

# Package ‘RFlux’

March 14, 2022

**Type** Package

**Title** Eddy Covariance Flux Data Processing

**Version** 3.2.0

**Date** 2022-02-07

**Author** Domenico Vitale [aut, cre],  
Dario Papale [com, ctb]

**Maintainer** Domenico Vitale <domvit@unitus.it>

**Description** An R graphical user interface for processing eddy covariance raw data and release high quality fluxes of the main GHGs exchanged by ecosystems and agricultural fields. Fluxes are estimated through a call to the open source EddyPro software (registered trademark, LI-COR, Biosciences, 2020). 'RFlux' provides tools for the metadata management as well as for the implementation of the robust data cleaning procedure described by Vitale et al (2020) <doi:10.5194/bg-17-1367-2020>.

**License** GPL(>=2)

**Imports** stats, utils, bit64, data.table, imputeTS, robustbase, zoo,  
xts, stringr, stlplus, egcm, alphaOutlier, timeDate, car,  
robfilter, future, future.apply, lubridate, MASS, gtools,  
tools, diptest, forecast

**Depends** R (>= 3.5.0)

**Suggests** knitr, rmarkdown, markdown

**VignetteBuilder** knitr

**NeedsCompilation** no

## R topics documented:

RFlux-package	2
cleanFlux	3
closed_path_rawdata	8
convert_rawdata	9
despiking	11
ecmd_table	12
ecworkset	14
eddypro_run	15
get_md	17
inst_prob_test	20
mahrt	21

mode . . . . .	22
qcStat . . . . .	22
robust_despiking . . . . .	24
flag_detection . . . . .	25

<b>Index</b>	<b>27</b>
--------------	-----------

---

RFlux-package	<i>Eddy Covariance Flux Data Processing</i>
---------------	---

---

## Description

An R graphical user interface for processing eddy covariance raw data and release high quality fluxes of the main GHGs exchanged by ecosystems and agricultural fields. Fluxes are estimated through a call to the open source EddyPro software (registered trademark, LI-COR, Biosciences, 2020). 'RFlux' provides tools for the metadata management as well as for the implementation of the robust data cleaning procedure described by Vitale et al (2020) <doi:10.5194/bg-17-1367-2020>.

## Details

RFlux package ingests eddy covariance rawdata sampled by either open- or closed-path system and implement the processing pipeline adopted by the ICOS-ETC (Integrated Carbon Observation System European Research Infrastructure - Ecosystem Thematic Center, <http://www.icos-etc.eu/icos/>). All metadata information have to be reported in the *filename\_ecmd.csv* table. Such information are then processed by the `get_md` function.

The processing of rawdata aims at

- i estimating fluxes and other micrometeorological parameters.
- ii performing data quality control.

Flux estimation involves the following options/methods:

- Planar fit for tilt correction (Wilczak et al, 2001).
- Maximum cross-covariance method for time lag determination (see Rebmann et al, 2012).
- Block average (see Moncrieff et al, 2004).
- WPL correction, only for open path systems (Webb et al, 1980).
- In-situ spectral corrections (Fratini et al, 2012).

The open source EddyPro software (registered trademark, LI-COR Biosciences, 2019) is used to this aim employing also the estimation of micrometeorological parameters useful in subsequent analyses. It is required the EddyPro software is installed on your system (for download see [www.licor.com/EddyPro](http://www.licor.com/EddyPro)).

Quality control involves the data cleaning procedure described in Vitale et al (2020). Its implementation involves a three-step procedure

**Step 1:** Estimation of the QC test statistics via the `qcStat` function.

**Step 2:** Generating the workset via the `ecworkset` function.

**Step 3:** Application of data cleaning procedure (including despiking) via the `cleanFlux` function.

*Acknowledgements.* RFlux has been developed in the context of the ICOS Ecosystem Thematic Centre. DV and DP thank the ENVRIPLUS H2020 European project (Grant Agreement 654182) and the ENVRIFAIR H2020 European project (Grant Agreement 824068) for the support.

**Author(s)**

Domenico Vitale [aut, cre], Dario Papale [com, ctb]

Maintainer: Domenico Vitale <domvit@unitus.it>

**References**

Fratini, G., Mauder, M. (2014). Towards a consistent eddy-covariance processing: an intercomparison of EddyPro and TK3. *Atmospheric Measurement Techniques*, 7(7), 2273–2281, doi: <https://doi.org/10.5194/amt-7-2273-2014>.

Fratini, G., Ibrom, A., Arriga, N., Burba, G., Papale, D. (2012). Relative humidity effects on water vapour fluxes measured with closed-path eddy-covariance systems with short sampling lines. *Agricultural and forest meteorology*, 165, pp 53–63, doi: <https://doi.org/10.1016/j.agrformet.2012.05.018>.

LI-COR Biosciences: EddyPro 7.0.4: Help and User's Guide, LI-COR Biosciences, Lincoln, Nebraska USA, [www.licor.com/EddyPro](http://www.licor.com/EddyPro), 2019.

Moncrieff, J., Clement, R., Finnigan, J., Meyers, T. (2004). Averaging, detrending, and filtering of eddy covariance time series. In *Handbook of micrometeorology*, pp. 7–31, Springer, Dordrecht, doi: [https://doi.org/10.1007/1-4020-2265-4\\_2](https://doi.org/10.1007/1-4020-2265-4_2).

Rebmann, C., Kolle, O., Heinesch, B., Queck, R., Ibrom, A., Aubinet, M. (2012). Data acquisition and flux calculations. In *Eddy covariance*, pp. 59–83, Springer, Dordrecht.

Vitale, D., Fratini, G., Bilancia, M., Nicolini, G., Sabbatini, S., Papale, D. (2019). A robust data cleaning procedure for eddy covariance flux measurements, *Biogeosciences*, 17, 1367–1391, doi: <https://doi.org/10.5194/bg-17-1367-2020>, 2020.

Webb, E.K., Pearman, G.I., Leuning, R. (1980). Correction of flux measurements for density effects due to heat and water vapour transfer. *Quarterly Journal of the Royal Meteorological Society*, 106(447), pp 85–100, doi: <https://doi.org/10.1002/qj.49710644707>.

Wilczak, J.M., Oncley, S.P., Stage, S.A. (2001). Sonic anemometer tilt correction algorithms. *Boundary-Layer Meteorology*, 99(1), pp 127–150, doi: <https://doi.org/10.1023/A:1018966204465>.

---

cleanFlux

*Cleaning eddy covariane flux measurements*

---

**Description**

This is the main function of the RFlux library. It performs the data cleaning procedure described by Vitale et al (2019).

**Usage**

```
cleanFlux(path_workset, path_ecmd, path_output=NULL, FileName=NULL,
          plotQC=FALSE, storage=FALSE)
```

**Arguments**

path_workset	path where the workset file generated from function ecworkset is stored.
path_ecmd	path where the eddy covariance metadata file ( <i>CC-xxx_ecmd.csv</i> ) is stored.
path_output	path where the output file will be stored. Default is NULL.

FileName	file name for the output file. Default is NULL.
plotQC	Logical. Should the details of data cleaning procedure be saved. If TRUE two .jpeg files (Details and Synthesis) for each flux variables will be stored in the path specified by path_output. Default is FALSE.
storage	Logical. Should the Net Ecosystem Exchange flux take into account the CO2 storage term. Default is FALSE.

