvals = TRUE.

Value

A data.frame having the same number of rows of cal_df, with additional columns added for calibrated CO2 mole fractions and d13C values.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`calibrate_standards_water` *calibrate_standards_water*

Description

calibrate_standards_water

Usage

```
calibrate_standards_water(cal_df, ref_df, r2_thres = 0.95)
```

Arguments

`cal_df`	Data.frame containing calibration parameters
`ref_df`	Data.frame containing reference gas measurements
`r2_thres`	Threshold for calibration regression to be used to calibrate standards data. Default is 0.95. Calibrated reference gas measurements occurring during calibration periods with r2 values less than `r2_thres` will be marked NA.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`calibrate_water` *calibrate_water*

Description

[Experimental] This function uses NEON validation data to apply drift corrections to measured ambient water isotope ratios. In brief, ambient water isotope ratios are calibrated by generating regressions using reference water measurements bracketing an ambient period. Three reference waters are measured once per day, with several injections per reference water. Due to memory effects, only the last three are used currently to generate calibration equations. Regressions between measured d18O and d2H values and NEON-provisioned known reference values are generated, and used to calibrate the period of ambient measurements between them if the r2 of the regression is greater than a threshold value (by default, this is 0.95). Most of this function deals with selecting the appropriate calibration data and determining calibration quality. This function also contains a wrapper for `calibrate_ambient_water_linreg`, which calibrates the ambient water data using the calibration parameters generated in this function. This function also copies over data in the qfqm and ucrt hdf5 data groups.

Usage

```
calibrate_water(
  inpath,
  outpath,
  site,
  calibration_half_width = 14,
  filter_data = TRUE,
  force_cal_to_beginning = FALSE,
  force_cal_to_end = FALSE,
  r2_thres = 0.95,
  slope_tolerance = 9999
)
```

Arguments

`inpath`	Directory path to input (monthly) NEON HDF5 files.
`outpath`	Directory path to save output data file. (For now, 1 per site).
`site`	Four-letter NEON code for site being processed.
`calibration_half_width`	Determines the range of standard measurements to use in determining the calibration regression dataset. Creates a moving window that is `2*calibration_half_width` days wide. Default is set to 14 for a 28 day moving window.
`filter_data`	Apply median absolute deviation filter from Brock 86 to remove impulse spikes?
`force_cal_to_beginning`	Extend first calibration to the beginning of the file?
`force_cal_to_end`	Extend last calibration to the end of the file?
`r2_thres`	Minimum r2 threshold of an "acceptable" calibration. Acts to remove calibration periods where a measurement error makes relationship nonlinear. Default = 0.95
`slope_tolerance`	How different from 1 should we allow 'passing' regression slopes to be? Experimental parameter, off by default (e.g., default slope parameter = 9999)

Details

IMPORTANT NOTE Currently this function does not apply a correction for humidity dependence of Picarro isotopic measurements. This is because the data to implement these corrections is not yet publicly available. Caution is suggested when analyzing data at low humidities, below ~5000 ppm, with likely higher biases at lower humidity values.

Additionally, please note that this function is meant to work on *all* files for a given site at the same time. A more flexible version that can handle all files or monthly files will be added to a future release.

Value

nothing to the workspace, but creates a new output file of calibrated water isotope data.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`calibrate_water_linreg_bymonth` *calibrate_water_linreg*

Description

[Deprecated] This function uses NEON validation data to apply drift corrections to measured ambient water isotope ratios. In brief, ambient water isotope ratios are calibrated by generating regressions using reference water measurements bracketing an ambient period. Three reference waters are measured once per day, with several injections per reference water. Due to memory effects, only the last three are used currently to generate calibration equations. Regressions between measured d18O and d2H values and NEON-provisioned known reference values are generated, and used to calibrate the period of ambient measurements between them if the r2 of the regression is greater than a threshold value (by default, this is 0.95). Most of this function deals with selecting the appropriate calibration data and determining calibration quality. This function also contains a wrapper for `calibrate_ambient_water_linreg`, which calibrates the ambient water data using the calibration parameters generated in this function. This function also copies over data in the qfqm and ucrt hdf5 data groups.

