DE GRUYTER OPEN

INTERNATIONAL
Agrophysics
www.international-agrophysics.org

Int. Agrophys., 2018, 32, 495-515
doi: 10.1515/intag-2017-0043

Eddy covariance raw data processing for CO_2 and energy fluxes calculation at ICOS ecosystem stations

Simone Sabbatini[1]*, *Ivan Mammarella*[2], *Nicola Arriga*[3], *Gerardo Fratini*[4], *Alexander Graf*[5], *Lukas Hörtnagl*[6], *Andreas Ibrom*[7], *Bernard Longdoz*[8], *Matthias Mauder*[9], *Lutz Merbold*[6,10], *Stefan Metzger*[11,17], *Leonardo Montagnani*[12,18], *Andrea Pitacco*[13], *Corinna Rebmann*[14], *Pavel Sedlák*[15,16], *Ladislav Šigut*[16], *Domenico Vitale*[1], *and Dario Papale*[1,19]

[1]DIBAF, University of Tuscia, via San Camillo de Lellis snc, 01100 Viterbo, Italy
[2]Institute for Atmosphere and Earth System Research/Physics, PO Box 68, Faculty of Science, FI-00014, University of Helsinki, Finland
[3]Research Centre of Excellence Plants and Ecosystems (PLECO), Department of Biology, University of Antwerp, Universiteitsplein 1, 2610, Wilrijk, Belgium
[4]LI-COR Biosciences Inc., Lincoln, 68504, Nebraska, USA
[5]Institute of Bio- and Geosciences, Agrosphere (IBG-3), Forschungszentrum Jülich, Wilhelm-Johnen-Straße, 52428 Jülich, Germany
[6]Department of Environmental Systems Science, Institute of Agricultural Sciences, ETH Zurich, Universitätstrasse 2, 8092 Zürich, Switzerland
[7]DTU Environment, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[8]TERRA, Gembloux Agro-Bio-Tech, University of Liège, 5030 Gembloux, Belgium
[9]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, KIT Campus Alpin, Kreuzeckbahnstraße 19, D-82467 Garmisch-Partenkirchen, Germany
[10]Mazingira Centre, International Livestock Research Institute (ILRI), P.O. Box 30709, 00100 Nairobi, Kenya
[11]National Ecological Observatory Network, Battelle, 1685 38th Street, CO 80301 Boulder, USA
[12]Faculty of Science and Technology, Free University of Bolzano, Piazza Università 1, 39100 Bolzano, Italy
[13]Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Via dell'Università 16, 35020 Legnaro, Italy
[14]Helmholtz Centre for Environmental Research – UFZ, Permoserstr. 15, 04318 Leipzig, Germany
[15]Institute of Atmospheric Physics CAS, Bocni II/1401, CZ-14131 Praha 4, Czech Republic
[16]Department of Matter and Energy Fluxes, Global Change Research Institute, CAS, Bělidla 986/4a, 603 00 Brno, Czech Republic
[17]Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, 1225 West Dayton Street, Madison, WI 53706, USA
[18]Forest Services, Autonomous Province of Bolzano, Via Brennero 6, 39100 Bolzano, Italy
[19]CMCC Euro Mediterranean Centre on Climate Change, IAFES Division, viale Trieste 127, 01100 Viterbo, Italy

Received January 29, 2018; accepted August 20, 2018

Abstract. The eddy covariance is a powerful technique to estimate the surface-atmosphere exchange of different scalars at the ecosystem scale. The EC method is central to the ecosystem component of the Integrated Carbon Observation System, a monitoring network for greenhouse gases across the European Continent. The data processing sequence applied to the collected raw data is complex, and multiple robust options for the different steps are often available. For Integrated Carbon Observation System and similar networks, the standardisation of methods is essential to avoid methodological biases and improve comparability of the results. We introduce here the steps of the processing chain applied to the eddy covariance data of Integrated Carbon Observation System stations for the estimation of final CO_2, water and energy fluxes, including the calculation of their uncertainties. The selected methods are discussed against valid alternative options in terms of suitability and respective drawbacks and advantages. The main challenge is to warrant standardised processing for all stations in spite of the large differences in *e.g.* ecosystem traits and site conditions. The main achievement of the Integrated Carbon Observation System eddy covariance data processing is making CO_2 and energy flux results as comparable and reliable as possible, given the current micrometeorological understanding and the generally accepted state-of-the-art processing methods.

Keywords: ICOS, protocol, method standardisation, biosphere-atmosphere exchange, turbulent fluxes

*Corresponding author e-mail: simone.sabbatini@unitus.it

INTRODUCTION

The eddy covariance (EC) technique is a reliable, widespread methodology used to quantify turbulent exchanges of trace gases and energy between a given ecosystem at the earth's surface and the atmosphere. The technique relies on several assumptions about the surface of interest and the atmospheric conditions during measurements, which however are not always met. An adequate setup of the instrumentation is needed, to minimise the rate of errors and increase the quality of the measurements. A consistent post-field raw data processing is then necessary, which includes several steps to calculate fluxes from raw measured variables (*i.e.* wind statistics, temperature and gas concentrations). Quality control (QC) and the evaluation of the overall uncertainty are also crucial parts of the processing strategy. The present manuscript is based on the official "Protocol on Processing of Eddy Covariance Raw Data" of the Integrated Carbon Observation System (ICOS), a pan-European Research Infrastructure, and is focused on the processing of EC raw data, defining the steps necessary for an accurate calculation of turbulent fluxes of carbon dioxide (F_{CO_2}), momentum (τ), sensible (H) and latent heat (LE). The processing chain for non-CO_2 fluxes is not included here. Besides concentrations and wind measurements, some auxiliary variables are measured in the field (*e.g.* climate variables), and additional parameters calculated during the processing (*e.g.* mass concentrations, air heat capacity), as detailed below. The main EC processing chain, used to derive the final data, is applied once per year; at the same time a near real time (NRT) processing is also performed day-by-day following a similar scheme. The general aim of the definition of the processing scheme is to ensure standardisation of flux calculations between ecosystems and comparability of the results. ICOS Class 1 and Class 2 stations are equipped with an ultrasonic anemometer-thermometer (SAT) Gill HS 50 or HS 100 (Gill Instruments Ltd, Lymington, UK), and an infrared gas analyser (IRGA) LICOR LI-7200 or LI-7200RS (LI-COR Biosciences, Lincoln, NE, USA), both collecting data at 10 or 20 Hz (for further details on ICOS EC setup).

In order to achieve a reliable calculation of the ecosystem fluxes to and from the atmosphere with the EC method, not only the covariance between high-frequency measured vertical wind speed (w) and the scalar of interest (s) has to be calculated, but corrections have to be applied to the raw data to amend the effect of instrumental limitations and deviances from the theory of application of the EC technique (Aubinet *et al.*, 2012).

The application of different corrections to the EC datasets has been widely discussed in the scientific literature, both for single options, steps and for the comprehensive set of corrections. Before the "Modern Age" of EC, began with the start of European and North-American networks (EUROFLUX, AMERIFLUX and FLUXNET, around the years 1995-1997), Baldocchi *et al.* (1988) depicted the state of the art of flux measurements, including EC and its most important corrections developed at the time. Afterwards, the presence of research networks and the technological development enhanced the opportunities for studies and inter-comparisons. A considerable number of publications on the topic followed: a brief and probably not exhaustive review includes the publications of Moncrieff *et al.* (1997), Aubinet *et al.* (2000), Baldocchi (2003), Lee *et al.* (2004), van Dijk *et al.* (2004), Aubinet *et al.* (2012), who described the EC methodology from both theoretical and practical perspectives, going through the several processing options required. On the side of comparisons, Mauder *et al.* (2007) tested five different combinations of post-field data processing steps, while Mammarella *et al.* (2016) performed an inter-comparison of flux calculation in two different EC setup, using two different processing software, and applying different options of processing. For the single options and corrections, the most important publications are listed in the text.

Among several possible options that can be selected and used for each different step, a comprehensive "processing chain" (main processing chain producing final data) has been conceived to be used in the treatment of the raw data collected at ICOS ecosystem stations. In this environmental research infrastructure and particularly in the ecosystem network, standardised processing is necessary to facilitate data inter-comparability between stations and to simplify the organisation of the centralised processing. The fact that all the stations use the same instrumentation setup helps this standardisation process. Some of the steps selected for the ICOS processing need the provisional calculation of specific parameters. This means that a pre-processing is necessary before the application of the actual processing, where the required parameters are extrapolated on a sufficiently long dataset. Two months of raw data covering unchanged site characteristics and instrument setup are deemed sufficient, with the exception of fast-growing crops where a shorter period is accepted. For a list of the most important characteristics see Table 1. When a sufficiently long dataset is not available, the application of a simplified – possibly less accurate – method is necessary, especially for NRT (producing daily output) data processing (see below).

METHODOLOGY

The ICOS processing is performed once every year for each ICOS station. The core processing is made by using LICOR EddyPro software. All the routines used in the processing chain are applied centrally at the Ecosystem Thematic Centre (ETC), and will be publicly available in the Carbon Portal (https://www.icos-cp.eu). In addition to data processing, the ETC is the ICOS facility also in charge

Table 1. List of the main characteristics that should be as constant as possible, and of the variables to be represented, during the pre-processing phase. SAT = sonic anemometer thermometer, IRGA = infrared gas analyser, *RH* = relative humidity

Constant characteristic	Overall variability
Measurement height	*RH*
Canopy height	Atmospheric stability
IRGA sampling line characteristics (tube length and diameter, flow rate)	
Relative position of the SAT and IRGA	
Roughness of the surface	
Absence/presence of leaf/vegetation	
Instruments calibration parameter	

of coordinating the general activities of the stations, providing support to the station teams, following the smooth proceeding of the measurements. The selected methods, however, allow a certain degree of flexibility, especially for those options depending on site characteristics, as shown below. The ICOS processing chain described here is characterised by a Fourier-based spectral correction approach on the covariances of interest. The steps performed in the ICOS processing chain are described in this section, numbered from Step 1 to Step 14, and schematically reported in Figs 1-3. The NRT processing mentioned in the Introduction follows a similar path, with some differences due to the short window used in this case (*i.e.* 10 days), as detailed in a dedicated sub-section.

Step 1: Raw data quality control

Raw data may be erroneous for different reasons. Several tests to flag and possibly discard data exist, and several combinations of them are also possible (Vickers and Mahrt, 1997; Mauder *et al.,* 2013). In principle, the tests selected for ICOS processing work independently on the dataset. However, as for some of them we discard bad quality records, the order of their application is also important. The QC includes a test on the completeness of the available dataset, test on unrealistic values exceeding absolute limits (including their deletion), elimination of spikes, and test on records in distorted wind sectors and their elimination (Fig. 1). The tests with deletion of bad data produce a flag according to Table 2, and the resulting gaps are filled with linear interpolation of neighbouring values only if the number of consecutive gaps is lower than 3, and with NaN if higher, as to reduce as much as possible the impact on the calculation of the covariances (Step 7).

Then, on this pre-cleaned dataset, additional tests are performed to determine the quality of the data, namely an amplitude resolution test, a test on drop-out values, and a test on instrument diagnostics.

Step 1.1: Completeness of the averaging interval (*AI*)

This step checks the number of originally missing records in the time-series. The flagging system is based on the rules shown in Table 2: if the number of missing records is below 1% of what expected based on the acquisition frequency (*i.e.* 18000 for a 30-min long *AI* and sampling

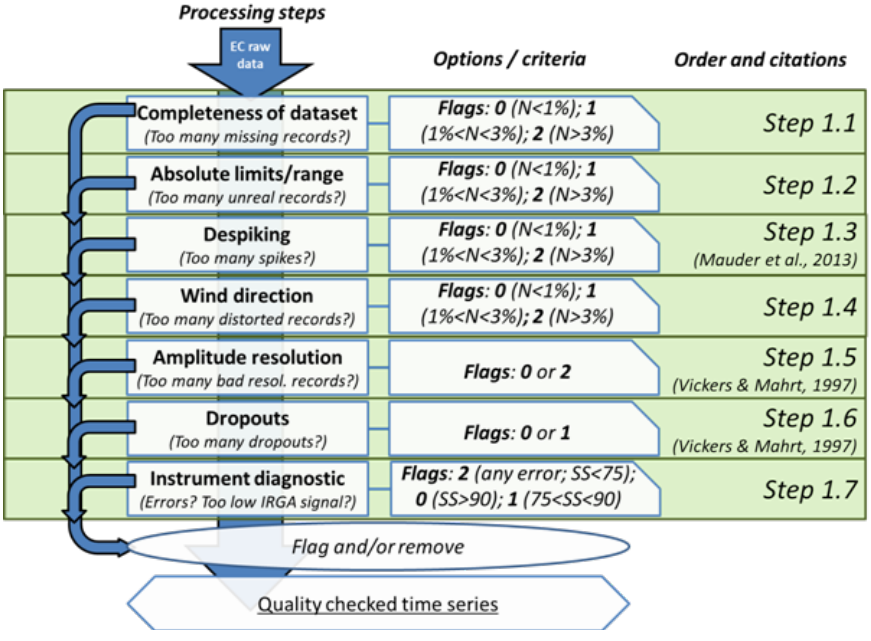

Fig. 1. ICOS flagging system applied to the raw time series in the processing schemes to flag and to eliminate corrupted data. Left: type of test, centre: criteria used in the flagging system, right: steps as numbered in the text and literature references, where it applies. N=number; SS=signal strength.