## Details

Comparing statistics estimated by `qcStat` with two threshold values, each test returns one of 3 possible statements:

**SevEr:** if the test provides strong evidence about the presence of a specific source of systematic error.

**ModEr:** if the test provides only weak evidence about the presence of a specific source of systematic error.

**NoEr:** if the test does not provide evidence about the presence of a specific source of systematic error.

The data cleaning is based on a two-step procedure. In the first stage, fluxes that inherited at least one SevEr statement are rejected, while fluxes that inherited no SevEr statements and any number of ModEr statements are retained. In the second stage, flux data that inherited no SevEr statement are subject to an outlier detection procedure and only flux data that are both detected as outlier and inherited at least a ModEr statement are conclusively rejected. This implies that data points that inherited any number of ModEr statements but were not detected to be outliers, as well as outliers which showed no evidence of systematic errors, are retained in the dataset and can be used for any analysis or modeling purposes.

## Value

Returns a dataframe containing:

TIMESTAMP_START	ISO timestamp start of averaging period (format: <code>yyyymmddHHMM</code> ).
TIMESTAMP_END	ISO timestamp end of averaging period (format: <code>yyyymmddHHMM</code> ).
TAU_UNCLEANED	Momentum flux (uncleaned).
TAU	Momentum flux (cleaned).
TAU_DATA_FLAG	Flag for TAU (0: observed flux for which any quality control (QC) tests provided negligible evidences of error; 1: outlying flux rejected because at least one of the QC tests provided a moderate evidence of error; 2: flux removed because at least one of the QC test provided a severe evidence of error).
H_UNCLEANED	Sensible heat turbulent flux (no storage correction, uncleaned).
H	Sensible heat turbulent flux (no storage correction, cleaned).
H_DATA_FLAG	Flag for H (0: observed flux for which any quality control (QC) tests provided negligible evidences of error; 1: outlying flux rejected because at least one of the QC tests provided a moderate evidence of error; 2: flux removed because at least one of the QC test provided a severe evidence of error).
LE_UNCLEANED	Latent heat turbulent flux (no storage correction, uncleaned).
LE	Latent heat turbulent flux (no storage correction, cleaned).

LE_DATA_FLAG	Flag for LE (0: observed flux for which any quality control (QC) tests provided negligible evidences of error; 1: outlying flux rejected because at least one of the QC tests provided a moderate evidence of error; 2: flux removed because at least one of the QC test provided a severe evidence of error).
FC	Carbon Dioxide (CO <sub>2</sub> ) turbulent flux (no storage correction).
SC	Carbon Dioxide (CO <sub>2</sub> ) storage flux.
NEE_UNCLEANED	Net Ecosystem Exchange (uncleaned).
NEE	Net Ecosystem Exchange (cleaned).
NEE_DATA_FLAG	Flag for NEE (i.e., 0: observed flux for which any quality control (QC) tests provided negligible evidences of error; 1: outlying flux rejected because at least one of the QC tests provided a moderate evidence of error; 2: flux removed because at least one of the QC test provided a severe evidence of error).
X_OUTLYING_FLAG	Flag for the X flux variable denoting outliers (0: no outlying flux; 1: outlying flux).
X_OOR_FLAG	Flag for the X flux variable denoting values out of the physically plausible range (0: within range; 2: out of range).
X_FMR_STAT	Fraction of Missing Records in raw, high-frequency, data used for the X flux variable estimation.
X_FMR_FLAG	Flag for the FMR test for the X flux variable (0: negligible evidences of error, IF $X\_FMR < 5$ ; 1: moderate evidences of error, 1 IF $5 \leq X\_FMR \leq 15$ ; 2: severe evidences of error, 2 IF $X\_FMR > 15$ ).
X_LGD_STAT	Longest Gap Duration in raw, high-frequency, data used for the X flux variable estimation.
X_LGD_FLAG	Flag for the LGD test for X flux variable (0: negligible evidences of error, IF $X\_LGD < 90$ ; 1: moderate evidences of error, IF $90 \leq X\_LGD \leq 180$ ; 2: severe evidences of error, IF $X\_LGD > 180$ ).
SA_DIAG_FLAG	Flag for Sonic Anemometer (SA) instrumental diagnostics (0: negligible evidences of error; 2: severe evidences of error).
GA_DIAG_FLAG	Flag for gas analyzer (GA) instrumental diagnostics (0: negligible evidences of error; 2: severe evidences of error).
WD	Wind direction.
WSECT_FLAG	Footprint quality flag indicating periods when wind was blowing from directions known to significantly affect the turbulent flow (0: negligible evidences of error; 2: severe evidences of error).
X_LSR_STAT	Statistic of the Low Signal Resolution test for the X flux variable.
X_LSR_FLAG	Flag for the LSR test for the X flux variable (0: negligible evidences of error, IF $X\_LSR\_STAT > 0.995$ ; 1: moderate evidences of error, IF $0.99 \leq X\_LSR\_STAT \leq 0.995$ ; 2: severe evidences of error, IF $X\_LSR\_STAT < 0.99$ ).
X_HF5_STAT	Statistic of the homogeneity test applied on the X variable (percentage of data exceeding $\mu + / - 5\sigma$ ).
X_HF5_FLAG	Flag for the homogeneity test applied on the X variable (0: negligible evidences of error, IF $X\_HF5\_STAT < 2$ ; 1: moderate evidences of error, IF $2 \leq X\_HF5\_STAT \leq 4$ ; 2: severe evidences of error, IF $X\_HF5\_STAT > 4$ ).
X_HF10_STAT	Statistic of the homogeneity test applied on the X variable (percentage of data exceeding $\mu + / - 10\sigma$ ).

X_HF10_FLAG	Flag for the homogeneity test applied on the X variable (0: negligible evidences of error, IF $X\_HF10\_STAT < 0.5$ ; 1: moderate evidences of error, IF $0.5 \leq X\_HF10\_STAT \leq 1$ ; 2: severe evidences of error, IF $X\_HF10\_STAT > 1$ ).
X_HD5_STAT	Statistic of the homogeneity test applied on differenced X variable (percentage of data exceeding $\mu + / - 5\sigma$ ).
X_HD5_FLAG	Flag for the homogeneity test applied on differenced X variable (0: negligible evidences of error, IF $X\_HD5\_STAT < 2$ ; 1: moderate evidences of error, IF $2 \leq X\_HD5\_STAT \leq 4$ ; 2: severe evidences of error, IF $X\_HD5\_STAT > 4$ ).
X_HD10_STAT	Statistic of the homogeneity test applied on differenced X variable (percentage of data exceeding $\mu + / - 10\sigma$ ).
X_HD10_FLAG	Flag for the homogeneity test applied on differenced X variable (0: negligible evidences of error, IF $X\_HD10\_STAT < 0.5$ ; 1: moderate evidences of error, IF $0.5 \leq X\_HD10\_STAT \leq 1$ ; 2: severe evidences of error, IF $X\_HD10\_STAT > 1$ ).
X_AL1_STAT	Autocorrelation at lag 1 for the X variable.
X_AL1_FLAG	Flag for the autocorrelation at lag 1 for the X variable (0: negligible evidences of error, IF $X\_AL1\_STAT > 0.75$ ; 1: moderate evidences of error, IF $0.5 < X\_AL1\_STAT \leq 0.75$ ; 2: severe evidences of error, IF $X\_AL1\_STAT \leq 0.5$ ).
X_DDI_STAT	Maximum number of binned data points for the X variable (number of classes (bins) calculated by Freedman-Diaconis rule based on the inte-quartile range proportional to the sample size).
X_DDI_FLAG	Flag for the X_DDI_STAT (0: negligible evidences of error, IF $X\_DDI\_STAT < hz * 60 * 2.5$ ; 1: moderate evidences of error, IF $hz * 60 * 2.5 \leq X\_DDI\_STAT < hz * 60 * 5$ ; 2: severe evidences of error, IF $X\_DDI\_STAT \geq hz * 60 * 5$ , where hz denoting the acquisition frequency).
X_DIP_STAT	P-value of the Hartigans' Dip Test for Unimodality.
X_DDI_FLAG	Flag for the X_DIP_STAT (0: negligible evidences of error, IF $X\_DIP\_STAT > 0.1$ ; 1: moderate evidences of error, IF $0.05 \leq X\_DDI\_STAT < 0.1$ ; 2: severe evidences of error, IF $X\_DIP\_STAT \geq 0.05$ ).
X_KID_STAT	Kurtosis Index of Differenced X variable.
X_KID_FLAG	Flag for the X_KID_STAT (0: negligible evidences of error, IF $X\_KID\_STAT < 30$ ; 1: moderate evidences of error, IF $30 \leq X\_KID\_STAT \leq 50$ ; 2: severe evidences of error, IF $X\_KID\_STAT > 50$ ).
ITC_STAT	Statistic of the Integral Turbulence Characteristics test (Foken and Wichura, 1996).
ITC_FLAG	Flag for the ITC test (0: negligible evidences of error, IF $ITC\_STAT < 30$ ; 1: moderate evidences of error, IF $30 \leq ITC\_STAT \leq 50$ ; 2: severe evidences of error, IF $ITC\_STAT > 50$ ).
X_SCF_STAT	Spectral correction factor for the X variable.
X_M98_STAT	Statistic of the nonstationarity ratio test by Mahrt (1998) for the X flux variable.
X_M98_FLAG	Flag of the X_M98_STAT (0: negligible evidences of error, IF $x\_M\_98\_STAT < 2$ ; 1: moderate evidences of error, IF $2 \leq X\_M98\_STAT \leq 3$ ; 2: severe evidences of error, IF $X\_M98\_STAT > 3$ ).
CO2	Carbon Dioxide (CO2) in mole fraction of wet air.
CO2_SIGMA	Standard deviation of carbon dioxide in mole fraction of wet air.
H2O	Water (H2O) vapor mole fraction.
H2O_SIGMA	Standard deviation of water vapor mole fraction.

