

# Science Demonstrator 2: The Eddy Covariance Fluxes of GHGs (Use Case IC\_13)

Please provide your feedback on this Science Demonstrator using the questionnaire at <https://survey2.icos-cp.eu/ENVRIplus-evaluator!>

## Overview

Eddy covariance (EC) fluxes calculation involves a complex set of data processing steps that, beyond the knowledge of the technique, requires a considerable amount of computational resources. This might constitute a constraint for RIs (e.g. ICOS) that aim to simultaneously process raw data sampled at multiple sites in Near Real Time (NRT) mode (i.e. provide each day fluxes estimates relative to the previous day). This demonstrator is the implementation results of [Use Case IC\\_13](#), and it showcases the service solution that integrates the gCube service to optimize the processing of EC data based on 4 different processing schemes resulting from a combination of block average (ba) or linear detrending (ld) and double rotation (dr) or planar fit (pf) processing options and make this available to other RIs.

## Scientific Objectives

The ambitious goal of this pilot investigation is to provide a computationally efficient tool able to process EC raw data and offer to users the possibility to calculate EC fluxes according to the multiple processing scheme described by Sabbatini et al. (2018)<sup>[1]</sup>. The ultimate aim is to establish a service that can be used by RIs that use this micrometeorological technique to measure exchanges of greenhouse gases and energy between terrestrial ecosystems and atmosphere (e.g. ICOS but also other RIs using the eddy covariance methods such for example LTER or ANAEE).

## Description

The EC technique involves high-frequency sampling (e.g. 10 or 20 Hz) of wind speed and scalar atmospheric concentration data, and yields vertical turbulent fluxes. EC fluxes are computed within a finite averaging time (normally 30 mins) from the covariance estimates between instantaneous deviations in vertical wind speed and gas concentration (e.g. CO<sub>2</sub>) from their respective mean values, multiplied by the mean air density (see Aubinet et al., 2012)<sup>[2]</sup>.

Despite the simplicity of this idea, a number of practical difficulties arise in transforming high-frequency data into reliable half-hourly flux measurements. To cope with these issues, here we used the tools implemented by the EddyPro® Fortran code (LI-COR Biosciences, 2017, Fratini and Mauder, 2014)<sup>[3]</sup> an open source software application available for free download at [https://www.licor.com/env/products/eddy\\_covariance/eddypro.html](https://www.licor.com/env/products/eddy_covariance/eddypro.html). The choice of EddyPro® software is motivated by i) the availability of different methods for data quality control and processing (e.g. coordinate rotation, time series detrending, time lag determination, spectral corrections, flux random uncertainty quantification, etc.), ii) the availability of the source code and iii) the fact that the software is based on a community developed set of tools.

Required for the processing of EC raw data through EddyPro® software, are 1) the availability of metadata information about the EC system setup and raw data file structure, and 2) the choice of a suitable combination of processing options.

Concerning 1), users have to provide a standardized metadata file in .csv format (metadata.csv, see [Table 1](#)). This file constitutes the input of an R script that automatically builds the mandatory files ingested into the EddyPro® software (i.e. the .metadata and .eddypro files) developed ad hoc for this exercise. The organization and name of the metadata variables is based on an international standard (BADM) used also in the USA network AmeriFlux. The format of the csv has been instead designed in order to develop a template easy to prepare by individual scientists and organized RIs. In case of NRT data processing, in order to perform part of the flux corrections (i.e. spectral corrections and planar fit), 5 additional configuration files are needed: planar\_fit.txt, spectral\_assessment\_badr.txt, spectral\_assessment\_lddr.txt, spectral\_assessment\_bapf.txt, spectral\_assessment\_ldpf.txt. They can be obtained by specific EddyPro® runs based on long periods of data (at least one month of data is usually required for a consistent parameters estimation). The above files have to be placed together with EC raw-data files in an archive folder (data.zip) which will constitute the input file of the current implementation (see [Figure 1](#)).