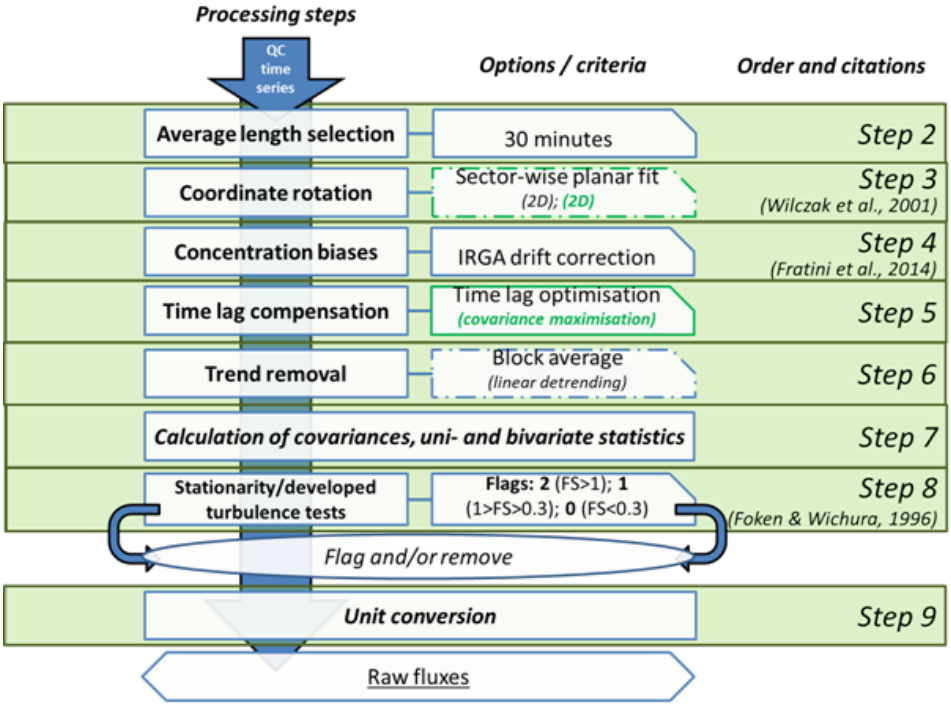

Fig. 2. ICOS scheme of the steps performed for the calculation of "raw" (not corrected) fluxes. Left: corrections applied, centre: options selected or criteria used in the flagging scheme, right: steps as numbered in the text and literature references, where it applies. Green boxes indicate methods that require a pre-processing phase: green text between brackets indicates the method used in case the dataset is not available for the pre-processing. Dashed boxes indicate the steps used for the estimation of the uncertainty: text in black between brackets indicates the different method applied. FS= indicator for flux stationarity.

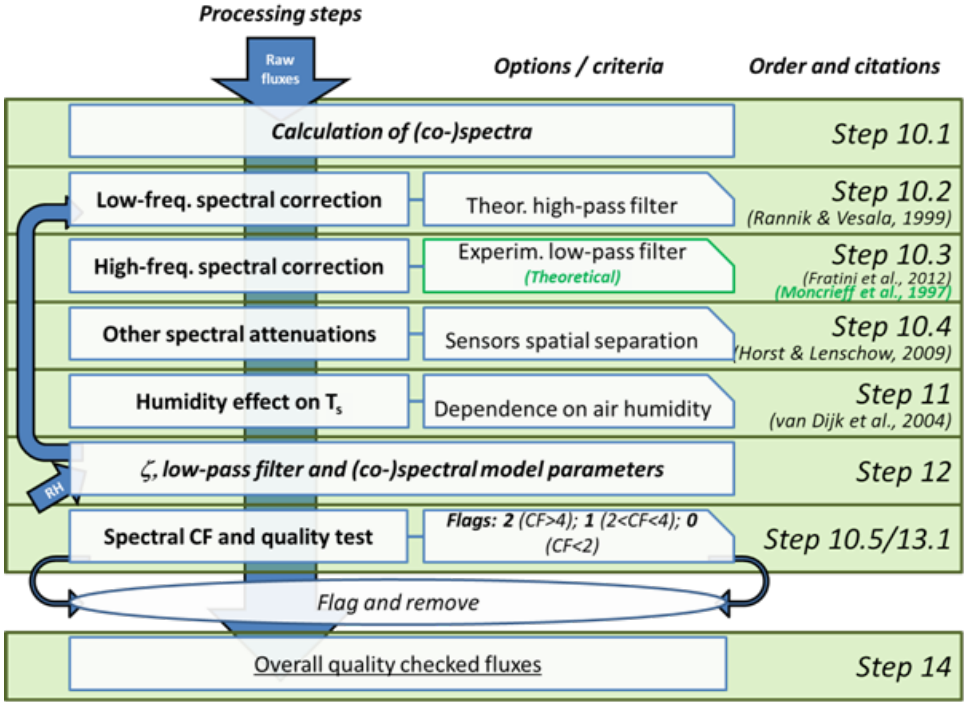

Fig. 3. Spectral corrections. Left: corrections performed, centre: options selected or criteria used in the flagging scheme, right: steps as numbered in the text and literature references, where it applies. ζ=stability parameter; T_s=sonic temperature; *RH*=relative humidity; *CF*=correction factor. Green boxes indicate methods that require a pre-processing on a long dataset: green text between brackets indicates the method used if the dataset is not available.

Table 2. Flagging logic applied to steps from 1.1 to 1.4: 0 = good quality; 1 = moderate quality; 2 = bad quality; N = number of records within the averaging interval (*AI*)

% missing/deleted records	Flag
N<=1%	0
1%<N<=3%	1
N>3%	2

frequency f_s = 10 Hz) the flag is '0'. If this number is between 1 and 3% flag is '1', while if more than 3% of records are missing the flag is '2'.

Step 1.2: Absolute limits test

Unrealistic data values are detected by comparing the minimum and maximum values of all points in the record to fixed thresholds: values above and below the thresholds are considered unphysical. Thresholds defining plausibility ranges for the exclusion of records are defined as follow:

1. Horizontal wind (*u* and *v*)*:* corresponding to the specification of the sonic anemometer (±30 m s^{-1} for Gill HS).
2. Vertical wind (*w*)*:* ±10 m s^{-1}.
3. Sonic temperature (T_s)*:* -25°C to +50°C (corresponding to the operation range of LI-7200).
4. Carbon dioxide (CO_2) dry mole fraction ($X_{CO_2,d}$)*:* 320 to 900 ppm.
5. Water vapour (H_2O) dry mole fraction ($X_{H_2O,d}$)*:* 0 to 140 ppt (corresponding to saturated water vapour pressure at max temperature, 12.3 kPa).

The records are discarded accordingly, and flags issued according to Table 2.

Step 1.3: De-spiking

Spikes are large, unrealistic deviations from the average signal, mainly due to electrical failures, instrument problems, or data transmission issues, that need to be eliminated from the raw dataset to avoid biased (or fictitious) flux values. At the same time, physical and large fluctuations of the measured variables must be retained. The method for spike determination, elimination and flagging is based on the median absolute deviation (*MAD*) as described in Mauder *et al.* (2013). *MAD* is calculated as:

$$MAD = \langle |x_i - \langle x \rangle| \rangle, \quad (1)$$

where: x_i is the single record and $<x>$ the median of the data within a 30-min window. Records out of the following range are considered spikes:

$$\langle x \rangle \pm \frac{h * MAD}{0.6745}, \quad (2)$$

where: *h* is a parameter which is set to 7. The factor of 0.6745 allows the correspondence between *MAD* = 1 and one arithmetic standard deviation in case of Gaussian frequency distribution (Mauder *et al.*, 2013). A flag is set according to Table 2, considering the number of discarded values (*i.e.* values identified as spikes).

Step 1.4: Wind direction test

Sonic anemometer arms and tower structures are known to distort the wind component. Also, other obstacles in the source area could create disturbances. According to this test, records are eliminated by default if the corresponding instantaneous wind direction lies within a range of ±10° with respect to the sonic arm, corresponding also to the direction of the tower. This range corresponds to the distorted wind sector in the calibration tunnel for the Gill HS due to the presence of the arm (Gill, personal communication). Different ranges and additional sectors can however be defined according to site-specific characteristics, such as tower size and transparency. Flags are raised according to Table 1.

Step 1.5: Amplitude resolution test

The amplitute resolution test identifies data with a too low variance to be properly assessed by the instrument resolution. It is performed according to Vickers and Mahrt (1997). A series of discrete frequency distributions for half-overlapping windows containing 1000 data points each is built, moving one-half the window width at a time through the series. For each window position, data are grouped within 100 bins, the total range of variation covered by all bins is the smaller between the full range of occurring data and seven standard deviations. When the number of empty bins in the discrete frequency distribution exceeds a critical threshold value, the record is flagged as a resolution problem. This test only produces a '0' or '2' flag. A flag = '2' is raised if the number of empty bins in the discrete frequency distribution exceeds 70%.

Step 1.6: Drop out test

The drop out test is performed to flag sudden "jumps" in the timeseries which continue for a too long period to be identified by the despiking procedure (offsets). It is based on the same window and frequency distributions used for the amplitude resolution test. Consecutive points that fall into the same bin of the frequency distribution are tentatively identified as dropouts. When the total number of dropouts in the record exceeds a threshold value, the record is flagged for dropouts. This test only produces flag '0' and '1'. The definition of the thresholds for this test varies depending on the location of the record in the distribution. For records between the 10th and the 90th percentile, if the number of consecutive points in the same bin in the discrete frequency distribution is higher than 10% of the total number of records, the flag is '1'; for records beyond this range the threshold is set to 6% to consider the higher sensitivity to the extremes of the distribution.

Step 1.7: Instruments diagnostics

The Gill HS ultrasonic anemometer outputs information on possible failures in transducer communication, while the LICOR LI-7200 diagnostics include information on several

possible failures of the instrument. If one of the diagnostics of the instrument detects a problem within the *AI*, no records are discarded, but the *AI* is flagged '2'. This is done based on the diagnostic codes produced by the sonic anemometer (issued at the selected frequency when an error occurs) and the IRGA (issued at 1 Hz). In addition, based on the average signal strength of the gas analyser, if this is above 90% in the *AI* a flag '0' is produced; if between 75 and 90%, the flag is '1'; if below 75%, flag '2' is produced. We stress that this value is not an absolute indicator of measurement offset, as the error induced by dirt in the cell depends on the spectral nature of the particles as well as on temperature. Furthermore, the user can reset the value to 100% after cleaning, setting the conditions of deployment at the cleaning as reference (LICOR, personal communication). For that reason, the flag based on signal strength is indicative of the amount of dirt in the cell, and no data are discarded from its value.

Step 2: Selection of the length of the *AI*

The *AI* must be long enough to include eddies in the low frequency range of the spectrum, and short enough to avoid inclusion of non-turbulent motions and potential overestimation of turbulent fluxes. The widely accepted length of 30 min (Aubinet *et al.*, 2000), associated with the spectral correction and the Ogive test described below, is deemed appropriate.

Step 3: Coordinate rotation of sonic anemometer wind data

One of the assumptions at the basis of the EC method is that, on average, no vertical motion is present ($\overline{w} = 0$, where the overbar represent the mean value, *i.e.* the value of w averaged over the *AI*). This means that the pole where the SAT is installed has to be perpendicular to the mean streamline. This assumption is not always met due to small divergences between the mean wind streamline and the ground, imperfect levelling of the instrument, or large scale movements, i.e. low frequency contributions (below 1/(length of *AI*) Hz). Hence, the coordinate frame has to be rotated so that the vertical turbulent flux divergence approximates the total flux divergence as close as possible (Finnigan *et al.*, 2003). ICOS uses a modified version of the planar-fit (*PF*) method proposed by Wilczak *et al.*, 2001 to rotate the coordinates of the sonic anemometer, often referenced as sector-wise *PF* (Mammarella *et al.,* 2007). This *PF* method assumes that $\overline{w} = 0$ only applies to longer averaging periods, in the scale of weeks. Firstly, a plane is fitted for each wind sector to a long, un-rotated dataset including all the possible wind directions:

$$\overline{w_0} = b_0 + b_1\overline{u_0} + b_2\overline{v_0}, \quad (3)$$

where $\overline{w_0}$, $\overline{u_0}$ and $\overline{v_0}$ are time series of mean un-rotated wind components and b_0, b_1 and b_2 are regression coefficients, calculated using a bilinear regression. Then, for each averaging period the coordinate system is defined as having the vertical (z) axis perpendicular to the plane, the first horizontal axis (x) as the normal projection of the wind velocity to the plane, and the y-axis as normal to the other axes. The rotation matrix results to be time-independent, and it does not need to be fitted for each *AI.*

Also in this case a period of two months to be pre-processed is deemed sufficient to fit the plane with the number of sectors set to 8 by default. A different number of sectors might be considered on the basis of vegetation and topographic characteristics. For the detailed equations of the method see the above-mentioned publications, and also van Dijk *et al.* (2004). SAT orientation and canopy structure have a significant influence on the fit; this means that a new plane is fitted in case relevant changes occur to these parameters. In case the plane for a given wind sector is not well defined (*e.g.* not enough wind data from this direction), the closest valid sector is used to extrapolate the planar-fit parameter. If the time period for which data with a given site configuration is available is shorter than the required two months, a different method is applied (the *2D* rotation method as in Aubinet *et al.* (2000) and a special flag is raised.