T_SONIC	Sonic temperature.
T_SONIC_SIGMA	Standard deviation of sonic temperature.
WS	Wind speed.
USTAR	Friction velocity.
W_SIGMA	Standard deviation of vertical velocity fluctuations.
U_SIGMA	Standard deviation of lateral velocity fluctuations (towards main-wind direction after coordinates rotation).
V_SIGMA	Standard deviation of lateral velocity fluctuations (cross main-wind direction after coordinates rotation).
ZL	Monin-Obukhov stability parameter.
MO_LENGTH	Monin-Obukhov length.
AT	Air temperature.
AP	Air pressure.
RHO	Air density.
CP	Air heat capacity.
CANOPY_HEIGHT	Canopy height.
SA_HEIGHT	Sonic anemometer height.
SA_NORTH_OFFSET	Sonic anemometer north offset.
INVALID_WIND_SECTOR_c1	Center of the first invalid wind sector, if any.
INVALID_WIND_SECTOR_w1	Width of the first invalid wind sector, if any.
INVALID_WIND_SECTOR_c2	Center of the second invalid wind sector, if any.
INVALID_WIND_SECTOR_w2	Width of the second invalid wind sector, if any.
INVALID_WIND_SECTOR_c3	Center of the third invalid wind sector, if any.
INVALID_WIND_SECTOR_w3	Width of the third invalid wind sector, if any.

### Author(s)

Domenico Vitale

### References

Vitale, D. Fratini, G. Bilancia, M. Nicolini, G. Sabbatini, S. Papale, D. A robust data cleaning procedure for eddy covariance flux measurements, Biogeosciences, 17, 1367–1391, doi: <https://doi.org/10.5194/bg-17-1367-2020>, 2020.

## Examples

```
PATH_WORKSET <- system.file("extdata", "ecworkset_example.csv", package = "RFlux")
PATH_ECMD <- system.file("extdata", "DE-HoH_ecmd.csv", package = "RFlux")

cleanset <- cleanFlux(path_workset=PATH_WORKSET,
                      path_ecmd=PATH_ECMD,
                      path_output=NULL,
                      FileName=NULL,
                      plotQC=FALSE,
                      storage=TRUE)

str(cleanset)
```

---

closed_path_rawdata	<i>Raw, high-frequency, data for closed path systems</i>
---------------------	--

---

## Description

An example of raw, high-frequency, eddy covariance data for a closed path systems.

## Usage

```
data("closed_path_rawdata")
```

## Format

A data frame with 36000 observations of

U Horizontal wind component along the Gill HS-50 sonic anemometer x axis (m/s).  
V Horizontal wind component along the Gill HS-50 sonic anemometer y axis (m/s).  
W Vertical wind component along the Gill HS-50 sonic anemometer z axis (m/s).  
T\_SONIC Temperature measured by the Gill HS-50 sonic anemometer (kelvin).  
CO2 Carbon dioxide atmospheric concentrations (mixing ratio) measured by the LI-7200 gas analyzer (ppm).  
H2O Water vapor atmospheric concentrations (mixing ratio) measured by the LI-7200 gas analyzer (ppt).  
SA\_DIAG Diagnostic flag output by the Gill HS-50 sonic anemometer (dimensionless).  
GA\_DIAG Diagnostic flag output by the LI-7200 gas analyzer (dimensionless).  
T\_CELL Average cell temperature of the LI-7200 gas analyzer (celsius).  
T\_CELL\_IN Temperature at the inlet of the LI-7200 gas analyzer cell (celsius).  
T\_CELL\_OUT Temperature at the outlet of the LI-7200 gas analyzer cell (celsius).  
PRESS\_CELL Cell pressure of the LI-7200 gas analyzer (Kpa).

## Source

Data are from the Integrated Carbon Observation System (ICOS) European Research Infrastructure and accessible through the ICOS Carbon Portal <http://www.icos-cp.eu>.



**Examples**

```
data(closed_path_rawdata)
str(closed_path_rawdata)
```

convert\_rawdata

*Raw Data File Conversion***Description**

Converting a generic ASCII raw data file in a format suitable for RFlux package

**Usage**

```
convert_rawdata(
  file_path_in,
  file_path_out,
  info_U=list(NULL,NULL,1,0),
  info_V=list(NULL,NULL,1,0),
  info_W=list(NULL,NULL,1,0),
  info_T_SONIC=list(NULL,NULL,1,0),
  info_CO2=list(NULL,NULL,1,0),
  info_H2O=list(NULL,NULL,1,0),
  info_T_CELL=list(NULL,NULL,1,0),
  info_T_CELL_IN=list(NULL,NULL,1,0),
  info_T_CELL_OUT=list(NULL,NULL,1,0),
  info_PRESS_CELL=list(NULL,NULL,1,0),
  info_SA_DIAG=NULL,
  info_GA_DIAG=NULL,
  na.strings,
  nrow.header,
  sep,
  dec,
  timestamp_loc=c(NULL, NULL),
  timestamp_format,
  siteID="Cc-xxx",
  labels=NULL)
```

**Arguments**

file_path_in	path where raw data file is stored.
file_path_out	path where converted raw data file will be stored.
info_U	a list of 4 elements denoting the column position (integer) of the horizontal U wind component, its units (character string, see details), the multiplier and the offset (1 and 0, respectively, in case no linear conversion should be applied).
info_V	a list of 4 elements denoting the column position (integer) of the horizontal V wind component, its units (character string, see details), the multiplier and the offset (1 and 0, respectively, in case no linear conversion should be applied).
info_W	a list of 4 elements denoting the column position (integer) of the vertical W wind component, its units (character string, see details), the multiplier and the offset (1 and 0, respectively, in case no linear conversion should be applied).