Usage

```
calibrate_water_linreg_bymonth(
  inname,
  outname,
  site,
  time_diff_betweeen_standards = 1800,
  filter_data = TRUE,
  force_cal_to_beginning = TRUE,
  force_cal_to_end = TRUE,
  r2_thres = 0.95
)
```

Arguments

`inname`	Name of the input file.
`outname`	Name of the output file.
`site`	Four-letter NEON code for site being processed.
`time_diff_betweeen_standards`	Time (in seconds) required between consecutive standard measurements.
`filter_data`	Apply median absolute deviation filter from Brock 86 to remove impulse spikes?
`force_cal_to_beginning`	Extend first calibration to the beginning of the file?

`force_cal_to_end`
: Extend last calibration to the end of the file?

`r2_thres`
: Minimum r2 threshold of an "acceptable" calibration. Acts to remove calibration periods where a measurement error makes relationship nonlinear. Default = 0.95

Details

IMPORTANT NOTE Currently this function does not apply a correction for humidity dependence of Picarro isotopic measurements. This is because the data to implement these corrections is not yet publicly available. Caution is suggested when analyzing data at low humidities, below ~5000 ppm, with likely higher biases at lower humidity values.

Value

nothing to the workspace, but creates a new output file of calibrated carbon isotope data.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`calibrate_water_reference_data`
calibrate_water_reference_data

Description

calibrate_water_reference_data

Usage

```
calibrate_water_reference_data(outname, standard, site, stdDf, calDf)
```

Arguments

`outname`	Output file name.
`standard`	Which reference material is being 'calibrated'? (Low, med, or high)
`site`	NEON 4-letter site code.
`stdDf`	Data frame of reference material measurements.
`calDf`	Calibration data frame - this is the output from fit_water_regression

Value

Nothing to the environment.

carbon_regression_plots *carbon_regression_plots*

Description

carbon_regression_plots

Usage

```
carbon_regression_plots(caldata, plot_filename, method, mtitle)
```

Arguments

caldata	Data frame corresponding to a specific calibration period.
plot_filename	What should the output file name for diagnostic plot be?
method	Which method are we using? Currently only works for gain/offset.
mtitle	Fed from above routine - what should the plot title be?

Value

Nothing to the environment, but a pdf plot to a file.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

convert_NEONhdf5_to_POSIXct_time *convert_NEONhdf5_to_POSIXct_time*

Description

convert_NEONhdf5_to_POSIXct_time

Usage

```
convert_NEONhdf5_to_POSIXct_time(intime)
```

Arguments

intime	Vector of datetimes in NEON data files (as string) to convert to POSIXct class

Value

Vector of datetimes from NEON data file now in POSIXct format.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Examples

```
convert_NEONhdf5_to_POSIXct_time("2019-06-01T12:00:00.000Z")
```

`convert_POSIXct_to_NEONhdf5_time`
convert_POSIXct_to_NEONhdf5_time

Description

Converts a POSIXct object back to the character format used by NEON in their HDF eddy covariance files. Output format, using strptime syntax, is %Y-%m-%dT%H:%M:%OSZ.

Usage

```
convert_POSIXct_to_NEONhdf5_time(intime)
```

Arguments

`intime`	POSIXct vector to convert to NEON time format.

Value

Returns character version of POSIXct object matching NEON time variable format.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Examples

```
convert_POSIXct_to_NEONhdf5_time(Sys.time())
```

copy_qfqm_group *copy_qfqm_group*

Description

copy_qfqm_group

Usage

```
copy_qfqm_group(data_list, outname, site, file, species)
```

Arguments

data_list	List of groups to retrieve qfqm data from.
outname	Output filename.
site	Four-letter NEON site code.
file	Input filename.
species	CO2 or H2O? Same function used for both CO2 and H2O isotopes.

Value

Nothing to the workspace, but copies qfqm group from input file to output file.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

copy_ucrt_group *copy_ucrt_group*

Description

copy_ucrt_group

Usage

```
copy_ucrt_group(data_list, outname, site, file, species)
```

Arguments

data_list	List of groups to retrieve ucrt data from.
outname	Output file name.
site	NEON 4-letter site code.
file	Input file name.
species	H2O or CO2.

Value

Nothing to the workspace, but copies ucrt group from input file to output file.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`correct_carbon_ref_cval` *correct_carbon_ref_cval*

Description

This ugly function is present out of necessity, and will only exist for as long as it is necessary. It is an internal correction within the NEONiso calibration routines that is required as there are some mismatches between the 'true' isotope reference values and those in the NEON HDF5 files. NEON is working on correcting this, and after it has been corrected, this function has no need to exist and will be immediately deprecated. As a result, this function is fairly messy but there is little incentive to improve it.