Alternative options: Different options suitable for the coordinate rotation are available in the scientific literature, in particular the so-called *2D* (Aubinet *et al.*, 2000) and *3D* (McMillen, 1988) rotations. The *3D* rotation was excluded from the ICOS routine, while the 2*D* rotation is used for the estimation of the uncertainty. An overall discussion on alternative procedures is reported in the Results and Discussion section.

Step 4: Correction for concentration/mole fraction drift of the IRGA

The LI-7200 is subject to drift in measured mole fractions due to thermal expansion, aging of components, or dirt contamination. If an offset exists in the absorptance determination from the IRGA, not only the estimation of the mole fraction of the gas is biased, but also its fluctuations, due to the polynomial (non-linear) calibration curve used to convert absorptance into gas densities, with an impact on the calculation of turbulent fluxes (Fratini *et al.*, 2014). Even if these issues are expected to be strongly reduced by the filter and heating in the tube and the periodic cleaning and calibrations, the problem might still arise. The method by Fratini *et al.*, 2014 is used to correct the data: assuming that the zero offset in mole fraction readings ($\Delta\chi_{g,w}$) increases linearly between two consecutive field calibrations, the offset measured in occasion of calibration can be converted into the corresponding zero offset absorptance bias (Δa) through the inverse of the factory calibration polynomial, and spreading it linearly between two calibrations leads

each *AI* to be characterised by a specific absorptance offset (Δa_i). The inverse of the calibration curve is then used to convert raw mole fraction values $\chi_{g,w,i}$ into the original absorptances a_i, and these latter corrected for Δa_i using:

$$a_{i,corr} = \frac{a_i - \Delta a_i}{1 - \Delta a_i}, \quad (4)$$

$a_{i,corr}$ is then easily converted to mole fraction $\chi_{g,w,\,corr,i}$ applying the calibration curve, yielding a corrected dataset which can be used for the calculation of fluxes. For the LI7200 the calibration function uses number densities: measurements of cell temperature and pressure are necessary for conversion to/from mixing ratios. After each calibration the method is applied to the data of the LI-7200.

Step 5: Time lag compensation

In EC systems, and particularly those using a closed path gas analyser, a lag exists between measurements of w and gas scalars (s_g), due to the path the air sample has to travel from the inlet to the sampling cell of the IRGA, and due to the separation between the centre of the anemometer path and the inlet of the IRGA. In closed-path systems like the LI-7200, this time lag varies mostly with sampling volume and flow in the tube, and depending on water vapour humidity conditions; more slightly with wind speed and direction. The different treatment of the electronic signal might also contribute. Time series of w and s_g need to be shifted consequently to properly correct this lag, or the covariances will be dampened and the fluxes underestimated. Even if a "nominal" time lag can be easily calculated based on the volume of the tube and the flow rate of the pump, this lag can vary with time due to different reasons such as accumulation of dirt in the filter of the tube, slight fluctuations of the flow rates, and relative humidity (*RH*) content. It may also be different between gases, as molecules of H_2O tend to adhere more to the walls of the tube than other gases like CO_2, resulting in a higher time lag (Ibrom *et al.*, 2007b; Massman and Ibrom, 2008; Mammarella *et al.*, 2009; Nordbo *et al.*, 2014). The method used to correct this lag in ICOS is called "time lag optimisation", as it consists in a modification of the well-known covariance maximisation approach (Aubinet *et al.*, 2012). Both methods seek an extremum in the cross correlation function between w and the s_g that has the same direction as the flux (*e.g.* a minimum for a negative flux). However, in the time lag optimisation the sizes of the window where to look for the time lag are computed continuously instead of being constant. For H_2O, different windows and nominal lags are automatically computed for different *RH* classes in order to take into account the dependence of the time lag on *RH*. The width and centre of the windows are statistically computed on the basis of an existing, long enough dataset, subjected to the cross correlation analysis: the nominal time lag (TL_{nom}) is calculated as the median of the single time lags (TL_i), and the limits of the window (TL_{range}) are defined by:

$$TL_{range} = TL_{nom} \pm 1.5\ \frac{\langle (TL_i - TL_{nom}) \rangle}{0.6745}, \quad (5)$$

where: <> indicates the median operator. For water vapour, this procedure is replicated for each *RH* class, and a different TL_{range} is calculated for any different *RH*. TL_{nom} is used as default value to be used as time lag in case the actual time lag cannot be determined in the plausibility window (*i.e.* if an extremum is not found). This method requests the execution of Steps 1-5 of the standard processing chain on a sufficiently long dataset to calculate the time lag parameters. The dataset has to cover a range of climatic conditions as broad as possible, and its minimum length is two months. The two time series are then shifted accordingly with the calculated time lag in every *AI*. In the case of a shorter period with a given site configuration, the method is switched to the "traditional" covariance maximisation (*i.e.* using a fixed window). A special flag indicates the occurrence of this different method applied to the final users.

Alternative options: The most widespread alternative method consists in maximising the covariance in fixed windows, a less flexible approach especially for H_2O concentrations at different values of *RH*. See also the Results and discussion section.

Step 6: Calculation of background signal and fluctuations

The procedure of separating background signals ($\bar{x}$) and fluctuations (x') of a time series $x=x(t)$ is based on Reynolds averaging (Lee *et al.*, 2004):

$$x' = x - \bar{x}. \quad (6)$$

This operation introduces a spectral loss, especially at the low frequency end of the spectra (Kaimal and Finnigan, 1994, see Step 10.2). The method used in the main processing chain of ICOS is the so-called block averaging (*BA*). The effect on the spectra can be conveniently represented with a high-pass filter, and then corrected (see below). Block averaging consists of time averaging all the instantaneous values in an *AI*, and the subsequent calculation of fluctuations as instantaneous deviations from this average ($\bar{x}_{BA}$):

$$\bar{x}_{BA} = \frac{1}{N}\sum_{j=1}^{N} x_j, \quad (7)$$

$$x_j'_{BA} = x_j - \bar{x}_{BA}, \quad (8)$$

where: N is the number of digital sampling instants j of the time series $x(t)$ (discrete form) in the *AI*. As the mean is constant in the averaging interval, *BA* fulfils Reynolds averaging rules (Lee *et al.*, 2004) and impacts at a lower degree the spectra (Rannik and Vesala, 1999).

Alternative options: two alternative methods are widely used in the EC community, namely the linear de-trending (*LD*, Gash and Culf, 1996) and the autoregressive filtering (*AF*, Moore, 1986; McMillen, 1988). The *LD* is used as an alternative method to calculate the uncertainty in the data. An overall discussion is reported in the respective section.

Step 7: Calculation of final covariances and other statistics

After the steps described in the above, the final covariances are calculated, together with other moment statistics. In turbulent flows, a given variable x can be statistically described by its probability density function (*PDF*) or associated statistical moments. Assuming stationarity and ergodicity of x (Kaimal and Finnigan, 1994) we are able to calculate statistical moments, characterizing the *PDF* of x, from the measured high frequency time series $x(t)$.

Single-variable statistics include moments used for flagging data: mean value (first order moment), variance (and thus standard deviation), skewness and kurtosis, *i.e.* second, third and fourth order moments. Indicating with n the moment, we can resume all of them in a unique equation:

$$\overline{x'^n} = \frac{1}{N}\sum_{j=1}^{N}(x_j - \bar{x})^n, \quad (9)$$

where: $\overline{x'^n}$ indicates the n-th moment, $N = l_{AI} f_s$, l_{AI} the length of the averaging interval (s) and f_s the sampling frequency (Hz), x_j the instantaneous value of the variable x, $\bar{x}$ the mean value.

In the same way, moments of any order associated with joint probability density function of two random variables can be calculated. In particular, covariances among vertical wind component w and all other variables also need to be calculated. In general, the covariance of any wind component u_k or scalar s_g with another wind component u_i is calculated as follows:

$$\overline{s_g{}'u_i{}'} = \frac{1}{N}\sum_{j=1}^{N}\left[\left(s_{g,j} - \bar{s_g}\right)\left(u_{i,j} - \bar{u_i}\right)\right] = \frac{1}{N}\sum_{j=1}^{N} s_{g,j}{}'u_{i,j}{}', \quad (10)$$

$$\overline{u_k{}'u_i{}'} = \frac{1}{N}\sum_{j=1}^{N}\left[\left(u_{k,j} - \overline{u_k}\right)\left(u_{i,j} - \bar{u_i}\right)\right] = \frac{1}{N}\sum_{j=1}^{N} u_{k,j}{}'u_{i,j}{}', \quad (11)$$

where: u_k, with $k = 1, 2, 3$, represents wind components u_i, v_i or w_i, and the subscript j indicates the instantaneous values of the corresponding scalar or wind component from 1 to N.

Please note that this step refers to the calculation of the final covariances used for the calculation of the fluxes. However, statistics are also calculated after each step of the processing.

Step 8: Quality control on covariances: steady-state and well developed turbulence tests

These two tests (Foken and Wichura, 1996), applied to the calculated covariances, are necessary to check the verification of basic assumptions of the EC method, *i.e.* the stationarity of the variables in the *AI* and the occurrence of conditions where the Monin-Obukhov similarity theory applies.

Step 8.1 Steady-state test (Foken and Wichura, 1996)

Typical non-stationarity is driven by the change of meteorological variables with the time of the day, changes of weather patterns, significant mesoscale variability, or changes of the measuring point relative to the measuring events such as the phase of a gravity wave. The method to define non-steady-state conditions within *AI* by Foken and Wichura (1996) uses an indicator for flux stationarity *FS*:

$$FS = \left|\frac{\overline{w's'_{(AI/m)}} - \overline{w's'_{AI}}}{\overline{w's'_{AI}}}\right|, \quad (12)$$

where: $m = 6$ is a divisor of the *AI* defining a 5 mins window. $\overline{x's'}_{(AI/m)}$ is the mean of the covariance between w and s calculated in the m windows of 5 min:

$$\overline{w's'}_{(AI/m)} = \frac{1}{m}\sum_{i=1}^{m}\overline{w's'}_i. \quad (13)$$

The flux is considered non-stationary if *FS* exceeds 30%.

Step 8.2 Well-developed turbulence test (Foken and Wichura, 1996)

Flux-variance similarity is used to test the development of turbulent conditions, where the normalized standard deviation of wind components and a scalar are parameterized as a function of stability (Stull, 1988; Kaimal and Finnigan, 1994). The measured and the modelled normalized standard deviations are compared according to:

$$ITC_\sigma = \left|\frac{(\sigma_x/x_*)_{mod} - (\sigma_x/x_*)_{meas}}{(\sigma_x/x_*)_{mod}}\right|, \quad (14)$$

where: variable x may be either a wind velocity component or a scalar, and x_* the appropriate scaling parameter. The test can in theory be done for the integral turbulence characteristics (*ITC*) of both variables used to determine the covariance, but it is applied only on w (more robust), with the exception of $w'u'$ for which both parameters are calculated, and the ITC_σ is then derived using the parameter that leads to the higher difference between the modelled and the measured value (numerator Eq. (14)). The models used for $(\sigma_x/x_*)_{mod}$ are the ones published in Thomas and Foken (2002) (Table 2). An ITC_σ exceeding 30% is deemed as the absence of well-developed turbulence conditions.