info_T_SONIC	a list of 4 elements denoting the column position (integer) of the sonic temperature, its units (character string, see details), the multiplier and the offset (1 and 0, respectively, in case no linear conversion should be applied).
info_CO2	a list of 4 elements denoting the column position (integer) of CO2 concentration, its units (character string, see details), the multiplier and the offset (1 and 0, respectively, in case no linear conversion should be applied).
info_H2O	a list of 4 elements denoting the column position (integer) of the H2O concentration, its units (character string, see details), the multiplier and the offset (1 and 0, respectively, in case no linear conversion should be applied).
info_T_CELL	a list of 4 elements denoting the column position (integer) of cell temperature (only for EC closed path systems), its units (character string, see details), the multiplier and the offset (1 and 0, respectively, in case no linear conversion should be applied).
info_T_CELL_IN	a list of 4 elements denoting the column position (integer) of in-cell temperature (only for EC closed path systems), its units (character string, see details), the multiplier and the offset (1 and 0, respectively, in case no linear conversion should be applied).
info_T_CELL_OUT	a list of 4 elements denoting the column position (integer) of out-cell temperature (only for EC closed path systems), its units (character string, see details), the multiplier and the offset (1 and 0, respectively, in case no linear conversion should be applied).
info_PRESS_CELL	a list of 4 elements denoting the column position (integer) of cell pressure (only for EC closed path systems), its units (character string, see details), the multiplier and the offset (1 and 0, respectively, in case no linear conversion should be applied).
info_SA_DIAG	integer. Column position of the sonic anemometer diagnostics.
info_GA_DIAG	integer. Column position of the gas analyzer diagnostics.
na.strings	a character string denoting missing values in the raw data file.
nrow.header	the number of header rows in the raw data file.
sep	the field separator character of the raw data file. See <a href="#">read.table</a> for specifications.
dec	the character used in the raw data file for decimal points.
timestamp_loc	a vector of 2 negative integers denoting the location of the timestamp in the raw data file name starting from the right (eg. -21, -2).
timestamp_format	a character string denoting the format of the timestamp included in the raw data file name. It must be specified according to specifications of the <a href="#">strptime</a> function.)
siteID	a 6-character string denoting the Site ID (eg. CC-Xxx).
labels	(optional) a character string denoting further information. Default is NULL.

## Details

Possible units specification in the info\_XX field are one among: mm s-1, cm s-1, m s-1 for U, V, W wind components; K, C for sonic temperature (T\_SONIC) and cell temperatures (T\_CELL, T\_CELL\_IN, T\_CELL\_OUT); ppt, ppm, ppb, umol m-3, mmol m-3, ug m-3, mg m-3, g m-3 for

CO<sub>2</sub>, H<sub>2</sub>O concentration variables; Pa, hPa, KPa for cell pressure (PRESS\_CELL). The multiplier and the offset values in info\_XX field aim at converting data expressed in Volt in their physical units. They must be set equals to 1 and 0, respectively, in case no (linear) conversion is required.

### Value

A .csv file stored in the file\_path\_out containing the following variables: U (m s<sup>-1</sup>), V (m s<sup>-1</sup>), W (m s<sup>-1</sup>), T\_SONIC (K), CO<sub>2</sub> (mmol m<sup>-3</sup> in case of molar density, umol/mol (ppm) in case of molar fraction/mixing ratio), H<sub>2</sub>O (mmol m<sup>-3</sup> in case of molar density, mmol/mol (ppt) in case of molar fraction/mixing ratio), SA\_DIAG (dimensionless), GA\_DIAG (dimensionless), T\_CELL (C), T\_CELL\_IN (C), T\_CELL\_OUT (C), PRESS\_CELL (KPa). Missing values are denoted by "-9999".

### Author(s)

Domenico Vitale

### Examples

```
##Single raw data file conversion (TO DO)
##Block of raw data file conversion in parallel mode (TO DO)
```

---

despiking	<i>Despiking algorithms</i>
-----------	-----------------------------

---

### Description

A suite of despiking algorithms for micro-meteorological variables

### Usage

```
despiking(x, mfreq, variant, wsignal, wscale, wby = 1, zth = 5, alpha = 0.01)
```

### Arguments

x	a vector of the observed time-series values.
mfreq	the main frequency of the observed time series (24 or 48 for hourly and half-hourly time series, respectively; 10 or 20 for raw EC data acquired at 10Hz or 20Hz, respectively).
variant	"v1" mainly designed for (half-)hourly EC fluxes (missing values are allowed); "v2" mainly designed for (half-)hourly meteo time series (a low percentage of missing values is allowed); "v3" mainly designed for high-frequency EC raw data.
wsignal	the window width used to estimate the underlying signal (only for met and raw variants).
wscale	the window width used to estimate the local scale parameter (only for met and raw variants).
wby	calculate the scale parameter at every by-th point rather than every point (default=1). Large values of wby reduce the computational time, but can introduce bias in the scale parameter estimation.
zth	the threshold value of the z-sigma rule ((only for met and raw variants).
alpha	the significance level used in the outlier detection rule (only for flx variant).

**Value**

A list containing the following components:

ts\_cleaned      A vector of the despiked time-series.  
 spike\_loc      Integer. Location of the detected spikes.

**Author(s)**

Domenico Vitale

**References**

Fried, R. Schettlinger, K. Matthias Borowski, M. robfilter: Robust Time Series Filters. R package version 4.1.1. <https://CRAN.R-project.org/package=robfilter>, 2018.

Hafen, R. stlplus: Enhanced Seasonal Decomposition of Time Series by Loess. R package version 0.5.1. <https://CRAN.R-project.org/package=stlplus>, 2016.

Vitale, D. Fratini, G. Bilancia, M. Nicolini, G. Sabbatini, S. Papale, D. A robust data cleaning procedure for eddy covariance flux measurements, Biogeosciences, 17, 1367–1391, doi: <https://doi.org/10.5194/bg-17-1367-2020>, 2020.

Vitale, D. A performance evaluation of despiking algorithms for eddy covariance data. Scientific Reports, 11, 11628, doi: <https://doi.org/10.1038/s41598-021-91002-y>, 2021.

---

ecmd_table	<i>Eddy covariance metadata table</i>
------------	---------------------------------------

---

**Description**

A collection of metadata useful for the proper settings of eddy covariance data processing options.

**Usage**

```
data("ecmd_table")
```

**Format**

A data frame with the following variables:

DATE\_OF\_VARIATION\_DB Optional. ISO timestamp of the variation in central database. Required format: *yyyymmddHHMM*.

DATE\_OF\_VARIATION\_EF ISO timestamp of effective date of variation. Required format: *yyyymmdd* or *yyyymmddHHMM*.

SITEID Character string indicative of the EC Site's ID (*CC-Xxx*).

LATITUDE Latitude in decimal degree

LONGITUDE Longitude in decimal degree

ALTITUDE Altitude in meters

CANOPY\_HEIGHT Canopy height in meters

SA\_MANUFACTURER Sonic anemometer manufacturer (see EddyPro manual.)

SA\_MODEL Sonic anemometer model (see EddyPro manual.)