Usage

```
correct_carbon_ref_cval(
  std_frame,
  site,
  omit_already_corrected = TRUE,
  co2_tol = 5,
  d13c_tol = 0.25
)
```

Arguments

`std_frame`	Standard data frame to perform swap on.
`site`	NEON four letter site code.
`omit_already_corrected`	Should we attempt correction, if it's already been corrected in the raw files.
`co2_tol`	Tolerance to use to select co2 values that need to be replaced, in ppm. Default = 5 ppm.
`d13c_tol`	Tolerance to use to select d13C values that need to be replaced, in ppm. Default = 0.25 per mil.

Details

Current sites and time periods affected:

Value

A data.frame, based on `std_frame`, where NEON-supplied reference values have been corrected if a mismatch has previously been identified.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`correct_carbon_ref_output` *Correct carbon ref output*

Description

Correct carbon ref output

Usage

```
correct_carbon_ref_output(
  std_list,
  site,
  omit_already_corrected = TRUE,
  co2_tol = 5,
  d13c_tol = 0.25,
  refGas
)
```

Arguments

`std_list`	List containing reference/validation gas measurements.
`site`	Four-letter NEON site code.
`omit_already_corrected`	Skip correction if the reference gas values have already been corrected in the files (default TRUE) If you have older versions of the files, you may want to set this to FALSE.
`co2_tol`	Tolerance used to identify a mismatch in CO2 values. Will correct measured CO2 values within +/- co2_tol within time period identified as having incorrect reference values.
`d13c_tol`	Tolerance used to identify a mismatch in d13C values. Will correct measured d13C values within +/- d13c_tol within time period identified as having incorrect reference values.
`refGas`	Which reference gas is being corrected? Expects "co2High", "co2Med", or "co2Low"

Value

A version of std_list with corrected reference values.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`delta_to_R` *delta_to_R*

Description

delta_to_R

Usage

```
delta_to_R(delta_values, element)
```

Arguments

`delta_values`	A vector of isotope ratios in delta notation.
`element`	Which element to return R values - carbon, oxygen, or hydrogen.

Value

Vector of isotope ratios (R values).

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Examples

```
delta_to_R(delta_values = 0, element = 'oxygen') # returns 2005.2e-6 for VSMOW.
```

`estimate_calibration_error` *estimate_calibration_error*

Description

estimate_calibration_error

Usage

```
estimate_calibration_error(formula, data)
```

Arguments

`formula`	Formula to pass to caret::train to perform cross validation.
`data`	Data frame to perform cross-validation on.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`extract_carbon_calibration_data` *extract_carbon_calibration_data.R*

Description

extract_carbon_calibration_data.R

Usage

```
extract_carbon_calibration_data(data_list)
```

Arguments

`data_list`	List containing data, from the /*/dp01/data/ group in NEON HDF5 file.

Value

Returns data frame of required variables.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

extract_water_calibration_data
extract_water_calibration_data

Description

extract_water_calibration_data

Usage

```
extract_water_calibration_data(
  data_list,
  ucrt_list = NULL,
  standard,
  ucrt_source = "data",
  method = "by_site"
)
```

Arguments

data_list	List containing data, from the /*/dp01/data/ group in NEON HDF5 file.
ucrt_list	List containing uncertainty data, from the /*/dp01/ucrt/ group in NEON HDF5 file. (only works if paired with ucrt_source = 'ucrt' and method = 'by_month')
standard	String indicating whether to grab data from the high, medium, or low standard.
ucrt_source	Where from HDF5 file should variance be extracted from? (Only "data" works now..."ucrt" will throw an error.)
method	Are we calling this function from the calibrate_water_linreg function (use "by_month") or the calibrate_water_linreg_bysite function (use "by_site)

Value

Returns data frame of required variables.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

filter_median_Brock86 *filter_median_Brock86*

Description

Median absolute deviation filter of Brock 1986.