**Table 1. Description of Metadata to provide in the metadata.csv file**

Column	Variable Label (file header)	Description (Units)
1	SITEID	Official EC station code following the FLUXNET standards (CC-Xxx)
2	LATITUDE	Geographic latitude ([-90,90] from S to N, decimal)
3	LONGITUDE	Geographic longitude ([-180,180] from W to E, decimal)
4	ALTITUDE	Altitude of ecosystem under study (m)
5	CANOPY_HEIGHT	Distance between the ground and the top of the plant canopy (m)
6	SA_MANUFACTURER	Manufacturer of the sonic anemometer (currently only gill)
7	SA_MODEL	Model of the SA (currently only SA-Gill HS-50 or -100)
8	SA_SW_VERSION	Embedded software version of the SA
9	SA_WIND_DATA_FORMAT	The format of wind data (currently only uvw)
10	SA_NORTH_ALIGNMENT	Specify whether the SA's axes are aligned to transducers (axis) or spars (spar)
11	SA_HEIGHT	The vertical distance between the ground and the center of the device sampling volume (m)

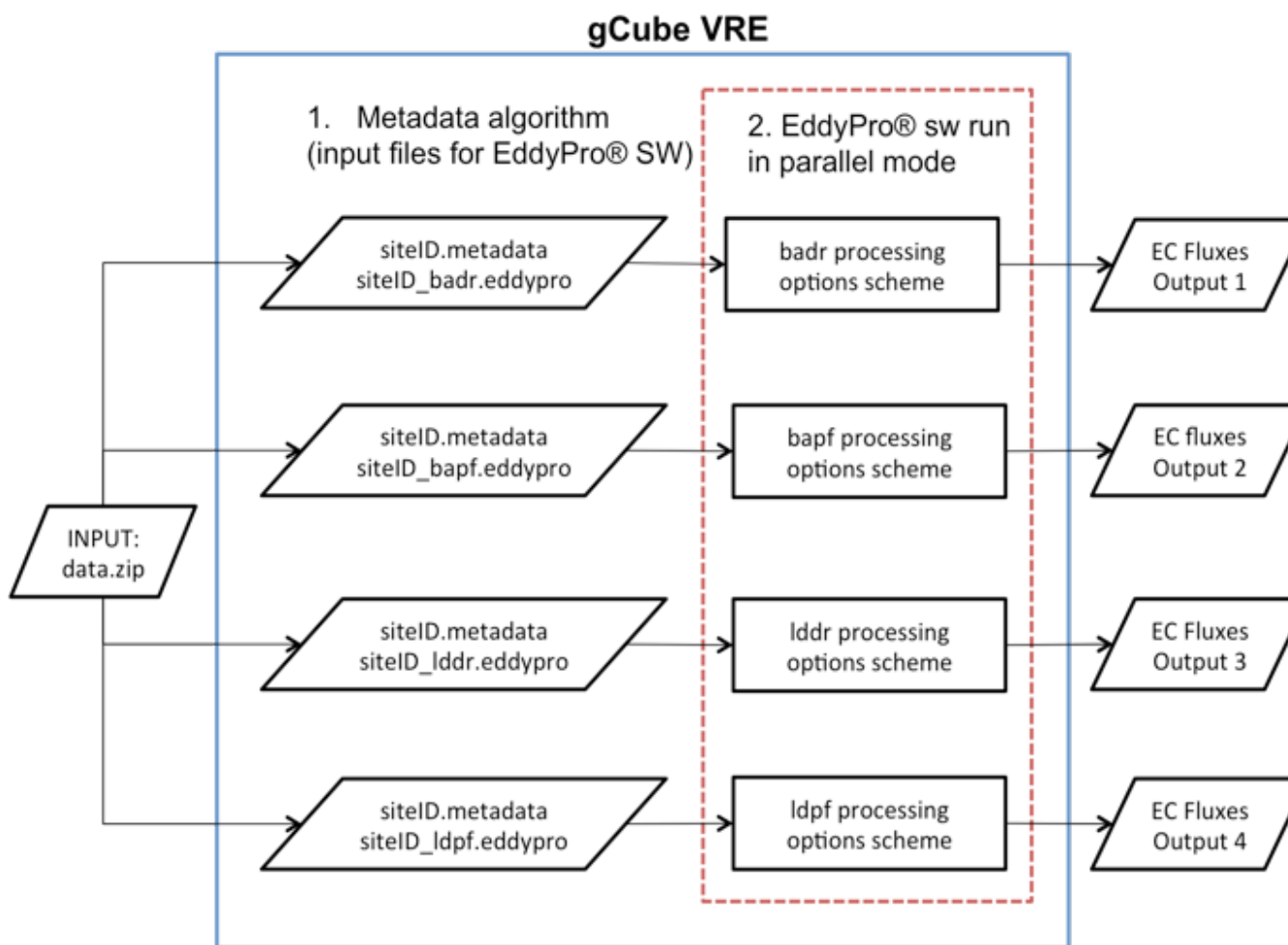
12	SA_NORTH_OFFSET	Specify the SA's yaw offset with respect to local magnetic north (degree positive eastward)
13	GA_MANUFACTURER	Manufacturer of the gas analyzer (currently only licor)
14	GA_MODEL	Model of the GA (currently only GA_CP-LI-COR LI7200)
15	GA_SW_VERSION	Embedded software version of the GA
16	GA_NORTHWARD_SEPARATION	The distance between the center of sample volume of the GA and the SA as measured horizontally along the north-south axis (cm)
17	GA_EASTWARD_SEPARATION	The distance between the center of sample volume of the GA and the SA as measured horizontally along the east-west axis (cm)
18	GA_VERTICAL_SEPARATION	The distance between the center of sample volume of the GA and the SA as measured along the vertical axis (cm)
19	GA_TUBE_DIAMETER	The inside diameter of the intake tube (mm)
20	GA_FLOWRATE	The flow rate of the intake tube (l/min)
21	GA_TUBE_LENGTH	The length of the intake tube (cm)