SA\_SW\_VERSION Sonic anemometer software version.  
 SA\_WIND\_DATA\_FORMAT Format of the wind components (see EddyPro manual.)  
 SA\_NORTH\_ALIGNMENT Sonic anemometer north alignment (see EddyPro manual.)  
 SA\_HEIGHT Sonic anemometer height (see EddyPro manual.)  
 SA\_NORTH\_OFFSET Sonic anemometer north offset (see EddyPro manual.)  
 SA\_NORTH\_MAGDEC Sonic anemometer magnetic declination (see EddyPro manual.)  
 SA\_INVALID\_WIND\_SECTOR\_c1 Center of the first invalid wind sector.  
 SA\_INVALID\_WIND\_SECTOR\_w1 Width of the first invalid wind sector.  
 SA\_INVALID\_WIND\_SECTOR\_c2 Center of the second invalid wind sector.  
 SA\_INVALID\_WIND\_SECTOR\_w2 Width of the second invalid wind sector.  
 SA\_INVALID\_WIND\_SECTOR\_c3 Center of the third invalid wind sector.  
 SA\_INVALID\_WIND\_SECTOR\_w3 Width of the third invalid wind sector.  
 GA\_PATH Specify the eddy covariance path system: closed or open.  
 GA\_MANUFACTURER Gas analyzer manufacturer (see EddyPro manual).  
 GA\_MODEL Gas analyzer model (see EddyPro manual).  
 GA\_SW\_VERSION Gas analyzer software version (see EddyPro manual).  
 GA\_NORTHWARD\_SEPARATION The distance between the center of the sample volume (or the inlet of the intake tube) of the current gas analyzer and the sonic anemometer, as measured horizontally along the north-south axis (see EddyPro manual.).  
 GA\_EASTWARD\_SEPARATION The distance between the center of the sample volume (or the inlet of the intake tube) of the current gas analyzer and the sonic anemometer, as measured horizontally along the east-west axis (see EddyPro manual.).  
 GA\_VERTICAL\_SEPARATION The distance between the center of the sample volume (or the inlet of the intake tube) of the current gas analyzer and the sonic anemometer, as measured vertically (see EddyPro manual).  
 GA\_TUBE\_DIAMETER The inside diameter of the intake tube in centimeters. Mandatory for closed path system.  
 GA\_FLOWRATE The flow rate in the intake tube. Mandatory for closed path system.  
 GA\_TUBE\_LENGTH The length of the intake tube in centimeters.  
 FILE\_DURATION File length duration in minutes  
 ACQUISITION\_FREQUENCY Number of sample records per second.  
 FILE\_FORMAT ASCII (fixed)  
 FILE\_EXTENSION csv (fixed)  
 LN 99 (fixed)  
 FN 1 (fixed)  
 EXTERNAL\_TIMESTAMP END (fixed)  
 EOL crlf (fixed)  
 SEPARATOR comma (fixed)  
 MISSING\_DATA\_STRING -9999 (fixed)  
 NROW\_HEADER 1 (fixed)  
 UVW\_UNITS Units of wind speed components (see EddyPro manual).  
 T\_SONIC\_UNITS Units of sonic temperature (see EddyPro manual).

T\_CELL\_UNITS Units of closed path gas analyzer cell temperature (see EddyPro manual).

P\_CELL\_UNITS Units of closed path gas analyzer cell pressure (see EddyPro manual).

CO2\_measure\_type Carbon dioxide concentration measurement type (e.g. mixing\_ratio, see EddyPro manual).

CO2\_UNITS Units of carbon dioxide concentration (e.g. ppm, see EddyPro manual).

H2O\_measure\_type Water vapor concentration measurement type (e.g. mixing\_ratio, see EddyPro manual).

H2O\_UNITS Units of water vapor concentration (e.g. ppt, see EddyPro manual).

SA\_DIAG Units of the sonic anemometer diagnostics (dimensionless, fixed)

GA\_DIAG Units of the gas analyzer diagnostics (dimensionless (fixed)).

### Source

Metadata elaborated by ICOS-ETC team (Integrated Carbon Observation System (ICOS-ETC) European Research Infrastructure - Ecosystem Thematic Center, <http://www.icos-etc.eu/icos/>).

### References

LI-COR Biosciences: EddyPro 7.0.4: Help and User's Guide, LI-COR Biosciences, Lincoln, Nebraska USA, [www.licor.com/EddyPro](http://www.licor.com/EddyPro), 2019.

### Examples

```
data(ecmd_table)
str(ecmd_table)
```

---

ecworkset	<i>Merge time series and returns the workset dataframe</i>
-----------	--

---

### Description

Merge time series with common indexes (times) and returns the *workset* data frame to be used as input for the data cleaning procedure via [cleanFlux](#) function.

### Usage

```
ecworkset(path_EPout, path_EPqc, path_EPmd, path_QCstat, path_output=NULL, FileName=NULL)
```

### Arguments

path_EPout	path where the fulloutput file generated by EddyPro software is stored.
path_EPqc	path where the qcdetails file generated by EddyPro software is stored.
path_EPmd	path where the metadata file generated by EddyPro software is stored.
path_QCstat	path where the QC statistics file generated by <a href="#">qcStat</a> is stored.
path_output	path where the results will be stored. Default is NULL.
FileName	a character string naming a file for writing. Default is NULL.

## Details

Returns a dataframe ordered by common indexes (times) containing a set of variables selected from the fulloutput, metadata, qcdetails files generated by EddyPro via `eddypro_run` and the statistics of the quality control routines (Vitale et al, 2019) generated by `qcStat`.

## Value

A dataframe object to be used as input of the `cleanFlux`. For the meaning of variables and units see the EddyPro manual and the `qcStat` function description.

## Author(s)

Domenico Vitale, Dario Papale

## References

LI-COR Biosciences: EddyPro 7.0.4: Help and User's Guide, LI-COR Biosciences, Lincoln, Nebraska USA, [www.licor.com/EddyPro](http://www.licor.com/EddyPro), 2019.

Vitale, D. Fratini, G. Bilancia, M. Nicolini, G. Sabbatini, S. Papale, D. A robust data cleaning procedure for eddy covariance flux measurements, Biogeosciences, 17, 1367–1391, doi: <https://doi.org/10.5194/bg-17-1367-2020>, 2020.

## Examples

```
PATH_EPOUT <- system.file("extdata", "eddypro_DE-HoH_full_output_example.csv", package = "RFlux")
PATH_EPQC <- system.file("extdata", "eddypro_DE-HoH_qc_details_example.csv", package = "RFlux")
PATH_EPMO <- system.file("extdata", "eddypro_DE-HoH_metadata_example.csv", package = "RFlux")
PATH_QCSTAT <- system.file("extdata", "qcStat_example.csv", package = "RFlux")

WorkSet <- ecworkset(path_EPout=PATH_EPOUT,
                    path_EPqc=PATH_EPQC,
                    path_EPMd=PATH_EPMO,
                    path_QCstat=PATH_QCSTAT,
                    path_output=NULL,FileName=NULL)

str(WorkSet)
```

---

eddypro\_run

*Fluxes estimation*


---

## Description

Estimates flux values and other micrometeorological parameters through a call to LI-COR EddyPro software.

## Usage

```
eddypro_run(siteID, path_eddypro_bin, path_eddypro_projfiles, showLOG = TRUE)
```

## Arguments

siteID	Character string indicative of the site's ID (i.e. CC-Xxx)
path_eddypro_bin	path where eddypro_rp and eddypro_fcc executables (i.e. the bin folder) are stored.
path_eddypro_projfiles	path where the EddyPro project files (i.e. the <i>filename.eddypro</i> file) is stored.
showLOG	logical. If TRUE (default), the EddyPro Output Console Page is shown on R console.

## Value

A set of .csv files (e.g. fulloutput, qcdetails and metadata) generated by EddyPro software will be stored in the path specified by `get_md`.

## Warning

It is required LI-COR EddyPro software is currently installed on your system (see [www.licor.com/EddyPro](http://www.licor.com/EddyPro)).