Usage

```
filter_median_Brock86(data, width = 7, threshold = 5)
```

Arguments

data	Vector to filter.
width	Width of filter, in rows.
threshold	Only filter values that are abs(threshold) away from median

Value

Returns filtered vector.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

fit_carbon_regression *fit_carbon_regression*

Description

fit_carbon_regression

Usage

```
fit_carbon_regression(
  ref_data,
  method,
  calibration_half_width,
  plot_regression_data = FALSE,
  plot_dir = "/dev/null",
  site
)
```

Arguments

ref_data	Reference data.frame from which to estimate calibration parameters.
method	Are we using the Bowling et al. 2003 method ("Bowling_2003") or direct linear regression of d13C and CO2 mole fractions ("linreg")?
calibration_half_width	Determines the period (in days) from which reference data are selected (period is 2*calibration_half_width).
plot_regression_data	True or false - should we plot the data used in the regression? Useful for debugging.
plot_dir	If plot_regression_data is true, where should the plots be saved?
site	Needed for regression plots.

Value

Returns a data.frame of calibration parameters. If `method == "Bowling_2003"`, then data.frame includes gain and offset parameters for 12CO2 and 13CO2, and r^2 values for each regression. If `method == "linreg"`, then data.frame includes slope, intercept, and r^2 values for d13C and CO2 values.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

---

fit_water_regression *fit_water_regression*

Description

fit_water_regression

Usage

```
fit_water_regression(stds, calibration_half_width, slope_tolerance, r2_thres)
```

Arguments

stds	Reference data.frame from which to estimate calibration parameters.
calibration_half_width	Determines the period (in days) from which reference data are selected (period is 2*calibration_half_width).
slope_tolerance	Allows for filtering of slopes that deviate from 1 by slope_tolerance.
r2_thres	What is the minimum r2 value permitted in a 'useful' calibration relationship.

Value

Returns a data.frame of calibration parameters. Output data.frame includes slope, intercept, and r^2 values for d13C and CO2 values.

`get_Rstd` *get_Rstd*

Description

get_Rstd

Usage

```
get_Rstd(element)
```

Arguments

`element`	Which element to return standard ratio - carbon, oxygen, or hydrogen.

Value

Heavy-to-light isotope ratio of most common stable isotope standard. VSMOW for water, VPDB for carbon.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`ingest_data` *ingest_data*

Description

ingest_data

Usage

```
ingest_data(inname, analyte, name_fix = TRUE)
```

Arguments

`inname`	A file (or list of files) to extract data from for calibration.
`analyte`	Carbon (Co2) or water (H2o)?
`name_fix`	Fix to data frame required for next-generation calibration functions, but breaks old 'by_month()' functions. This parameter provides a necessary work around until these functions are removed.

Value

List of data frames, taken from files specified in `inname`

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

---

loocv *loocv*

Description

loocv

Usage

```
loocv(mod)
```

Arguments

`mod`	Fitted model to estimate leave-one-out CV on.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

helper function for the leave-one-out cross variance

---

manage_local_EC_archive *manage_local_EC_archive*

Description

Utility function to help retrieve new EC data and/or prune duplicates, as NEON provisions new data or re-provisions data for an existing site and month.

Usage

```
manage_local_EC_archive(
  file_dir,
  get = TRUE,
  unzip_files = TRUE,
  trim = FALSE,
  dry_run = TRUE,
  sites = "all"
)
```

Arguments

file_dir	Specify the root directory where the local EC store is kept.
get	Pull down data from NEON API that does not exist locally?
unzip_files	NEON gzips the hdf5 files, should we unzip any gzipped files within file_dir? (Searches recursively)
trim	Search through local holdings, and remove older file where there are duplicates?
dry_run	List files identified as duplicates, but do not actually delete them? Default true to prevent unintended data loss.
sites	Which sites to retrieve data from? Default will be all sites with available data, but can specify a single site or a vector here.

Value

Returns nothing to the environment, but will download new NEON HDF5 files for selected sites (if get = TRUE), unzip them in the local file directory (if unzip_files = TRUE), and identify and remove suspected duplicate files (if trim = TRUE and dry_run = FALSE).

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

NEONiso *NEONiso: A package for calibrating NEON atmospheric isotope observations.*

Description

This package provides functions for retrieving, calibrating, and generating diagnostic plots of NEON atmospheric isotope data.

---

restructure_ambient_data *restructure_ambient_data*

Description

restructure_ambient_data

Usage

```
restructure_ambient_data(inpath, analyte)
```

Arguments

inpath	Folder containing data to stack.
analyte	Carbon (Co2) or water (H2o)?