Step 9: From covariances to fluxes: conversion to physical units

Calculated covariances as described above need to be converted in physical units, mainly *via* multiplication by meteorological parameters and constants. The subscript '0' indicates that the fluxes are still not corrected for spectral attenuation, and the sonic temperature not yet corrected for the humidity effect:

Flux of buoyancy (W m^{-2}):

$$H_0 = \overline{w'T_s'}\rho_{a,m}c_p\,, \quad (15)$$

Flux of CO_2 (µmol m^{-2} s^{-1}):

$$F_{CO_2,0} = \overline{w'\chi_{co2,d}'}\,1/(\frac{RT_a}{P_d}), \quad (16)$$

Flux of H_2O (mmol m^{-2} s^{-1}):

$$F_{H_2O,0} = \overline{w'\chi_{v,d}'}\,1/(\frac{RT_a}{P_d}), \quad (17)$$

Flux of evapotranspiration (kg m^{-2} s^{-1})

$$E_0 = \overline{w'\chi_{v,d}'}\,1/(\frac{RT_a}{P_d})\,M_v, \quad (18)$$

Flux of latent heat (W m^{-2})

$$LE_0 = \overline{w'\chi_{v,d}'}\,1/(\frac{RT_a}{P_d})\,M_v\lambda\,, \quad (19)$$

Flux of momentum (kg m^{-1} s^{-2})

$$\tau_0 = \sqrt[2]{\overline{u'w'}^2 + \overline{v'w'}^2}\rho_{a,m}\,, \quad (20)$$

Friction velocity (m s^{-1})

$$u_* = \sqrt[4]{\overline{u'w'}^2 + \overline{v'w'}^2}, \quad (21)$$

where $\chi_{CO_2,d}$ is the mixing ratio of CO_2, $\chi_{v,d}$ that of water vapour; M_v the molecular weight of water vapour: 0.01802 kg mol^{-1}; λ the specific evaporation heat: $(3147.5\text{-}2.37T_a)$ 10^3; and $p_d = p_a - e$ is the dry air partial pressure. Gas fluxes are calculated using the dry mole fraction data, which are not affected by density fluctuations, differently from *e.g.* mole fraction (Kowalski and Serrano Ortiz, 2007). As the ICOS selected IRGA is able to output the gas quantities in terms of dry mole fraction, this is a mandatory variable to be acquired at ICOS stations. The IRGA applies a transformation in its software to convert mole fraction in dry mole fraction using the high-frequency water vapour concentration data.

Steps to calculate the air density $\rho_{a,m}$:

a) $v_a = \frac{R\,T_a}{p_a}$ (air molar volume); (22)

b) $\rho_{v,m} = \frac{\chi_v M_v}{v_a}$ (water vapour mass concentrations); (23)

c) $\rho_{d,m} = \frac{(1-\chi_v)M_d}{v_a}$ (dry air mass concentrations); (24)

d) $\rho_{a,m} = \rho_{d,m} + \rho_{v,m} = \frac{M_d + \chi_v(M_v - M_d)}{v_a}$. (25)

Steps to calculate the air heat capacity at constant pressure c_p:

a) $c_v = 1859 + 0.13RH + (0.193 + 5.6\,10^{-3}RH)T_a + (10^{-3} + 5\,10^{-5}RH){T_a}^2$ (26)

(water vapour heat capacity at constant pressure);

b) $c_d = 1005 + \frac{(T_a+23.12)^2}{3364}$ (27)

(dry air heat capacity at constant pressure);

c) $q = \rho_{v,m}/\rho_{a,m}$ (specific humidity); (28)

d) $c_p = c_d(1-q) + c_v q$. (29)

Step 10: Spectral correction

In the frequency domain, the EC system acts like a filter in both high and low frequency ranges. Main cause of losses at the low frequency range is the finite *AI* that limits the contribution of large eddies, together with the method used to calculate turbulent fluctuations (see Step 6), which acts as a high-pass filter. Main cause for losses at the high frequency range is the air transport system of the IRGA, together with the sensor separation and inadequate frequency response of instruments leading to the incapability of the measurement system to detect small-scale variations. To correct for these losses, the correction is based on the knowledge of the actual spectral characteristics of the measured variables, and on the calculation of the difference with a theoretical, un-attenuated spectrum. A loop is present in this part of the processing chain (Fig. 3), where the corrections are performed iteratively to refine the calculation of the spectral parameters.

Step 10.1: Calculation and quality of spectra and cospectra

Calculation of power spectra (and cospectra) is a fundamental step to perform spectral analysis and correction of the fluxes. Calculated fluxes are composed of eddies of different lengths, *i.e.* signals of different contributions in the frequency domain: the knowledge of the spectral characteristics of the EC system and the shape of the model spectra is crucial for the execution of spectral corrections. In ICOS processing the calculation of power (co)spectra, at the basis of the analysis in the frequency domain, is performed using the Fourier transform (Kaimal and Finnigan, 1994), which allows to correct the spectral losses analysing the amount of variance associated with each specific frequency range. Validity of the Fourier transform for EC measurements is supported by the Taylor's hypothesis of frozen turbulence (Taylor, 1938). For further details on the theory (Stull, 1988).

Full spectra and cospectra are calculated for each *AI* over a frequency range up to half the acquisition frequency (Nyquist frequency). The algorithm used to calculate the Fourier transform is the FFT (Fast Fourier Transform), which is applied to the time series after their reduction and tapering (Kaimal and Kristensen, 1991). In order to apply the FFT algorithm, the number of samples used will be equal to the power-of-two closest to the available

number of records (*i.e.* 2^{14} for 10 Hz and 2^{15} for 20 Hz data). Binned cospectra are also calculated to reduce the noise dividing the frequency range in exponentially spaced frequency bins, and averaging individual cospectra values that fall within each bin (Smith, 1997). Finally, spectra and cospectra are normalised using relevant run variances and covariances, recalculated for the records used for (co)spectra calculation, and averaged into the exponentially spaced frequency base. The integral of the spectral density represents the corresponding variance and the integral of the cospectral density the corresponding covariance (Aubinet *et al.*, 2000). Normalisation by the measured (co)variances forces the area below the measured (co)spectra to 1, but the ratio between the areas (model/measured (co)spectra) is preserved. Thus, addressing losses in the different ranges of frequency allows having a flux corrected for spectral losses due to different causes.

The approach to compensate for spectral losses is different in the low and high frequency ranges, and for fluxes measured by the SAT and gas scalars. All the methods used in ICOS processing are based on Fourier transforms, and rely on the calculation of an ideal (co)spectrum unaffected by attenuations ($CS_{Id,F}$), on the definition of a transfer function (*TF*) characteristic of the filtering made by the measuring system (both functions of frequency *f* (Hz)), and on the estimation of a spectral correction factor (*CF*):

$$CF = \frac{\int_0^{\infty} CS_{Id,F}(f)df}{\int_0^{\infty} CS_{Id,F}(f) * TF(f)df}. \quad (30)$$

Multiplying the attenuated flux for *CF* leads to the spectral corrected flux. Several parameterisations exist of the ideal cospectrum, which depends mainly upon atmospheric stability conditions, frequency, horizontal wind speed, measurement height and scalar of interest. One of the most used model cospectrum is based on the formulation by Kaimal *et al.* (1972), and is implemented also in the ICOS processing when an analytic approach is adopted. In the following we describe the application of spectral corrections separately for the different frequency ranges and the different types of flux.

Step 10.2: Low-frequency range spectral correction (Moncrieff *et al.*, 2004)

Losses at the low end of the (co)spectra mainly arise when the necessarily finite average time is not long enough to capture the full contribution of the large eddies to the fluxes. The methods used to separate fluctuations from the background signal enhance this event, excluding eddies with periods higher than the averaging interval. This correction is based on an analytic approach.

In this case the ideal cospectrum is based on the formulation of Moncrieff *et al.* (1997), while *TF* represents the dampening at the low frequency range. Application of *TF* to the cospectra gives the attenuation due to frequency losses, and because of the proportional relationship between a flux and the integral of its cospectrum, a spectral correction factor can be easily calculated. This can be done on the basis of the method used to separate the fluctuations from the rest, which acts as a filter as said above. From Rannik and Vesala (1999) and Moncrieff *et al.* (2004) we get the high-pass transfer function corresponding to the block average:

$$HPTF_{BA} = \frac{\sin^2(\pi f l_{AI})}{(\pi f l_{AI})^2}. \quad (31)$$

where: l_{AI} is the length of *AI* (s). The corresponding correction factor (*HPSCF*) is calculated as the ratio:

$$HPSCF = \frac{\int_0^{f_{max}} CS_{Id,F}(f)dn}{\int_0^{f_{max}} CS_{Id,F}(f) * HPTF(f)dn}, \quad (32)$$

where: f_{max} is the higher frequency contributing to the flux, and $CS_{Id,F}(f)$ represents the model describing the normalised un-attenuated cospectrum for a given flux *F*, depending on the natural frequency *f*.

Step 10.3: High-frequency range spectral correction

Corrections of the high frequency part of the spectra and cospectra are more detailed as this end of the spectra is defined and several publications exist to correct the corresponding attenuations.

A fully analytic method is applied to correct *H* and τ, an experimental method for the other fluxes.

Step 10.3.1: Fully analytic method (Moncrieff *et al.*, 1997)

This method is used in ICOS processing only to correct anemometric fluxes (*i.e. H* and τ). With this method both the ideal cospectral densities $CS_{Id,F}$ and the low-pass transfer function $LPTF_{SAT}$ are calculated in an analytical (theoretical) way. $LPTF_{SAT}$ results from a combination of transfer functions each characterising a source of attenuation, as described among others in Moore *et al.* (1986); Moncrieff *et al.* (1997), from where all of the transfer functions reported below are taken, while the cospectral model is based on Kaimal *et al.* (1972):

1. $LPTF_{SAT}$ of sensor-specific high frequency loss:

$$LPTF_{SAT,s} = \frac{1}{\sqrt{1+(2\pi f \tau_s)^2}}, \quad (33)$$

where: τ_s the instrument-specific time constant.

2. *TF* of the sonic anemometer path averaging for *w* and for *Ts,* respectively:

$$LPTF_{SAT,pw} = \frac{2}{\pi f_w}\left\{1 + \frac{e^{(-2\pi f_w)}}{2} - 3\frac{\left[1-e^{(-2\pi f_w)}\right]}{4\pi f_w}\right\}, \quad (34)$$

$$LPTF_{SAT,pTs} = \frac{1}{2\pi f_w}\left\{3 + e^{(-2\pi f_w)} - 4\frac{\left[1-e^{(-2\pi f_w)}\right]}{2\pi f_w}\right\}, \quad (35)$$

where f_w is the normalised frequency:

$$f_w = \frac{f l_p}{\bar{u}}, \quad (36)$$

and l_p is the path length of the sonic anemometer transducers.

The final transfer function of the analytic method results from the product:

$$LPTF_{SAT} = LPTF_{SAT,s} \sqrt{LPTF_{SAT,pw}} \sqrt{LPTF_{SAT,pTs}} . \quad (37)$$

Application of Eq. (36) using $LPTF_{SAT}$ allows the calculation of the corresponding low-pass spectral correction factor for sonic fluxes ($LPSCF_{SAT}$).

Step 10.3.2: Experimental method (Ibrom *et al.*, 2007a; Fratini *et al.*, 2012)

The experimental method firstly described in Ibrom *et al.* (2007a) and then refined in Fratini *et al.*, 2012 is applied to gas scalars. Assuming spectral similarity, this method derives *in-situ* the un-attenuated spectral density, assuming that the spectrum of the sonic temperature is unaffected by the main sources of dampening, and then represents a proxy for the ideal spectrum of gas concentrations.

The effect of the measuring system is approximated with a first-order recursive, infinite impulse response filter (*IIR*), whose amplitude in the frequency domain is well approximated by the Lorentzian:

$$LPTF_{GAS} = \frac{S_{meas}}{S_{Ts}} = \frac{1}{1+(f/f_c)^2}, \quad (38)$$

where: S_{meas} represents the ensemble averaged measured spectra, and S_{Ts} the ensemble averaged spectra of sonic temperature representing the ideal unaffected spectrum. This way, the cut-off frequency f_c can be determined *in-situ* based on the actual conditions of the measuring system and of the ecosystem using Eq. (38). For water vapour in closed-path systems, this fit is performed for different classes of *RH*, and then the general cut-off frequency is extrapolated fitting the single cut-off frequencies f_{ci} to an exponential function:

$$f_c(RH) = e^{(aRH^2+bRH+c)}, \quad (39)$$

where: *a*, *b* and *c* are fitting parameters.

The low-pass spectral correction factor for gas scalars is then calculated as:

$$LPSCF_{GAS} = \frac{A_1 \bar{u}}{A_2+f_c} + 1, \quad (40)$$

where: A_1 and A_2 are parameters determined by filtering sonic temperature time series with the low pass recursive filter using varying values of the filter constant, in both stable and unstable atmospheric stratifications (see Ibrom *et al.*, 2007a for details). This leads to a $LPSCF_{GAS}$ that not only depends on stratification, but is also site-specific.

Both the gas (attenuated) and the sonic temperature (un-attenuated) spectra are intended as ensemble-averaged spectra. Averaging is performed on binned spectra (see Step. 10.1) selected for high-quality. This quality control is based on the same tests applied to the time series (see above), including also skewness and kurtosis. Moreover, only spectra for *AI*s with high fluxes are used for the determination of high-quality spectra, thus excluding periods of low turbulence and low signal-to-noise ratio. Thresholds are therefore applied on the calculated fluxes before spectral correction. Thresholds are defined differently between stable and unstable atmospheric conditions: in the latter case thresholds are less strict, in order to reduce the risk of missing night-time spectra. The situations excluded are reported in Table 3. In addition, H_2O spectra are sorted into nine relative humidity classes.