22	FILE_DURATION	The time span covered by each raw file (min)
23	ACQUISITION_FREQUENCY	The number of records per second in raw files (10 or 20 Hz)
24	FILE_FORMAT	Specify the format of raw files (ASCII or BIN)
25	FILE_EXTENSION	Specify the raw files extension (e.g. .csv, .txt, .dat)
26	LN	Logger number (from 1 to 10)
27	FN	Number of the file generated by the logger (from 1 to 10)
28	EXTERNAL_TIMESTAMP	0 or 1 if the timestamp in the file name refers to the beginning or the end of averaging period, respectively.
29	INTERNAL_TIMESTAMP	1 if there is a timestamp internal to raw files, otherwise 0.
30	EOL	Specify the end of line of raw files (e.g. lf)
31	SEPARATOR	The character that separates individual values in raw files
32	MISSING_DATA_STRING	Specify the character string used for missing data in raw files (e.g. NA, NaN, -9999)
33	NROW_HEADER	The number of rows in the header of the raw file
33+1	COLNAMES_1	Variable name in the first column of the raw data file
33+j	COLNAMES_j	Variable name in the j-th column of the raw data file
33+N	COLNAMES_N	Variable name in the last column of the raw data file

It is important to note that in the current implementation only few sensors are supported (the one used in ICOS) but the structure has been prepared in order to be ready to add new sensors and new processing methods, options and combinations.

The use of different processing options lead to different flux estimates. Discrepancies in flux estimates are caused by systematic errors introduced by methods used in the raw-data processing stage. Since there are not tools to establish a priori which is the best combination of processing options providing unbiased flux estimates, the viable solution, proposed by Sabbatini et al. (2018) and implemented here, involves a multiple processing scheme where EC flux data are calculated according to different combinations of methods.