Value

List of data extracted from files listed in inpath.

---

restructure_ambient_data2 *restructure_ambient_data2*

Description

restructure_ambient_data2

Usage

```
restructure_ambient_data2(inpath, analyte)
```

Arguments

inpath	Folder containing data to stack.
analyte	Carbon (Co2) or water (H2o)?

Value

List of data extracted from files listed in inpath.

---

restructure_carbon_variables *restructure_carbon_variables*

Description

restructure_carbon_variables

Usage

```
restructure_carbon_variables(dataframe, varname, mode, group)
```

Arguments

dataframe	Input data.frame, from neonUtilities::stackEddy
varname	Which variable are we applying this function to? There's a list of ~10 common ones to write to the hdf5 file.
mode	Are we fixing a reference data frame or an ambient data frame?
group	Data, ucrt, or qfqm?

Value

data.frame formatted for output to hdf5 file.

`restructure_water_variables`
restructure_water_variables

Description

restructure_water_variables

Usage

```
restructure_water_variables(dataframe, varname, mode)
```

Arguments

dataframe	Input data.frame, from neonUtilities::stackEddy
varname	Which variable are we applying this function to? There's a list of ~10 common ones to write to the hdf5 file.
mode	Are we fixing a reference data frame or an ambient data frame?

Value

data.frame formatted for output to hdf5 file.

R_to_delta *R_to_delta*

Description

R_to_delta

Usage

```
R_to_delta(R_values, element)
```

Arguments

R_values	A vector of isotope ratios (e.g., R values).
element	Which element to return delta values - carbon, oxygen, or hydrogen.

Value

Vector of isotope ratios in delta notation.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Examples

```
R_to_delta(R_values = 2005.20e-6, element = 'oxygen') # returns 0.
```

---

select_daily_reference_data
select_daily_reference_data

Description

select_daily_reference_data

Usage

```
select_daily_reference_data(standard_df, analyte, min_nobs = NA)
```

Arguments

standard_df	Input reference data.frame.
analyte	Are we calibrating CO2 and H2O? (Use argument 'co2' or 'h2o', or else function will throw error)
min_nobs	Minimum number of high-frequency observations to define a peak. If not supplied, defaults are 200 for `analyte = 'co2'` or 30 for `analyte = 'h2o'`

Value

Smaller data.frame where only the reference data selected to use in the calibration routines is returned. Assumes that we are calibrating on a daily basis, and not on a longer time scale. Data are selected based on two criteria: cannot be missing, and must be at least a certain number of high-frequency observations in order to qualify as a valid measurement. For the water system, this function also keeps only the last three injections for each reference water per day.

`setup_output_file` *setup_output_file*

Description

Creates a skeleton hdf5 file for the calibrated data.

Usage

```
setup_output_file(inname, outname, site, analyte)
```

Arguments

`inname`	Input file name.
`outname`	Output file name.
`site`	NEON 4-letter site code.
`analyte`	Carbon ('Co2') or water ('H2o') system?

Value

Nothing to the environment, but creates a new data file with the most basic output HDF5 structure consistent with NEON's data files.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

swap_standard_isotoperatios
swap_standard_isotoperatios

Description

There are a few suspected instances where the water isotope ratios for oxygen and hydrogen have been flipped in the reference data. This function corrects them until they are corrected in the NEON database using a d-excess filter.

Usage

```
swap_standard_isotoperatios(std_frame, dxs_thres = 500)
```

Arguments

std_frame	Standard data frame to perform swap on.
dxs_thres	d-excess threshold to indicate when to swap.

Value

A data.frame based on std_frame, where d18O and d2H values have been swapped from NEON input files if determined to have a reference value mismatch. Mismatch is determined based on the d-excess of the standard (= d2H - 8*d18O), using a value of 500 by default.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

terrestrial_core_sites
terrestrial_core_sites

Description

terrestrial_core_sites

Usage

```
terrestrial_core_sites()
```

Value

A vector listing NEON core terrestrial sites.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Examples

```
terrestrial_core_sites()
```

---

terrestrial_relocatable_sites *terrestrial_relocatable_sites*

Description

terrestrial_relocatable_sites

Usage

```
terrestrial_relocatable_sites()
```

Value

A vector listing NEON core terrestrial sites.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Examples

```
terrestrial_relocatable_sites()
```

---

validate_analyte *validate_analyte*

Description

validate_analyte

Usage

```
validate_analyte(analyte)
```

Arguments

analyte	Co2 or H2o?