## Author(s)

Domenico Vitale, Dario Papale

## References

LI-COR Biosciences: EddyPro 7.0.4: Help and User's Guide, LI-COR Biosciences, Lincoln, Nebraska USA, [www.licor.com/EddyPro](http://www.licor.com/EddyPro), 2019.

## Examples

```
## Not run:
workdir <- getwd()
siteID <- "DE-HoH"
PATH_ECMD <- system.file("extdata", "DE-HoH_ecmd.csv", package = "RFlux")
PATH_RAWDATA <- system.file("extdata", package = "RFlux")
PATH_OUTPUT <- paste0(workdir, "/eddypro/processing")
PATH_SA_FILE <- system.file("extdata", "spectral_assessment.txt", package = "RFlux")
PATH_PF_FILE <- system.file("extdata", "planar_fit.txt", package = "RFlux")

MD <- get_md(path_ecmd=PATH_ECMD,
             path_rawdata=PATH_RAWDATA,
             path_output=PATH_OUTPUT,
             online=TRUE,
             path_sa_file=PATH_SA_FILE,
             path_pf_file=PATH_PF_FILE,
             tlag_meth=2,
             despikes_meth="VM97",
             detrend_meth="BA",
             tilt_correction_meth="PF")

PATH_EDDYPRO_BIN <- "/Applications/eddypro.app/Contents/MacOS/bin" ## put your path!
PATH_EDDYPRO_PROJ <- PATH_OUTPUT
eddypro_run(siteID="DE-HoH",
```



```

    path_eddypro_bin=PATH_EDDYPRO_BIN,
    path_eddypro_projfiles=PATH_EDDYPRO_PROJ,
    showLOG=TRUE)

## End(Not run)

```

get\_md

*Metadata file management*

## Description

Returns the input files required by EddyPro software (LI-COR registered trademark; Fratini and Mauder, 2014).

## Usage

```

get_md(path_ecmd = NULL,
       path_rawdata = NULL,
       path_output = NULL,
       online,
       path_sa_file = NULL,
       path_pf_file = NULL,
       path_biomet_file=NULL,
       roughness_length=NULL,
       displacement_height=NULL,
       tlag_meth,
       despikes_meth=c("None", "VM97", "M13"),
       detrend_meth=c("BA", "LD"),
       tilt_correction_meth=c("DR", "PF"))

```

## Arguments

path_ecmd	path where the CC-Xxx_ecmd.csv file containing metadata information is stored.
path_rawdata	path where eddy covariance rawdata files are stored.
path_output	path where the results of <a href="#">eddypro_run</a> will be stored (it will be created, if any).
online	logical. Indicating whether the parameters of the planar fit method and the spectral correction factors are taken from results of previous processing (online=TRUE) or estimated by using the current set of EC rawdata (online=FALSE).
path_sa_file	path where the spectral assessment file generated by EddyPro software is stored.
path_pf_file	path where the planar fit file generated by EddyPro software is stored.
path_biomet_file	path where the file containing the biometeorological variables is stored (the file format must follow the EddyPro instructions).
roughness_length	a numeric value indicating the surface roughness. If NULL (default) it is estimated based on the canopy height.
displacement_height	a numeric value indicating the zero plane displacement length. If NULL (default) it is estimated based on the canopy height.

tlag_meth	time lag detection method: 0 None; 1 Constant time lag, 2 Maximum covariance with default; 3 Maximum covariance; 4 Automatic optimization.
despike_meth	despiking method: "None", not apply; "VM97", algorithm by Vickers and Mahrt (1997); "M13", algorithm by Mauder et al (2013).
detrend_meth	trend removal method: "BA" Block Average; "LD" Linear Detrending (Rannik and Vesala, 2001).
tilt_correction_meth	axis rotation method for tilt correction: "DR" Double Rotation; "PF" Planar Fit (Wilczak et al, 2001).

## Details

This function builds the input files required by EddyPro software to process eddy covariance raw-data files.

## Value

Returns the following files

CC-Xxx.eddypro

CC-Xxx.metadata

CC-Xxx\_dynamic\_metadata.txt

where CC-Xxx denoting the site's ID.

## Warning

If online=FALSE, the number of rawdata needs to be large enough to allow robust estimates of the planar fit parameters (Wilczak et al 2001) and of the spectral correction factors (Fratini et al, 2012). In case of few rawdata, the double rotation method for sonic anemometer tilt correction and the analytical method by Moncrieff et al (1997) for spectral correction will be performed.

## Note

Rawdata must be provided as .csv file with the following name: *CC-Xxx????yyyymmddHHMM????.csv*, where *CC-Xxx* is the site's ID, *yyyymmddHHMM* is the ISO timestamp, and ?s denoting free characters. For example: DE-HoH\_EC\_201901010030\_v01.csv.

## Author(s)

Domenico Vitale, Dario Papale

## References

- Fratini, G., Mauder, M. (2014). Towards a consistent eddy-covariance processing: an intercomparison of EddyPro and TK3. *Atmospheric Measurement Techniques*, 7(7), 2273-2281, doi: <https://doi.org/10.5194/amt-7-2273-2014>.
- Fratini, G., Ibrom, A., Arriga, N., Burba, G., Papale, D. (2012). Relative humidity effects on water vapour fluxes measured with closed-path eddy-covariance systems with short sampling lines. *Agricultural and forest meteorology*, 165, pp 53-63, doi: <https://doi.org/10.1016/j.agrformet.2012.05.018>.

LI-COR Biosciences: EddyPro 7.0.4: Help and User's Guide, LI-COR Biosciences, Lincoln, Nebraska USA, [www.licor.com/EddyPro](http://www.licor.com/EddyPro), 2019.

Mauder, M., Cuntz, M., Drue, C., Graf, A., Rebmann, C., Schmid, H. P., ..., Steinbrecher, R. (2013). A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements. *Agricultural and Forest Meteorology*, 169, 122-135, doi: <https://doi.org/10.1016/j.agrformet.2012.09.006>.

Moncrieff, J. B., Massheder, J. M., De Bruin, H., Elbers, J., Friborg, T., Heusinkveld, B., ..., Verhoef, A. (1997). A system to measure surface fluxes of momentum, sensible heat, water vapour and carbon dioxide. *Journal of Hydrology*, 188, 589-611, doi: [https://doi.org/10.1016/S0022-1694\(96\)03194-0](https://doi.org/10.1016/S0022-1694(96)03194-0).

Rannik, U., Vesala, T. (1999). Autoregressive filtering versus linear detrending in estimation of fluxes by the eddy covariance method. *Boundary-Layer Meteorology*, 91(2), 259-280, doi: <https://doi.org/10.1023/A:1001840416858>.

Vickers, D., Mahrt, L. (1997). Quality control and flux sampling problems for tower and aircraft data. *Journal of atmospheric and oceanic technology*, 14(3), 512-526, doi: [https://doi.org/10.1175/1520-0426\(1997\)014<0512:QCAFSP>2.0.CO;2](https://doi.org/10.1175/1520-0426(1997)014<0512:QCAFSP>2.0.CO;2).

Wilczak, J.M., Oncley, S.P., Stage, S.A. (2001). Sonic anemometer tilt correction algorithms. *Boundary-Layer Meteorology*, 99(1), 127-150, doi: <https://doi.org/10.1023/A:1018966204465>.