Fratini *et al.* (2012) suggested a modification of this method to provide a refinement in case of large fluxes. The calculation of $LPSCF_{GAS}$ is as above (Eq. (40), but see the original papers for a difference in the approach) only for small fluxes, while for large fluxes it becomes:

$$LPSCF_{GAS} = \frac{\int_{fmin}^{fmax} CS_H(f)dn}{\int_{fmin}^{fmax} CS_H(f) * TF(f|f_c)dn}, \quad (41)$$

thus applying the general Eq. (38) with measured data, using the lower frequency allowed by the averaging interval (f_{min}) and the Nyquist frequency ($f_{max}=f_s/2$); using the current cospectral density of sensible heat (CS_H) as a proxy for the ideal, unaffected cospectral density for each *AI*; and using the *TF* determined by the low-pass filter with the f_c determined *in-situ*, dependent from *RH* in the case of H_2O fluxes.

"Small" fluxes are defined on the basis of the following thresholds: |H| and |LE| < 20 W m^{-2}; |F_{CO_2}|< 2 µmol m^2 s^{-1}.

The dataset used to calculate the spectral characteristics and to extrapolate $LPSCF_{GAS}$ should be representative of the overall variability of micrometeorological conditions, especially *RH* and the overall ecosystem characteristics (dynamics of canopy development) and EC system (not

Table 3. Thresholds used to discard spectra from the calculation of ensemble averages in spectral corrections, depending on atmospheric stability. Stability conditions are defined by the Obhukov length *L*

Flux	Unstable conditions (-650 < *L* < 0)	Stable conditions (0 < *L* < 1000)
$\|H_0\|$	< 20 W m^{-2}; > 1000 W m^{-2}	< 5 W m^{-2}; > 1000 W m^{-2}
$\|LE_0\|$	< 20 W m^{-2}; > 1000 W m^{-2}	< 3 W m^{-2}; > 1000 W m^{-2}
$\|FCO_{2,0}\|$	< 2 µmol m^{-2} s^{-1}; > 100 µmol m^{-2} s^{-1}	< 0.5 µmol m^{-2} s^{-1}; > 100 µmol m^{-2} s^{-1}
	< 0.2 m s^{-1}; > 5 m s^{-1}	< 0.05 m s^{-1}; > 5 m s^{-1}

changing setup). In Table 1 the most important characteristics that have to be homogeneous throughout the overall period of spectral characterisation are reported, together with the most important conditions whose variability has to be represented as much as possible. This means that even if in theory the full-year period used in ICOS processing is sufficiently long, every time the setup is changed a new characterisation of (co)spectra is needed. For fast-growing species an exception is needed, due to both the fast changes in canopy height, and the consequent adaptation of the measurement height made at the stations to keep the distance between measurement and displacement height as constant as possible. In this case a two-week period is deemed enough to calculate spectral parameters, and the station team should ensure as much as possible periods of minimum two weeks with a constant measurement height and the same EC configuration in general. In any different case when a given configuration is operational for less than two weeks, the fully analytic method for spectral corrections is applied (Moncrieff *et al.*, 1997), more generic but with no need of calculating spectral parameters on a long dataset. A special flag is issued accordingly.

Alternative options: several alternatives are published in the scientific literature for the high-frequency correction for gas scalars, belonging to the two big groups of analytic and experimental methods. In the Results and discussion section is presented a wider discussion on alternative options.

Step 10.4: Losses due to instrumental separation (Horst and Lenschow, 2009)

The spatial separation between the gas analyser (inlet) and the sonic anemometer path causes attenuation in the high-frequency range of the cospectra which can be often neglected. However, it may become important especially over smooth surfaces at low measurement heights, due to the dependence of the cospectral peak frequency (f_p) on the atmospheric stratification (f_p shifts to lower frequencies with increasing (z-d)/L, Kaimal and Finnigan, 1994). To account for these losses a correction method was developed by Horst and Lenschow (2009), which depicts the flux of a scalar measured at a distance l_s from the centre of the sonic path as a function of the distance itself and the frequency at the peak of the corresponding cospectrum according to:

$$F(l_s) = F_0 * e^{(-n_p l_s)}, \tag{42}$$

where: F_0 is the un-attenuated flux, n_p=f_p (z-d)/$\overline{u}$, f_p being the frequency at the peak of the cospectrum and z the measurement height. This formulation can then be modulated for the three linear components of the 3D distance respective to the wind direction, *i.e.* along-wind, cross-wind and vertical. However, the along-wind component is mostly corrected when compensating for the time lag. For that reason, the method implemented in ICOS only uses the cross-wind and the vertical component of this correction.

Step 10.5: Calculation of the spectral correction factor

The implementation of the above-described methods is based on the calculation of different transfer functions in the frequency domain representing the dampening at both low and high frequencies, characterised by a specific cut-off frequency. Application of transfer functions to the cospectra gives the attenuation due to frequency losses. A correction factor including all the spectral corrections as a whole can be calculated. In practical terms, this is done by multiplying $HPTF_{BA}$ and $LPTF_{SA}$, resulting in a band pass transfer function for sonic variables correction ($BPTF_{SA}$) which is then inserted in Eq. (30) to derive a unique spectral correction factor. For the gas variables, whose high-frequency attenuation was instead calculated with the experimental approach, the HPSCF is first used to correct the gas fluxes for the high-pass filtering effects and then the *LPSCF* is applied for the low-pass filtering effects.

Step 11: Humidity effect on sonic temperature

Sonic temperature (T_s) is used to calculate the covariance $\overline{w'Ts'}$, which approximates the buoyancy flux, as T_s is close to virtual temperature (T_v). The sensible heat flux (H) is defined as the covariance of w with the real air temperature, which may deviate from T_s for 1-2% due to the dependency of sound velocity on water vapour pressure (Aubinet *et al.*, 2012):

$$T_s = T(1 + 0.32\, {}^{e}/_{p}), \tag{43}$$

where: T is the real absolute temperature, e the partial pressure of water vapour and p the air pressure. A correction for this effect is thus needed, and is based on papers by Schotanus *et al.* (1983) and van Dijk *et al.* (2004), using the following equation:

$$H_{corr} = H_{sp} - \rho_{a,m} c_p 0.51(\overline{T}_s \frac{E_{sp}}{\rho_{a,m}} + \bar{q}\overline{w'T_s'}), \tag{44}$$

where: H_{corr} represents the sensible heat flux corrected for humidity effect, H_{sp} approximates the buoyancy flux corrected for spectral losses, $\rho_{a,m}c_p$ is the product of air density and air heat capacity, E_{sp} is the evapotranspiration flux corrected for spectral losses, and q the specific humidity.

Step 12: Iteration and calculation of atmospheric stability

Some of the above mentioned methods depend on atmospheric stability. This characteristic is commonly described using the parameter ζ, defined as:

$$\xi = \frac{z-d}{L}, \tag{45}$$

where: d is the displacement height, *i.e.* the height at which the wind speed would assume zero if the logarithmic wind profile was maintained in absence of vegetation. It is mostly calculated from canopy height (h_c) as $d = 0.67h_C$. However,

in case the ICOS station has a wind profile (not mandatory), this parameter can be calculated more accurately. L is instead the Obukhov length:

$$L = -\frac{T_p * u_*^3}{\kappa * g * \overline{w'T_s'}}, \quad (46)$$

where: T_p is the potential temperature (K), calculated from air temperature as $T_p = T_a(\frac{p_0}{p})^{0.286}$ (p_0 is the reference pressure set to 10^5 Pa), κ = 0.41 is the von Kármán constant and g is the acceleration due to earth's gravity (9.81 m s^{-2}). Fluxes are thus involved in the calculation of L, meaning that some of the spectral correction methods above lead to a modification of L itself, and hence of ζ, which is used in spectral correction, thus there is the need for iteration to achieve higher precision (Fig. 3, Clement, 2004).

Step 13: Quality-control tests on calculated fluxes

After the calculation of the fluxes as described above, tests can be performed for a further exclusion of bad quality data, which are based on the spectral analysis.

Step 13.1: Spectral correction factor test

Spectral correction at both low and high frequency ranges basically estimates the attenuations of the EC signal due to different causes, and corrects for this losses. If however the losses turn out to be very big, *i.e*. if the EC system is capable of detecting only a small portion of the spectrum, the *AI* is discarded. The method for filtering data after fluxes calculation takes into consideration the portion of the power spectrum measured by the EC system: if the spectral correction factor obtained applying the spectral correction method is below 2, *i.e.* if more than half of the power spectrum has been detected by the EC system, the flag is '0'. If the spectral correction factor is between 2 and 4, *i.e*. if the EC system is capable of detecting less than half but more than one third of the power spectrum, the flag is '1'. Otherwise (less than one third of the power spectrum detected), flag is '2' and the calculated flux eliminated.

Step 13.2: Ogives test

The finite ogive (*Og*) is defined as the cumulative co-spectrum. At each given frequency, the ogive represents the integration of the co-spectrum from the current frequency to the Nyquist frequency. Then its value at the lowest frequency provides the integration of the full co-spectrum, which corresponds to the covariance (Berger *et al.*, 2001). Ogives can be used to evaluate the suitability of the chosen *AI*. Provided that conditions of stability are met in the *AI*, the typical form of an Ogive for EC fluxes has an S-shape: significant deviations from this shape indicate problems in the measured covariance. In particular, if the ogive has a large slope at the lowest frequencies, it means that significant fractions of the flux occur at these timescales, and then that the chosen *AI* may not be long enough. To use Ogives as a test, a smoothing filter is applied to reduce scatter at the higher frequencies, and the Ogives are normalised to the corresponding covariances to make them comparable. To determine the slope in the low frequency range, two thresholds are selected as in Spirig *et al.* (2005): the frequency threshold (0.002 Hz) determining the lower part of the frequency range, and the threshold (15%) for the value of the normalised Ogive at this given frequency, above which the corresponding *AI* is considered too short and a flag '1' raised. However, this flag is not used for the calculation of the overall data quality.

Step 14: Calculation of overall data quality flags

First the indicators of the quality-control tests applied to the covariances in Step 8 are combined in a single flag using the flagging system based on Mauder and Foken (2006). A system of nine flags is developed for both tests corresponding to as many ranges of *FS* and ITC_σ, and then combined in a 0-1-2 flag system: if both tests yield a value lower than 30% the flag is '0', if at least one of them is between 30% and 100% the flag is '1', and if at least one is above 100% the flag is '2'. All the flags produced as described in this manuscript are combined to decide whether or not the flux value for an *AI* will be discarded. If at least one of the flags is '2', the corresponding flux value is discarded. If four or more of the other tests have a flag '1' the *AI* is discarded as well. In Table 4 we report a summary of all the tests performed and used in the overall QC. From that list we intentionally excluded the following indicators, which are not part of the overall data quality control:

1. signal strength information from IRGA;
2. flag indicating periods with methods for spectral correction, coordinate rotation and time-lag compensation different from the standard ones (lack of data for the pre-processing);
3. Ogives test.

Post-processing stages

In the following a synthetic description of the ICOS treatment applied to the processed raw-data is presented. This section is organised in sub-sections called Stages to avoid confusion with the Steps above. It starts from the description of the assessment of data uncertainty, includes

Table 4. Tests used in the overall quality estimation

TEST	DATA	FLAGS
Completeness of the AI	raw	0, 1, 2
Absolute limits	raw	0, 1, 2
Spike detection	raw	0, 1, 2
Wind direction	raw	0, 1, 2
Amplitude resolution	raw	0, 2
Drop-out	raw	0, 1
Instruments diagnostics	raw	0, 2
Steady-state + ITC	covariances	0, 1, 2
Spectral correction factor	fluxes	0, 1, 2

the organisation of the Near Real Time (NRT) processing applied daily to the raw data, and finally lists the actions needed to filter out low quality periods or suspicious data points, to account for non-turbulent fluxes and to fill the gaps created by filtering and missing data. The footprint analysis is introduced. The description of the post-processing in this section is general, but it has the aim of clarifying that after the turbulent flux calculation a number of additional and crucial steps are needed in order to calculate the final NEE estimate.

Stage 1: Uncertainty estimation

Due to the complexity of the method, the overall uncertainty of the EC technique derives from a diverse pattern of sources of errors (Richardson *et al.*, 2012): from the instrument-related measuring errors to the non-occurrence of the conditions at the basis of the method, from the random error due to one-point sampling to the uncertainty introduced by the processing chain itself. Also the filtering of periods with too low turbulence (Stage 5) is responsible for a fraction of uncertainty. The estimation of uncertainty is under discussion in the EC community (Vickers and Mahrt, 1997; Finkelstein and Sims, 2001; Hollinger and Richardson, 2005; Dragoni *et al.*, 2007; Mauder *et al.*, 2013). Here we describe the estimation of random uncertainty and that introduced by the choice of the data processing scheme. Finally we mention how the different approaches used in the post-processing (namely the u_* thresholds calculation and the partitioning methods) depict the overall variability of the final fluxes.