Value

Standardized string for the water ('H2o') or carbon ('Co2') systems to make sure strings are standardized across package functions.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`validate_output_file` *validate_output_file*

Description

validate_output_file

Usage

```
validate_output_file(inname, outname, site, analyte)
```

Arguments

`inname`	Input file name.
`outname`	Output file name.
`site`	NEON 4-letter site code.
`analyte`	Carbon ('Co2') or water ('H2o') system?

Value

Nothing to environment, simply checks to make sure expected groups are in output.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

water_isotope_sites *water_isotope_sites*

Description

water_isotope_sites

Usage

```
water_isotope_sites()
```

Value

A vector listing NEON sites measuring water vapor isotope ratios.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Examples

```
water_isotope_sites()
```

write_carbon_ambient_data *write_carbon_ambient_data*

Description

Write out ambient observations from the NEON EC towers where the isotope data (either H2O or CO2) have been calibrated using this package.

Usage

```
write_carbon_ambient_data(outname, site, amb_data_list)
```

Arguments

outname	Output file name.
site	NEON 4-letter site code.
amb_data_list	Calibrated list of ambient data - this is the output from one of the calibrate_ambient_carbon* functions.

Value

Nothing to the environment, but writes data in amb_data_list to file.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`write_carbon_calibration_data` *write_carbon_calibration_data*

Description

write_carbon_calibration_data

Usage

```
write_carbon_calibration_data(outname, site, calDf, method)
```

Arguments

`outname`	Output file name.
`site`	NEON 4-letter site code.
`calDf`	Calibration data frame - this is the output from fit_carbon_regression
`method`	Was the Bowling et al. 2003 or the linear regression method used in fit_carbon_regression?

Value

Nothing to the environment, but writes out the calibration parameters (e.g., gain and offset or regression slopes and intercepts) to the output hdf5 file.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`write_carbon_reference_data` *write_carbon_reference_data*

Description

write_carbon_reference_data

Usage

```
write_carbon_reference_data(inname, outname, site, calDf)
```

Arguments

inname	Input file name.
outname	Output file name.
site	NEON 4-letter site code.
calDf	Calibration data frame - this is the output from fit_carbon_regression

Value

Nothing to the environment, but writes calibrated reference data to hdf5 file.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

---

write_qfqm *write_qfqm*

Description

Write NEON's qfqm data for an isotope species to output file. Wraps copy_qfqm_group.

Usage

```
write_qfqm(inname, outname, site, analyte)
```

Arguments

inname	Input file name.
outname	Output file name.
site	NEON 4-letter site code.
analyte	Carbon ('Co2') or water ('H2o') system?

Value

Nothing to the environment, but writes qfqm data to file.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

write_ucrt *write_ucrt*

Description

Write NEON's ucrt data for an isotope species to output file. Wraps copy_ucrt_group.

Usage

```
write_ucrt(inname, outname, site, analyte)
```

Arguments

inname	Input file name.
outname	Output file name.
site	NEON 4-letter site code.
analyte	Carbon ('Co2') or water ('H2o') system?

Value

Nothing to the environment, but writes ucrt data to file.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

write_water_calibration_data *write_water_calibration_data*

Description

write_water_calibration_data

Usage

```
write_water_calibration_data(outname, site, calDf)
```

Arguments

outname	Output file name.
site	NEON 4-letter site code.
calDf	Calibration data frame - this is the output from fit_water_regression

Value

Nothing to the environment, but writes out the calibration parameters (e.g., regression slopes and intercepts) to the output hdf5 file.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

`write_water_reference_data` *write_water_reference_data*

Description

write_water_reference_data

Usage

```
write_water_reference_data(inname, outname, site, lowDf, medDf, highDf, calDf)
```

Arguments

`inname`	Input file name.
`outname`	Output file name.
`site`	NEON 4-letter site code.
`lowDf`	Dataframe corresponding to the "low" reference water.
`medDf`	Data frame corresponding to the "med" reference water.
`highDf`	Data frame corresponding to the "high" reference water.
`calDf`	Calibration data frame - this is the output from fit_water_regression

Value

Nothing to the environment, but writes calibrated reference data to hdf5 file.

Author(s)

Rich Fiorella <rfiorella@lanl.gov>

Index