## See Also

[closed\\_path\\_rawdata](#)

## Examples

```
## Not run:
workdir <- getwd()
siteID <- "DE-HoH"
PATH_ECMD <- system.file("extdata", "DE-HoH_ecmd.csv", package = "RFlux")
PATH_RAWDATA <- system.file("extdata", package = "RFlux")
PATH_OUTPUT <- paste0(workdir, "/eddypro/processing")
PATH_SA_FILE <- system.file("extdata", "spectral_assessment.txt", package = "RFlux")
PATH_PF_FILE <- system.file("extdata", "planar_fit.txt", package = "RFlux")

MD <- get_md(path_ecmd=PATH_ECMD,
             path_rawdata=PATH_RAWDATA,
             path_output=PATH_OUTPUT,
             online=TRUE,
             path_sa_file=PATH_SA_FILE,
             path_pf_file=PATH_PF_FILE,
             path_biomet_file=NULL,
             roughness_length=NULL,
             displacement_height=NULL,
             tlag_meth=2,
             despikes_meth="VM97",
             detrend_meth="BA",
             tilt_correction_meth="PF")

## End(Not run)
```

---

inst_prob_test	<i>Detection of instrumental malfunctions, errors in eddy covariance systems</i>
----------------	--

---

### Description

A set of tests aims at detecting instrumental malfunctions, errors in eddy covariance systems.

### Usage

```
inst_prob_test(x,
               detrend=FALSE,
               hz=c(10,20),
               plot=FALSE,
               var_name=c("U", "V", "W", "T_SONIC", "CO2", "H2O"),
               cex.test=1)
```

### Arguments

x	raw high frequency eddy covariance time series.
detrend	logical. If TRUE a linear trend is removed from data (see details).
hz	integer. The data acquisition frequency (ie 10 or 20 HZ).
plot	logical. If TRUE a graphical representation of the results is provided. Default FALSE.
var_name	character. One of "U","V","W","T_SONIC","CO2","H2O" variable name can be specified to enrich the graphcial output when plot=TRUE.
cex.test	A numerical value giving the amount by which plotting test results. Default set to 1.

### Details

The detrend option should be set in agreement with the processing options adopted for fluxes estimation. In case fluxes are estimated after mean removal (i.e. block average), then detrend should be set to FALSE. In case fluxes are estimated after linear trend removal then detrend should be set to TRUE.

### Value

Skew	Skewness.
Kurt	Kurtosis index on original data.
KID0	Kurtosis index on difference data.
KID	Kurtosis index on difference data excluding low resolution problems.
HFx	Homogeneity test statistic based on fluctuation.
HDx	Homogeneity test statistic based on difference data.
AL1	Autocorrelation function estimate at lag 1.
DDI	Data Distribution Integrity test statistic. Maximum number of data points in a bin of the histogram with bins defined according to Freedman-Diaconis rule.
DIP	P-value of the Hartigans' Dip test for unimodality.

**Author(s)**

Domenico Vitale

**References**

Vitale, D. Fratini, G. Bilancia, M. Nicolini, G. Sabbatini, S. Papale, D. A robust data cleaning procedure for eddy covariance flux measurements, *Biogeosciences*, 17, 1367–1391, doi: <https://doi.org/10.5194/bg-17-1367-2020>, 2020.

Freedman, D. and Diaconis, P. On the histogram as a density estimator: L2 theory. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 57, 453–476, doi: <https://doi.org/10.1007/BF01025868>, 1981.

Hartigan, P.M. Computation of the Dip Statistic to Test for Unimodality. *Applied Statistics (JRSS C)* 34, 320–325. Corresponding Fortran code of AS 217 available from Statlib, <http://lib.stat.cmu.edu/apstat/217>, 1985.

Hartigan, J.A. and Hartigan, P.M. The Dip Test of Unimodality. *Annals of Statistics*, 13, 70–84, 1985.

**Examples**

```
data(closed_path_rawdata)
inst_prob_test(closed_path_rawdata$W, detrend=FALSE, h=20, plot=TRUE, var_name="W")
```

mahrt

*Stationary test for eddy covariance fluxes***Description**

Performs the non-stationary ratio test described by Mahrt (1998)

**Usage**

```
mahrt(x)
```

**Arguments**

**x** two-column dataframe containing raw high-frequency time series of vertical wind component (i.e. W) and scalar atmospheric variable (e.g. CO<sub>2</sub>).

**Value**

Returns the non-stationary ratio test statistic.

**Author(s)**

Domenico Vitale

**References**

Mahrt L (1998) Flux sampling errors for aircraft and towers, *J. Atmos. Ocean. Tech.*, 15, 416–429, [https://doi.org/10.1175/1520-0426\(1998\)015<0416:fsefaa>2.0.co;2](https://doi.org/10.1175/1520-0426(1998)015<0416:fsefaa>2.0.co;2).

**Examples**

```
PATH_RAWDATA <- system.file("extdata", "DE-HoH_EC_201907301200_v01.csv", package = "RFlux")
data(closed_path_rawdata)
mahrt(data.frame("W"=closed_path_rawdata$W, "CO2"=closed_path_rawdata$CO2))
```

---

mode	<i>Mode estimation</i>
------	------------------------

---

**Description**

Calculate the modal value

**Usage**

```
mode(x)
```

**Arguments**

x                      a vector of data.

**Value**

Returns the mode.

**Author(s)**

Domenico Vitale

---

qcStat	<i>Quality control tests for eddy covariance fluxes</i>
--------	---

---

**Description**

Returns the test statistics of the quality control routines described by Vitale et al (2019).

**Usage**

```
qcStat(path_rawdata, ext_tstamp=c("START", "END"), path_output=NULL, FileName=NULL)
```

**Arguments**

path_rawdata	path where raw high-frequency eddy covariance data are stored.
ext_tstamp	Character string specifying whether the timestamp in the file name is indicative of the beginning ("START") or the end ("END") of time series involved.
path_output	path where the results will be stored. Default is NULL.
FileName	a character string naming a file for writing. Default is NULL.

**Value**

A data frame containing:

TSTAMP	ISO timestamp start of averaging period (format: <i>yyyymmddHHMM</i> ).
SADiag	Diagnostic for the sonic anemometer.
FMR_X	Fraction of Missing Records in raw, high-frequency, data used for X flux variable estimation.
LGD_X	Longest Gap Duration in raw, high-frequency, data used for X flux variable estimation.
Skew_X	Skewness of X variable.
Kurt_X	Kurtosis of X variable.
KID0_X	Kurtosis Index of Differenced X variable without esclusion of zero values.
KID_X	Kurtosis Index of Differenced X variable excluding zeros, ie the effect of possible low resolution problems.
HF5_X	Statistic of the homogeneity test applied on X variable fluctuations (percentage of data exceeding $+/- 5\sigma$ ).
HF10_X	Statistic of the homogeneity test applied on X variable fluctuations (percentage of data exceeding $+/- 10\sigma$ ).
HD5_X	Statistic of the homogeneity test applied on differenced X variable (percentage of data exceeding $+/- 5\sigma$ ).
HD10_X	Statistic of the homogeneity test applied on differenced X variable (percentage of data exceeding $+/- 10\sigma$ ).
AL1_X	Autocorrelation at lag 1.
DDI_X	Maximum number of binned data points (number of bins proportional to the sample size).
DIP_X	P-value of the Hartigans' Dip Test for Unimodality.
COV_XY	Covariance between X and Y variables.
N0_X	Number of repeated consecutive values affecting X flux variable.
LSR_X	Statistic of the Low Signal Resolution test for X flux variable.
M98_X	Statistic of the non-stationary ratio test by Mahrt (1988) for X flux variable.

**Note**

The output of [qcStat](#) constiutes one of input files required by the [ecworkset](#) function.