In particular, EC fluxes are calculated according to four different processing schemes resulting from a combination of block average (ba) or linear detrending (ld) and double rotation (dr) or planar fit<sup>[4]</sup> (pf) processing options (for details see Aubinet et al, 2012)<sup>[5]</sup>. All other processing options remain unchanged: maximum cross-covariance method for time lag determination, spectral correction method proposed by Fratini et al. (2012)<sup>[6]</sup>, statistical tests by Vickers and Mahrt (1997)<sup>[7]</sup> and by Foken and Wichura (1996)<sup>[8]</sup> for data quality control, method by Finkelstein and Sims (2001)<sup>[9]</sup> to estimate random uncertainty.

To reduce the computational runtime, the implementation of the four processing schemes above is performed in parallel mode in the gCube Virtual Research Environment (VRE). The processing path is defined as in Figure 1. When using EC raw data from a single observation tower, the estimated computational time required for a NRT run is about 4 minutes, similar to those required for the run of a single processing scheme.



**Figure 1. EC data processing path**

## Advantages

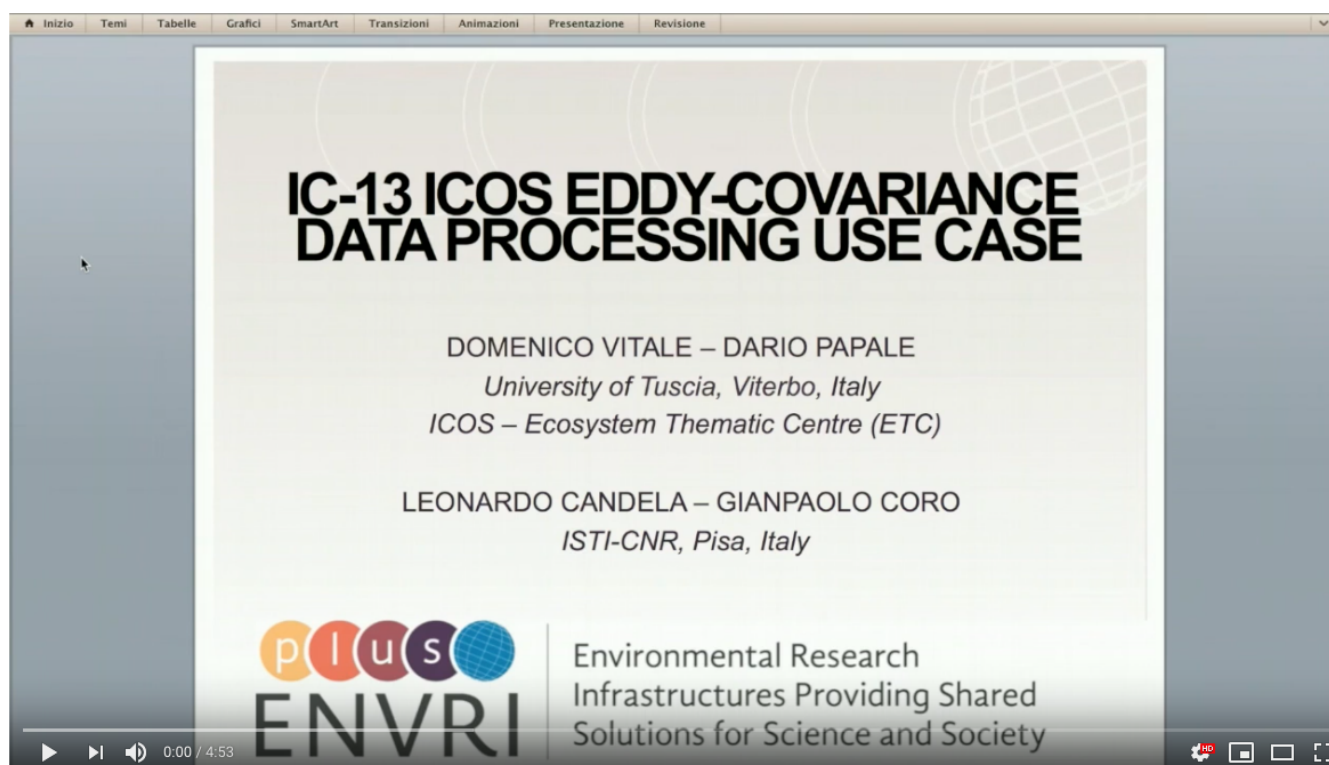
The implementation of a multiple processing schemes as illustrated above and the direct management and use of metadata according to international standard in the eddy covariance community constitutes a novelty in the context of EC data analysis. The main advantage of the multiple processing is twofold. From one hand, it offers the possibility of an extensive evaluation of the effect each method has on flux data estimation. On the other, by combining the output results as described by Sabbatini et al. (2018), it is possible to obtain more consistent estimates of the uncertainty associated to EC fluxes.. The direct use of metadata instead ensure the needed flexibility for a large use of the tool if the new sensors are added in the system.