Stage 1.1: Flux random uncertainty

The flux random uncertainty due to the limitations of sampling, also known as turbulence sampling error (*TSE*), is estimated following the approach by Finkelstein and Sims (2001). The Integral Turbulence time-Scale (*ITS*) is derived from the integral of the auto-covariance function of the time-series $\varphi = w's'$ normalised by the variance of φ (Rannik *et al.,* 2016), where φ contains the instantaneous covariance values between w and s measured over *AI*. Then the random uncertainty ($\varepsilon_{F,rand}$) is estimated, based on the calculation of the variance of covariance:

$$\varepsilon_{F,TSE} = \frac{1}{N}\left(\sum_{j=-m}^{m} \hat{\gamma}_{w,w}(j)\hat{\gamma}_{s,s}(j) + \sum_{j=-m}^{m} \hat{\gamma}_{w,s}(j)\hat{\gamma}_{s,w}(j)\right), \quad (47)$$

where: N is the number of raw measurements in the *AI*, represents a number of samples large enough to capture the integral timescale, calculated as $ITS*f_s$, and $\hat{\gamma}_{w,w}$ and $\hat{\gamma}_{w,s}$ are auto-covariance and cross-covariance terms for atmospheric fluxes, which can be estimated for lag h as:

$$\hat{\gamma}_{w,w}(h) = \hat{\gamma}_{w,w}(-h) = \frac{1}{N}\sum_{t=1}^{N-h}(w(t) - \bar{w})(w(t+h) - \bar{w}), \quad (48)$$

$$\hat{\gamma}_{w,s}(h) = \hat{\gamma}_{s,w}(-h) = \frac{1}{N}\sum_{t=1}^{N-h}(w(t) - \bar{w})(s(t+h) - \bar{s}). \quad (49)$$

In the ICOS processing h is set to 200 s, as recommended by Rannik *et al.* (2016).

Stage 1.2: Uncertainty in the half-hourly turbulent fluxes due to the processing scheme

The processing options used to calculate the fluxes contribute to the overall uncertainty. At present there is no a widely agreed strategy conceived to calculate this uncertainty in the scientific community, but few studies exist (Kroon *et al.*, 2010; Nordbo *et al.*, 2012; Richardson *et al.*, 2012; Mauder *et al.*, 2013; Mammarella *et al.*, 2016).

The main processing scheme proposed currently for ICOS ecosystems and described in the present manuscript aims at a maximum degree of standardisation. However, as shown above, other processing options might also be valid and reliable. This leads to an uncertainty due to the processing scheme. For evaluating this portion of the uncertainty, some alternative processing methods are proposed in a factorial approach together with the methods used in the main processing chain, which are assumed to contribute the most to the variability in the flux estimation, and which are equally reliable for the ICOS setup: the *2D* coordinate rotation and the linear de-trending (*LD*) approach. Hence, the processing scheme described is replicated four times, corresponding to the factorial combination of the suggested options (*BA/LD* combined with *2D/PF*). This scheme does not apply in the pre-processing, neither in the NRT processing scheme (Stage 2).

For each flux of each *AI*, the factorial combination results in four quantities of calculated fluxes. Since there are not tools to establish a priori which is the combination of processing options providing unbiased flux estimates, we assume that the "true unobserved" flux quantity is equally likely to fall anywhere between the upper and lower limit derived from the 4 processing schemes. This translates in assuming that the PDF of flux estimated by a multiple processing scheme is better approximated by a continuous uniform distribution (BIPM *et al.*, 2008). Hence, the mean value is calculates as:

$$F_{s,mean} = \frac{max(F_{s,j}) + \min(F_{s,j})}{2}, \quad (50)$$

and the uncertainty is quantified as:

$$\varepsilon_{F,proc} = \frac{max(F_{s,j}) - \min(F_{s,j})}{\sqrt{12}}. \quad (51)$$

F_s denotes the flux of a generic scalar s, while max ($F_{s,j}$) and min ($F_{s,j}$) are the maximum and minimum flux values respectively among those calculated in j=1,…,4 processing schemes.

$\varepsilon_{F,proc}$ represents the between-flux variability due to different processing schemes. However, different processing schemes lead also to different values of $\varepsilon_{F,TSE}$, whose variability needs to be taken into account for a proper quantification of the random uncertainty affecting half-hourly flux time series. To this aim, the *TSE* associated with the half-hourly $F_{s,mean}$ is estimated by averaging the four values of $\varepsilon_{F,TSE,j}$ estimated for each processing scheme:

$$\overline{\varepsilon_{F,TSE}} = \frac{\sqrt{\sum_{j=1}^{4} \varepsilon_{F,TSE,j}{}^2}}{4}. \quad (52)$$

Stage 1.3: Combination and expanded random uncertainty

The combined random uncertainty $\varepsilon_{F,comb}$ is then obtained by combining $\overline{\varepsilon_{F,TSE}}$ and $\varepsilon_{F,proc}$ *via* summation in quadrature:

$$\varepsilon_{F,comb} = \sqrt{\overline{\varepsilon_{F,rand}}^2 + \varepsilon_{F,proc}{}^2}. \quad (53)$$

The expanded uncertainty (see also BIPM *et al.*, 2008) is achieved by multiplying the combined random uncertainty with a coverage factor of 1.96 in order to define the 95% confidence interval of the true unobserved flux estimates as follows:

$$CI = F_{s,mean} \pm 1.96\varepsilon_{F,comb}. \quad (54)$$

The magnitude of the confidence interval is then proportional to the importance of the processing scheme in the flux estimates uncertainty.

Stage 2: Near-real time (NRT) data processing

Near-real time (NRT) data processing is an important tool for ICOS, as it allows updating fluxes on a daily basis for real time visualisation of the fluxes and possible ingestion in modelling activities. It also enables ETC to send warnings to principal investigators (PIs) in case of errors or failures of the system. A prompt reaction is demanded to the PI so that problems at the station can be solved swiftly, and the high-quality standard of data maintained. The NRT processing is applied on a moving window of 10 days. In order to achieve a high quality standard, the same processing chain as for the final data (main processing) is applied as far as possible. However, certain options are not applicable to NRT processing, as detailed below. This translates in applying Steps 1 – 14 of the main processing, with the exclusion of Step 4 and some differences in Steps 3, 5, 10 (Figs 2-3). Step 4 corresponds to the correction of concentration/mole fraction drift of the IRGA, which is not included in the NRT processing scheme, as it uses data from calibrations that may only become available afterwards. Differences in Steps 3, 5 and 10 occur due to the need for a pre-processing to calculate parameters used in these corrections. The corresponding parameters are calculated once on a given dataset of two months, and then used in the daily runs for flux calculation. It is therefore important to keep the distance between the measurement height and the displacement height as constant as possible, by dynamically move the system according to the height increment in stations with fast-growing vegetation.

A different method is used when the two-month dataset is not available. This applies, in addition to the beginning of the measurement period, every time a relevant change occurs in the instrument setup, or the ecosystem characteristics become different from those used in the pre-processing. This translates in a temporary (possibly less accurate) flux computation.

NRT processing stops after the calculation of net ecosystem exchange (NEE) by summing up the storage term to F_{CO_2}, excluding gap-filling and filtering based on friction velocity (u_*) methods (see below). Also, the uncertainty estimation is limited to the random component, *i.e.* the factorial combination of options is not applied to estimate the uncertainty due to the selection of processing methods.

Step 1 and Step 2 are the same as in the main processing. The rotation of the coordinate system of the sonic is also the same (sector-wise planar fit, Step 3) only if the corresponding parameters, calculated on a period of at least two months, are available: these coefficients are then used for the daily processing, as long as no relevant changes occur, or as long as their calculation will be updated. When the two-month dataset is not available, the 2*D* rotation is applied until a new two-month dataset will be complete.

As stated above, Step 4 is not applied in the NRT. Step 5, *i.e.* the time lag optimisation to align the data streams of the sonic and the IRGA, is applied with the same limitations of Step 3: a two-month dataset is used once for the calculation of the statistic parameters of the method, which are then used for the daily processing. When relevant changes happen in the setup of the EC system the parameters need to be updated. In this case, and whenever a coherent two-month dataset is not available, the classic approach of covariance maximisation is applied. Step 6, Step 7, Step 8, and Step 9 are executed following the same methods as in the main processing. Step 10 is partly different in case a long enough dataset of two months is not available to estimate the low-pass filter characteristics and the spectral model parameters of the experimental high-frequency spectral corrections (Step 10.3.2). Even if in theory the size of the data window depends on the characteristics listed in Table 1, for standardisation a period of minimum two months is required. As in the main processing scheme, for fast-growing species a two-week period is deemed enough to calculate spectral parameters, and the PI has to avoid a more frequent change of the EC system configuration. When the dataset is not available, the fully analytic method (Moncrieff *et al.*, 1997) is temporarily applied, and the Step 10.4 is not needed. In any case, the sub-Steps 10.1, 10.2 and 10.3.1 are the same of the main processing scheme. Step 11, Step 12, Step 13 and Step 14 in the NRT processing scheme are identical to the main processing scheme.

Stage 3: Storage component

The NEE is calculated from the turbulent flux of CO_2 by summing it up with the fluxes arising from the change in storage below the eddy covariance instrumentation (Nicolini *et al.*, 2018). The storage change term is calculated from CO_2 and H_2O concentration measurements along

a vertical profile on the tower. If there are half hours with gaps in the CO_2 profile or if the measurement height is below 3 m, a discrete approach is used, that is based on the measurement of CO_2 content from the IRGA used for the EC. These half hours are flagged but used in the following processing steps.

Stage 4: Spike detection

The half-hourly dataset of NEE can contain spikes. These spikes are being interpreted as outliers and are identified and flagged. The method described in Papale *et al.* (2006), based on MAD approach using a z value of 5.5 as in the original implementation, is used and spikes are flagged and removed from the subsequent processing steps.

Stage 5: u_*-filtering

Low turbulence conditions need to be identified and filtered out because there is the risk of strong advective fluxes that are not detected by neither the EC nor the storage profile systems. Half-hourly data sets measured below a certain threshold value of u_* have to be filtered out (Aubinet *et al.*, 2012). The threshold value is ecosystem-specific and can be identified using automatic methods. Two different u_*-threshold selection methods are applied (Reichstein *et al.*, 2005, modified in Barr *et al.*, 2013 respectively), using a bootstrapping approach: this allows the calculation of the uncertainty in the threshold definition, and the calculation of a threshold distribution that is used to flag and filter both daytime and night-time data. This uncertainty is expected to provide useful information for future steps in the data filtering. In fact a large uncertainty in the u_* threshold warns on difficulties in detecting and filtering out the advection, on the presence of large noise in the storage flux, as well as on large variability of the footprint.

Stage 6: Gap filling of CO_2 flux data

Gaps in the half-hourly fluxes resulting from not measured data or missing results after data quality control filtering can be filled with different types of imputation procedures, parameterised on good quality measurements to build relations between drivers and fluxes, and then applied using only the drivers in the presence of gaps. For ICOS purposes, the well-consolidated method of Marginal Distribution Sampling (MDS, Reichstein *et al.*, 2005) is used. Missing values in half-hourly flux time series are imputed according to an average calculated in a moving window of 7-14 days of observed values under similar meteorological conditions. Similarity is defined by a deviation in short-wave incoming radiation, T_a and vapour pressure deficit lower than 50 W m^{-2}, 2.5°C and 5 hPa, respectively. This approach is deemed suitable to consider the different regimes of crop phenology. In case of ecosystems where clearly distinguished functional periods are present due to human activities (management) or rapid naturally induced changes like in case of fires, the MDS is parameterised individually for the different periods (Béziat *et al.*, 2009).

Stage 7: Overall uncertainty quantification

Uncertainty in NEE is estimated creating an ensemble of the processing step combinations described above, *i.e.* uncertainty calculated from the post-field raw data processing (Stage 1.2) and from u_* threshold calculation and application (Stage 5). Percentiles describing the distribution for each data point are then extracted. Total uncertainty is calculated by propagating the single uncertainties.

A consistency indicator is also applied to detect issues in the measurements. It is derived from the calculation of the two main components of NEE (gross primary production, GPP; and total ecosystem respiration, TER) with two independent models parameterized using night-time and day-time data following Reichstein *et al.*, 2005 and Lasslop *et al.*, 2010, respectively. The indicator is based in the difference between the results of these two models. The parameterisation of the models using independent (night- and day- time) data indicates the occurrence of advective phenomena and decoupling, or of relevant changes in the footprint, which will lead to large differences in the results of the two methods.

Stage 8: Footprint analysis

The source area of the measured fluxes in relation to its spatial extent and distribution of intensity can be represented by the footprint function Φ, defined as the integral equation of diffusion (Wilson and Swaters, 1991; Pasquill and Smith, 1983):

$$\eta = \int_R \Phi(\vec{x}, \vec{x'}) Q(\vec{x'}) d\vec{x'}, \tag{55}$$

where η is the signal being measured at location $\vec{x}$ ($\vec{x}$ is a vector), and $Q(\vec{x})$ is the source emission rate/sink strength in the surface-vegetation volume R. The footprint is sensitive to atmospheric stability and surface roughness (Leclerc and Thurtell, 1990), in addition to measurement height and wind direction. The percentage-level for the source area is used to represent the origin of the fluxes (Schmid, 1994).