**Author(s)**

Domenico Vitale

**References**

- Mahrt L (1998) Flux sampling errors for aircraft and towers, J. Atmos. Ocean. Tech., 15, 416-429, [https://doi.org/10.1175/1520-0426\(1998\)015<0416:fsefaa>2.0.co;2](https://doi.org/10.1175/1520-0426(1998)015<0416:fsefaa>2.0.co;2).
- Vitale, D. Fratini, G. Bilancia, M. Nicolini, G. Sabbatini, S. Papale, D. A robust data cleaning procedure for eddy covariance flux measurements, Biogeosciences, 17, 1367–1391, doi: <https://doi.org/10.5194/bg-17-1367-2020>, 2020.

**Examples**

```
PATH_RAWDATA <- system.file("extdata", "DE-HoH_EC_201907301200_v01.csv", package = "RFlux")
QC_STAT <- qcStat(path_rawdata=PATH_RAWDATA, ext_tstamp="END", path_output=NULL, FileName=NULL)
QC_STAT
```

---

robust_despiking	<i>Despiking eddy covariance raw data</i>
------------------	---

---

**Description**

A parallel implementation of robust functionals for despiking high-frequency eddy covariance raw data

**Usage**

```
robust_despiking(x, mfreq, file_length)
```

**Arguments**

x	a vector of the observed time-series values.
mfreq	the main frequency of the observed time series (24 or 48 for hourly and half-hourly time series, respectively; 10 or 20 for raw EC data acquired at 10Hz or 20Hz, respectively).
file_length	the raw data file length in minutes (e.g. 30, 60 minutes).

**Value**

A list containing the following components:

despiked_ts	A vector of the despiked time-series.
spike_loc	Integer. Location of the detected spikes.

**Author(s)**

Domenico Vitale

**References**

Fried, R. Schettlinger, K. Matthias Borowski, M. robfilter: Robust Time Series Filters. R package version 4.1.1. <https://CRAN.R-project.org/package=robfilter>, 2018.

Vitale, D. A performance evaluation of despiking algorithms for eddy covariance data. Scientific Reports, 11, 11628, doi: <https://doi.org/10.1038/s41598-021-91002-y>, 2021.



---

tlag_detection	<i>Time-lag detection</i>
----------------	---------------------------

---

## Description

Optimal time-lag detection by means of the pre-whitening procedure.

## Usage

```
tlag_detection(x, y, mfreq, x.model = ar.res,
              AIC = FALSE, LAG.MAX = mfreq*10,
              show.plot = FALSE)
```

## Arguments

x	first component series (e.g. CO2 concentrations).
y	second component series (e.g. vertical wind speed).
mfreq	acquisition frequency.
x.model	an ARIMA model; if provided, it is used to prewhiten both series. Otherwise, an AR model is fitted to the x-series and used to pre-whiten both series.
AIC	Logical. If an AR model is used for prewhitening, the AR order is chosen by minimizing the Akaike Information Criteria (AIC=TRUE), otherwise the AR model is chosen as $10 \cdot \log_{10}(N)$ where N is the number of observations. Default is FALSE.
LAG.MAX	maximum lag at which to calculate the ccf. Default is 10 seconds.
show.plot	Logical. If TRUE, the CCFs will be displayed (default FALSE).

## Value

A list containing the following components:

ccf	Output from the ccf function on the prewhitened data.
ar	The AR model fit to the x-series, or x.model if it is provided.
opt_tlag	Optimal time lag detected (timestep).
tlag_pw	Time lag of the main peak of the cross-correlation function after pre-whitening.
tlag_lmax	Local maxima of the cross-covariance function closest to tlag_pw.
tlag_lmin	Local minima of the cross-covariance function closest to tlag_pw.
corr_est	Estimated correlation at .
cv1pct	Critical (abs) value of the ccf at 0.01 level.
cv5pct	Critical (abs) value of the ccf at 0.05 level.

## Author(s)

Domenico Vitale

## References

Vitale et al. (in prep) Optimal Time Lag Detection for Eddy Covariance Data Acquisition Systems

**See Also**

See also `prewhiten` function in the TSA R package.

**Examples**

```
## Not run:
library(RFlux)
library(forecast)
library(xts)
data("closed_path_rawdata")
rawdata <- closed_path_rawdata
N <- nrow(rawdata)
timestamp_orig <- strptime("0000.00", format="%H%M.%OS", tz="GMT") + seq(1, N, 1);
data.xts <- xts(rawdata, order.by=timestamp_orig);

## Detecting time lag between H2O and vertical wind speed (W) by using an AR model (default)
tlag_h2o_out <- tlag_detection(x=rawdata$H2O, y=rawdata$W, mfreq=20,
                             show.plot=TRUE)

tlag_h2o_out$opt_tlag
tlag_h2o_out$corr_est

## Detecting time lag between CO2 and Sonic Temperature by using an ARIMA model
mod <- auto.arima(rawdata$CO2, stationary=FALSE)
tlag_co2_out <- tlag_detection(x=rawdata$CO2, y=rawdata$T_SONIC, mfreq=20,
                              show.plot=TRUE, x.model=mod)

tlag_co2_out$opt_tlag
tlag_co2_out$corr_est

## Time series alignment
tlag_co2 <- tlag_co2_out$opt_tlag
tlag_h2o <- tlag_h2o_out$opt_tlag
timestamp_co2 <- strptime("0000.00", format="%H%M.%OS", tz="GMT") +
  seq(-tlag_co2+1, N - tlag_co2, 1);
timestamp_h2o <- strptime("0000.00", format="%H%M.%OS", tz="GMT") +
  seq(-tlag_h2o+1, N - tlag_h2o, 1);

co2s <- xts(rawdata$CO2, order.by=timestamp_co2);
h2os <- xts(rawdata$H2O, order.by=timestamp_h2o);
rawdata_aligned <- na.omit(data.frame(merge(data.xts[,c(1:4)],
      co2s, h2os, data.xts[,c(7:12)], tzone="GMT")));
colnames(rawdata_aligned) <- c("U", "V", "W", "T_SONIC", "CO2", "H2O",
  "SA_DIAG", "GA_DIAG",
  "T_CELL", "T_CELL_IN", "T_CELL_OUT", "PRESS_CELL")

## End(Not run)
```

# Index

- \*Topic **datagen**
  - cleanFlux, [3](#)
  - ecworkset, [14](#)
  - qcStat, [22](#)
- \*Topic **datasets**
  - closed\_path\_rawdata, [8](#)
  - convert\_rawdata, [9](#)
  - ecmd\_table, [12](#)
- \*Topic **file**
  - get\_md, [17](#)
- \*Topic **htest**
  - tlag\_detection, [25](#)
- \*Topic **interface**
  - eddypro\_run, [15](#)
- \*Topic **methods**
  - despiking, [11](#)
  - robust\_despiking, [24](#)
  - tlag\_detection, [25](#)
- \*Topic **optimize**
  - robust\_despiking, [24](#)
- \*Topic **package**
  - RFlux-package, [2](#)
- \*Topic **robust**
  - despiking, [11](#)
  - robust\_despiking, [24](#)
- \*Topic **univar**
  - inst\_prob\_test, [20](#)
  - mahrt, [21](#)
  - mode, [22](#)

cleanFlux, [2](#), [3](#), [14](#), [15](#)

closed\_path\_rawdata, [8](#), [19](#)

convert\_rawdata, [9](#)

despiking, [11](#)

ecmd\_table, [12](#)

ecworkset, [2](#), [14](#), [23](#)

eddypro\_run, [15](#), [15](#), [17](#)

get\_md, [2](#), [16](#), [17](#)

inst\_prob\_test, [20](#)

mahrt, [21](#)

mode, [22](#)

qcStat, [2](#), [4](#), [14](#), [15](#), [22](#), [23](#)

read.table, [10](#)

RFlux (RFlux-package), [2](#)

RFlux-package, [2](#)

robust\_despiking, [24](#)

strptime, [10](#)

tlag\_detection, [25](#)