The efficiency of parallel computing implemented in the VRE, drastically reduces the computational runtime required to obtain flux estimates from different processing options schemes. This constitutes a clear advantage for any user and in particular, for RIs aiming at analyzing routinely large amount of data.

Although, here we selected only 4 processing option schemes, the efficiency of parallel computing implemented in the VRE offers the possibility to increase the number of processing schemes suitable for the EC data processing and also post-processing steps.

This might considerably improve our understanding about the performance of methods developed for EC raw-data processing and about interpretation of resulting fluxes.

## Link to the Demonstrator



Youtube video is at: <https://youtu.be/hod2WksKzV8>

## Contributors

- Dr Domenico Vitale, University of Tuscia, [domvit@unitus.it](mailto:domvit@unitus.it)
- Dr Dario Papale, University of Tuscia, [darpap@unitus.it](mailto:darpap@unitus.it)
- Dr Leonardo Candela, CNR-ISTI, [leonardo.candela@isti.cnr.it](mailto:leonardo.candela@isti.cnr.it)
- Dr Gianpaolo Coro, CNR-ISTI, [gianpaolo.coro@isti.cnr.it](mailto:gianpaolo.coro@isti.cnr.it)

## Reference

- <sup>[1]</sup> Sabbatini S. et al (2018). Eddy covariance raw data processing for CO<sub>2</sub> and energy fluxes calculation at ICOS ecosystem stations, International Agrophysics.
- <sup>[2]</sup> Aubinet, M., Vesala, T., & Papale, D. (Eds.) (2012). *Eddy covariance: a practical guide to measurement and data analysis*. Springer Science & Business Media.
- <sup>[3]</sup> Fratini, G., & Mauder, M. (2014). Towards a consistent eddy-covariance processing: an intercomparison of EddyPro and TK3. *Atmospheric Measurement Techniques*, 7(7), 2273-2281.
- <sup>[4]</sup> Wilczak, J. M., Oncley, S. P., & Stage, S. A. (2001). Sonic anemometer tilt correction algorithms. *Boundary-Layer Meteorology*, 99(1), 127-150.
- <sup>[5]</sup> Aubinet, M., Vesala, T., & Papale, D. (Eds.) (2012). *Eddy covariance: a practical guide to measurement and data analysis*. Springer Science & Business Media.
- <sup>[6]</sup> 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, 53-63.
- <sup>[7]</sup> 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.
- <sup>[8]</sup> Foken, T., & Wichura, B. (1996). Tools for quality assessment of surface-based flux measurements. *Agricultural and forest meteorology*, 78(1-2), 83-105.
- <sup>[9]</sup> Finkelstein, P. L., & Sims, P. F. (2001). Sampling error in eddy correlation flux measurements. *Journal of Geophysical Research: Atmospheres*, 106 (D4), 3503-3509.