Several theoretical approaches exist for the determination of Φ (Kormann and Meixner, 2001; Hsieh *et al.*, 2000; Kljun *et al.*, 2004), that require in input the mean wind speed, the lateral wind speed standard deviation, the Obukhov length, the displacement height and the roughness length z_0, corresponding to the height above the displacement height where the wind speed is actually zero, and representing the roughness of the vegetation elements (z_0=0.15 h_c; can be calculated if wind profile is measured).

The parameterization by Kljun *et al.*, 2015 is adopted in ICOS, which uses an optimal combination of different existing methods. For a more detailed description we refer the reader to the cited papers.

Overview of the dataset obtained with the described corrections

At the end of the main processing chain and of the post-processing procedures, the dataset is composed of twelve months of half-hourly data, including fluxes of CO_2, latent heat, momentum and sensible heat. In addition, flux of evapotranspiration is also calculated, as well as friction velocity u_* and several atmospheric parameters and statistics. NEE includes the calculate change in storage, while gaps deriving from technical field problems as well as half hours filtered out for bad quality in Step 14 or after the definition of the u_* threshold are filled as described above and flagged accordingly. In addition, from the NRT processing every day we have a 48-half-hour dataset of fluxes, meteorological parameters and statistics relative to the day before, and calculated over a window of 10 days, publicly visible on the Carbon Portal website. Both the datasets (main and NRT) are open access, and can be cited in publications thanks to a dataset identifier system. In addition, also the raw data gets an identifier, as well as all the companion variables and the metadata sent by the PIs.

RESULTS AND DISCUSSION

The EC main processing chain described in the previous section has been selected in ICOS as the standard for the post-field elaboration of raw data to calculate high-quality, corrected fluxes. Even if using the same instrumentation, ICOS stations are characterised by different ecosystems, measurement heights, topography, and climatic conditions. In addition, the instrumentation ages due to the usage, and this must be recognised by automated data screening. The selected methods have to be applicable to different situations, and not specific for any special case. The experimental method for the spectral corrections, for example, or the planar fit method to rotate the coordinate frame of the sonic anemometer, are options suited to be flexible enough to cover a wide range of situations. However, numerous alternative options exist, some of which have been excluded because of a known risk that these introduce additional errors. Other options, considered as equally valid as the main ones, are used to estimate a range of uncertainty due to the choice of the methods. We discuss in the following the advantages and drawbacks of the options selected, also in comparison of some of the alternatives available. Moreover, we explain why some very well-known steps seem to be missing in the main processing chain.

The main risk of the quality control is the possible false flagging of good-quality values instead of flawed/wrong data. For that reason some of the tests proposed by Vickers and Mahrt (1997) are applied but not used for the quality control of ICOS. In particular, skewness and kurtosis are calculated and stored; a discontinuity test and the test on non-stationarity of the horizontal wind are applied.

SATs calculate sonic temperature based on three paths, and thus crosswind has to be considered (Liu *et al.*, 2001). Gill HS 50 and HS 100 include this correction in their embedded software, so it has not to be considered in the processing of data.

SAT support arms and transducers constitute a physical obstacle to the wind flowing through them, thus creating a disturbance to the vertical wind speed (w) that could lead to an underestimation of the fluxes. At present, there is no wide agreement in the scientific community on the application of the existing algorithms aiming at correcting this flow distortion error (aka angle of attack error; *e.g.* Van der Molen *et al.*, 2004; Nakai *et al.*, 2006, 2012). The Gill HS-50 and HS-100, yoke-mounted models, are less sensitive to the issue because no spars are present in the vicinity of the sampling volume, but the wind sector where the arm of the instrument lies has to be excluded from the raw dataset. However, transducer geometry is still an issue to be solved. Therefore, at present we do not perform any correction for the angle of attack, but this may change in future depending on new results and a higher degree of consensus within the scientific community. In this case, a retrospective application of new corrections is foreseen, as the raw data are archived along with the calculated fluxes.

Other options suitable for the coordinate rotation are the so-called *2D* (as in Aubinet *et al.*, 2000) and 3*D* (McMillen, 1988) rotation. While the latter is known to result in sometimes unphysical orientation of the vector basis (Finnigan, 2004), the 2*D* rotation is a valuable and widely adopted option. The 2*D* method is suitable in most cases, but, for stations having complex topography, the sector-wise *PF* method has better performances (Aubinet *et al.*, 2012), while the original *PF* method is not differentiated per wind sector. The *PF* method on the other hand needs the pre-processing described above; also it nullifies only the long-term mean vertical wind speed, while the half-hourly may be different from zero. Therefore the 2*D* method has been selected as an alternative method to be performed for the calculation of the uncertainty in combination with the *PF* method.

At present, we are not aware of alternative methods for the correction of the mole fraction drift of the IRGA. A drawback of the method proposed is the assumption of a linear drift between two consecutive calibrations. The error due to this assumption depends on the magnitude of the non-linearity of the drift. However, the frequency of the field calibration of the IRGA in ICOS (minimum 4 times in the first year, and once per year in the following) should reduce the likelihood of occurrence of non-linearity, together with the mandatory maintenance routine.

Block averaging as a method for separating fluctuations from its background signal might lead to biased estimation of the fluctuations if any trend exists in the time series, due to the lack of the condition of stationarity in the time series (drifts in the instrumentation, slow changes in the atmospheric conditions in the average time). This could be solved using one of the two alternative methods, namely the linear detrending (*LD*, Gash and Culf, 1996) or the autoregressive filtering (*AF*, Moore, 1986; McMillen, 1988). The three methods have been compared by Rannik and Vesala, 1999: while solving the trend issue, the two alternative methods do not obey Reynolds' averaging rules. *AF* also has a strong impact in terms of the spectral attenuation, and for both these reasons we decided to avoid it. The *LD* could be a good compromise when some trends exist in the *AI*, and for that reason has been chosen in ICOS as a valid method for the calculation of the uncertainty. It is also worth mentioning that Richardson *et al.*, 2012 report that *LD* and *AF* reduce the random error as compared to *BA* when combined with a proper high-pass filter for the spectral correction. However, *AF* method needs a proper calculation of the filter constant, and the *LD* should be applied only when a trend is present. This introduces a subjective component in the choice of the method, and for that reason in the main (standardised) processing chain we selected *BA*. The combination with *LD* in the uncertainty calculation is conceived also to include these differences in the variability range of the calculated fluxes.

Like most of the other methods selected in ICOS, also the option of time lag optimisation has the advantage of providing a certain degree of adaptation to the local conditions, because the searching lag is differentiated and selected according to the *RH* values. On the other hand, the drawback of this option is also the need for a pre-processing, which increases the computational demands. The more important alternative method is the maximisation of the covariance using fixed, predetermined windows, which can provide flawed time lags for H_2O (Taipale *et al.*, 2010; Mammarella *et al.*, 2016).

Fluctuations of air density (related to fluctuations of ambient or cell temperature, pressure and water vapour content) would entail unphysical fluctuations of gas concentration/mole fraction, while the dry mole fraction is instead a conserved quantity in the presence of changes in these ambient/cell conditions (Kowalski and Serrano Ortiz , 2007). The IRGA LI-7200 outputs dry mole fraction, hence the fluxes can be calculated using the reported processing chain. Otherwise, an additional step would have been needed to compensate *a posteriori* for air density fluctuations (aka as WPL term, Webb *et al.*, 1980; Ibrom *et al.*, 2007b).

The spectral correction option at the low-frequency range is *per se* a standard method, and no relevant alternatives are known. A different filter function has to be selected when the *LD* method is applied for the calculation of the fluctuations. However, some concerns exist in a part of the EC community in using a model (theoretical) (co) spectrum: the risk is that the spectral models, *e.g.* the spectra shape and the position of the peak frequency in respect to atmospheric stability, do not fit the data well. Efforts are needed to find an alternative solution which makes use of a reference (co)spectrum calculated *in-situ* as in the proposed method for the high-frequency range for gas scalars.

Several alternatives are possible for the high-frequency correction of gas scalars, which can be gathered in two groups: analytic and experimental methods. For example, in addition to the methods used in ICOS, the method by Massman (2000) is analytic, while Horst (1997) can be classified as experimental as it is parameterised with *in situ* information. As usual, one of the experimental methods has been selected in ICOS for gas scalars due to its flexibility, as using the *in-situ* spectra of *T* as ideally un-attenuated allows adaptation to local conditions. The analytic alternatives are not suitable for the ICOS setup (closed-path IRGA), and the other experimental methods are less accurate, especially in terms of dependence on *RH*. However, as with previous steps, the need of a two months dataset where to calculate the spectral parameters during the pre-processing is a shortcoming of this method.

For some of the above listed corrections some other alternative options exist, which are so far less used in the EC community and need further testing before being adopted in ICOS. This applies *e.g.* to the quality control of EC data: several alternatives to the tests of Vickers and Mahrt (1997) exist, based on different approaches, which can be combined with the currently applied tests on the raw data, and with the tests on the covariances (Foken and Wichura, 1996) or on the calculated fluxes. Other examples are the alternatives to the Fourier-based spectral corrections, mainly based on the wavelet decomposition (Nordbo and Katul, 2013). Future methodological development is then expected to take place in this part of the processing, possibly leading to significant improvements.

The uncertainty of the EC method has several sources. In both the main and NRT ICOS data processing schemes we quantify the random uncertainty according to the well-defined method of Finkelstein and Sims (2001). In addition, in the main processing chain we also determine the range of uncertainty due to the most important causes of flux variability, *i.e.* the selection of processing options, the selection of the u_* thresholds and the partitioning method used. The variability of the fluxes calculated with different processing options is based on the factorial combination of four options involving two steps, namely the coordinate rotation and the trend removal. These are expected to be the more important sources of variability, however some residual variability can also come from the corrections we decided to exclude for computational reasons (the full factorial of all the options available would mean hundreds of processing runs per station). The u_* threshold selection, while being a robust method for the assessment of low-turbulence periods, is prone to some subjectivity. Using two

different methods and bootstrapping the thresholds lead to consider the big variability in the fluxes due to the selection of different, plausible thresholds. Finally, the two alternative models used to parameterise the partitioning of NEE into GPP and TER use respectively daytime and night-time data, and are for that independent. The comparison of the two resulting time series gives an indication of the consistency and uncertainty.

CONCLUSIONS

1. The eddy covariance processing chain is composed of fourteen steps, and is standardised for all the Integrated Carbon Observation System ecosystem stations.

2. Standardisation of procedures in an infrastructure made of >70 ecosystem stations in the whole Europe is an important requirement to avoid local methodological biases when comparing the C and GHG budgets at the continental scale. With this standardised approach, ICOS aims at maximising the comparability, the reliability and the methodological transparency and traceability of the final estimated flux data and other variables, including the quantitative estimation of their quality and uncertainty.

3. The EC processing chain was developed with the contribution of different experts in the field of EC technique at global level and members of the Ecosystem Monitoring Station Assembly (MSA) of ICOS.

4. The processing chain is considered the currently best possible compromise between standardisation and adaptation of the state of the art methods to the specific conditions of each station. It is clear that it will evolve in time when new findings and methods will be available.

5. Random and processing uncertainties are also estimated. In particular, the method to determine uncertainty due to the processing is described, adding a quantification of the station-specific variability in the results. The estimation of the uncertainty evolving from the processing scheme is based on factorial combination of alternative processing techniques.

6. The centralisation of the processing at the ICOS ETC warrants the standardisation, adding at the same time value to the work made by the station team in establishing the stations and collecting the data. The data streams are dynamically linked to the metadata of the stations, resulting in an organic operational system. Central automated error detection and continuous communication with the station's team helps fast identification and elimination of measurement problems.

ACKNOWLEDGMENTS

The authors wish to thank the ICOS working group on EC Processing for the enthusiastic contribution in drafting the Protocol. This work is a step in the direction of the standardization and S.S. thanks the COOP+ H2020 EU project (Project Number 654131, March 2016-July 2018) for the support. L.H. and L.M. acknowledge funds received under the Swiss National Science Foundation funded project ICOS-CH (grant agreements: 20FI21_148992, July 2013 to June 2017; 20FI21_173691, since July 2017) and of ETH Zurich (0-23184-15, June 2015 to Dec. 2017). N.A. thanks support from ICOS-Belgium and Fonds Wetenschappelijke Onderzoek (FWO). L.Š. was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the CzeCOS program, grant number LM2015061, and within the National Sustainability Program I (NPU I), grant number LO1415 (2015-2019). The National Ecological Observatory Network is a project sponsored by the National Science Foundation and managed under cooperative agreement by Battelle Ecology, Inc. This material is based upon work supported by the National Science Foundation (grant DBI-0752017, 2008-2018). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation and European Union.

Conflict of interest: The Authors declare no conflict of interest.

REFERENCES

Aubinet M., Grelle A., Ibrom A., *et al.*, 2000. Estimates of the annual net carbon and water exchange of forests: the EUROFLUX methodology. Adv Ecol Res., 30, 113-175.

Aubinet M., Vesala T., and Papale D., 2012. Eddy covariance, a practical guide to measurement and data analysis, Springer.

Baldocchi D.D., Hincks B.B., and Meyers T.P., 1988. Measuring biosphere-atmosphere exchanges of biologically related gases with micrometeorological methods. Ecology, 69(5), 1331-1340.

Baldocchi D.D., 2003. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. Glob. Change Biol., 9(4), 479-492.

Barr A.G., Richardson A.D., Hollinger D.Y., *et al.*, 2013. Use of change-point detection for friction-velocity threshold evaluation in eddy-covariance studies. Agr. Forest Meteorol., 171, 31-45.

Berger B.W., Davis K.J., Yi C., Bakwin P.S., and Zhao C.L., 2001. Long-term carbon dioxide fluxes from a very tall tower in a northern forest: Flux measurement methodology. J. Atmos. Ocean Tech., 18(4), 529-542.

Béziat P., Ceschia E., and Dedieu G., 2009. Carbon balance of a three crop succession over two cropland sites in South West France. Agr. Forest Meteorol., 149(10), 1628-1645.

BIPM, IEC, IFCC, ILAC, IUPAC, IUPAP, ISO and OIML, **2008.** Evaluation of measurement data-guide for the expression of uncertainty in measurement. JCGM 100. Available at: http://www.bipm.org/en/publications/guides/gum.html

Clement R., 2004. Mass and energy exchange of a plantation forest in Scotland using micrometeorological methods. IAES. Edinburgh, The University of Edinburgh: 597.

Dragoni D., Schmid H.P., Grimmond C.S.B., and Loescher H.W., 2007. Uncertainty of annual net ecosystem productivity estimated using eddy covariance flux measurements. J. Geophys. Res.-Atmos., 112(D17).

Finkelstein P.L. and Sims P.F., 2001. Sampling error in eddy correlation flux measurements. J. Geophys. Res.-Atmos., 106, 3503-3509.

Finnigan J.J., 2004. A re-evaluation of long-term flux measurement techniques part II: coordinate systems. Boundary-Layer Meteorology, 113(1), 1-41.

Finnigan J.J., Clement R., Malhi Y., Leuning R., and Cleugh H.A., 2003. A re-evaluation of long-term flux measurement techniques - Part I: Averaging and coordinate rotation. Bound-Lay. Meteorol., 107, 1-48.

Foken T. and Wichura B., 1996. Tools for quality assessment of surface-based flux measurements. Agr. Forest Meteorol., 78: 83-105.

Fratini G., Ibrom A., Arriga N., Burba G., and Papale D., 2012. Relative humidity effects on water vapour fluxes measured with closed-path eddy-covariance systems with short sampling lines. Agr. Forest Meteorol., 165, 53-63.

Fratini G., McDermitt D.K., and Papale D., 2014. Eddy-covariance flux errors due to biases in gas concentration measurements: origins, quantification and correction. Biogeosciences, 11, 1037-1051.

Gash J.H.C. and Culf A.D., 1996. Applying a linear detrend to eddy correlation data in realtime. Bound-Lay. Meteorol. , 79(3), 301-306.

Hollinger D.Y. and Richardson A.D., 2005. Uncertainty in eddy covariance measurements and its application to physiological models. Tree Physiol., 25(7), 873-885.

Horst T.W., 1997. A simple formula for attenuation of eddy fluxes measured with first-order-response scalar sensors. Boundary-Layer Meteorology, 82(2), 219-233.

Horst T.W. and Lenschow D.H., 2009. Attenuation of scalar fluxes measured with spatially-displaced sensors. Bound-Lay. Meteorol., 130(2), 275-300.

Hsieh C.I., Katul G., and Chi T.W., 2000. An approximate analytical model for footprint estimation of scalar fluxes in thermally stratified atmospheric flows. Adv. Water Resour., 23(7), 765-772.

Ibrom A., Dellwik E, Flyvbjerg H., Jensen N.O., and Pilegaard K., 2007a. Strong low-pass filtering effects on water vapour flux measurements with closed-path eddy correlation systems. Agr. Forest Meteorol., 147: 140-156.

Ibrom A., Dellwik E., Larsen S.E., and Pilegaard K., 2007b. On the use of the Webb - Pearman - Leuning - theory for closed-path eddy correlation measurements. Tellus B, 59B: 937-946.

Kaimal J.C. and Finnigan J.J., 1994. Atmospheric Boundary Layer Flows. Their Structure and Measurement. Oxford University Press.

Kaimal J.C. and Kristensen L., 1991. Time series tapering for short data samples. Bound-Lay. Meteorol., 57(1-2), 187-194.

Kaimal J.C., Wyngaard J., Izumi Y., and Coté O.R., 1972. Spectral characteristics of surface-layer turbulence (No. AFCRL-72-0492). Air Force Cambridge Research Labs Hanscom Afb Ma.

Kljun N., Calanca P., Rotach M.W., and Schmid H.P., 2004. A simple parameterisation for flux footprint predictions. Bound-Lay. Meteorol., 112(3), 503-523.

Kljun N., Calanca P., Rotach M.V., and Schmid H.P., 2015. A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP). Geosci. Model Dev., 8, 3695-3713.

Kormann R. and Meixner F.X., 2001. An analytical footprint model for non-neutral stratification. Bound-Lay. Meteorol., 99(2), 207-224.

Kowalski A.S. and Serrano-Ortiz P., 2007. On the relationship between the eddy covariance, the turbulent flux, and surface exchange for a trace gas such as CO_2. Bound-Lay. Meteorol., 124(2), 129-141.

Kroon P.S., Hensen A., Jonker H.J.J., Ouwersloot H.G., Vermeulen A.T., and Bosveld F.C., 2010. Uncertainties in eddy covariance flux measurements assessed from CH_4 and N_2O observations. Agr. Forest Meteorol. 150: 806-816.

Lasslop G., Reichstein M., Papale D., *et al.*, 2010. Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. Glob. Change Biol., 16(1), 187-208.

Leclerc M.Y. and Thurtell G.W., 1990. Footprint prediction of scalar fluxes using a Markovian analysis. Bound-Lay. Meteorol., 52(3), 247-258.

Lee X., Massman W., and Law B. (Eds), 2004. Handbook of micrometeorology: a guide for surface flux measurement and analysis (Vol. 29). Kluwer Academic Publishers, Dordrecht, the Netherlands.

Liu H., Peters G., and Foken T., 2001. New equations for sonic temperature variance and buoyancy heat flux with an omnidirectional sonic anemometer. Bound-Lay. Meteorol., 100(3), 459-468.

Mammarella I., Kolari P., Vesala T., and Rinne J., 2007. Determining the contribution of vertical advection to the net ecosystem exchange at Hyytiälä forest, Finland. Tellus 59B, 900-909, DOI: 101111/j.1600-0889.2007.00306.x

Mammarella I., Launiainen S., Grönholm T., Keronen P., Pumpanen J., Rannik Ü., and Vesala T., 2009. Relative humidity effect on the high-frequency attenuation of water vapour flux measured by a closed-path eddy covariance system. J.Atmos. Ocean.Technol., 26: 1856-1866.

Mammarella I., Peltola O., Nordbo A., Järvi L., and Rannik Ü., 2016. Quantifying the uncertainty of eddy covariance fluxes due to the use of different software packages and combinations of processing steps in two contrasting ecosystems, Atmos. Meas. Tech., 9, 4915-4933.

Mauder M., Cuntz M., Drüe C., *et al.*, 2013. A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements. Agr. Forest Meteorol., 169, 122-135.

Massman W.J. and Ibrom A., 2008. Attenuation of concentration fluctuations of water vapour and other trace gases in turbulent tube flow. Atmos. Chem. Phys., 8: 6245-6259.

Massman W.J., 2000. A simple method for estimating frequency response corrections for eddy covariance systems. Agr. Forest Meteorology, 104: 185-198.

Mauder M. and Foken T., 2006. Impact of post-field data processing on eddy covariance flux estimates and energy balance closure. Meteorologische Zeitschrift, 15(6), 597-609.

Mauder M., Oncley S.P., Vogt R., *et al.*, 2007. The energy balance experiment EBEX-2000. Part II: Intercomparison of eddy-covariance sensors and post-field data processing methods. Bound-Lay. Meteorol., 123(1), 29-54.

McMillen R.T., 1988. An Eddy-correlation Technique With Extended Applicability to Non-simple Terrain. Bound-Lay. Meteorol., 43: 231-245.

Moncrieff J., Clement R., Finnigan J., and Meyers T., 2004. Averaging, detrending, and filtering of eddy covariance time series. In: Handbook of Micrometeorology. Springer, Netherlands.

Moncrieff J.B., Massheder J.M., De Bruin H., *et al.*, 1997. A system to measure surface fluxes of momentum, sensible heat, water vapour and carbon dioxide. J. Hydrol., 188, 589-611.

Moore C.J., 1986. Frequency response corrections for eddy correlation systems. Bound-Lay. Meteorol., 37(1-2), 17-35.

Nakai T. and Shimoyama K., 2012. Ultrasonic anemometer angle of attack errors under turbulent conditions. Agr. Forest Meteorol., 162, 14-26.

Nakai T., Van Der Molen M.K., Gash J.H.C., and Kodama Y., 2006. Correction of sonic anemometer angle of attack errors. Agr. Forest Meteorol., 136(1), 19-30.

Nicolini G., Aubinet M., Feigenwinter C., *et al.*, 2018. Impact of CO_2 storage flux sampling uncertainty on net ecosystem exchange measured by eddy covariance. Agr. Forest Meteorol., 248, 228-239.

Nordbo A., Järvi L., and Vesala T., 2012. Revised eddy covariance flux calculation methodologies - effect on urban energy balance. Tellus Ser. B-Chem. Phys. Meteorol., 64: 18184.

Nordbo A. and Katul G., 2013. A wavelet-based correction method for eddy-covariance high-frequency losses in scalar concentration measurements. Bound-Lay. Meteorol., 146(1), 81-102.

Nordbo A., Kekäläinen P., Siivola E., Mammarella I., Timonen J. and Vesala T., 2014. Sorption-caused attenuation and delay of water vapor signals in eddy covariance sampling tubes and filters. J. Atmos. Oceanic Technol., 31, 2629-2649.

Papale D., Reichstein M., Aubinet M., *et al.*, 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences, 3(4), 571-583.

Pasquill F. and Smith F.B., 1983. Atmospheric diffusion.: Study of the dispersion of windborne material from industrial and other sources. Wiley Press, New York, USA.

Rannik Ü., Peltola O., and Mammarella I., 2016. Random uncertainties of flux measurements by the eddy covariance technique, Atmos. Meas. Tech., 9, 5163-5181, doi:10.5194/amt-9-5163-2016.

Rannik Ü. and Vesala T., 1999. Autoregressive filtering versus linear detrending in estimation of fluxes by the eddy covariance method. Bound-Lay. Meteorol., 91: 259-280.

Reichstein M., Falge E., Baldocchi D.D., and Papale D., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob. Change Biol., 11:1-16.

Richardson A.D., Aubinet M., Barr A.G., Hollinger D.Y., Ibrom A., Lasslop G., and Reichstein M., 2012. Uncertainty Quantification. In: Eddy Covariance: A Practical Guide to Measurement and Data Analysis (Eds M. Aubinet, T. Vesala, D. Papale). Springer Atmospheric Sciences, Dordrecht, The Netherlands.

Schmid H.P., 1994. Source areas for scalars and scalar fluxes. Bound-Lay. Meteorol., 67(3), 293-318.

Schotanus P., Nieuwstadt F., and De Bruin H.A.R., 1983. Temperature measurement with a sonic anemometer and its application to heat and moisture fluxes. Bound-Lay. Meteorol., 26(1), 81-93.

Smith S.W., 1997. The scientist and engineer's guide to digital signal processing. California Technical Pub., San Diego CA, USA.

Spirig C., Neftel A., Ammann C., *et al.*, 2005. Eddy covariance flux measurements of biogenic VOCs during ECHO 2003 using proton transfer reaction mass spectrometry. Atmos. Chem. Phys., 5(2), 465-481.

Stull R.B., 1988. An Introduction To Boundary Layer Meteorology. Kluwer Academic Publishers.

Taipale R., Ruuskanen T.M., and Rinne J., 2010. Lag time determination in DEC measurements with PTR-MS, Atmos. Meas. Tech., 3, 853-862.

Taylor G.I., 1938. The spectrum of turbulence. Proc. R. Soc. Lon. Ser-A, 164(919), 476-490.

Thomas C. and Foken T., 2002. Re-evaluation of integral turbulence characteristics and their parameterisations. 15th Symp. Boundary Layers and Turbulence. Am. Meteorol. Soc., July 14-19, Wageningen, The Netherlands.

Van der Molen M.K., Gash J.H.C., and Elbers J.A., 2004. Sonic anemometer (co) sine response and flux measurement: II. The effect of introducing an angle of attack dependent calibration. Agr. Forest Meteorol., 122(1), 95-109.

Van Dijk A., Moene A.F. and De Bruin H.A.R., 2004. The principles of surface flux physics: theory, practice and description of the ECPACK library. University of Wageningen, Wageningen.

Vickers D. and Mahrt L., 1997. Quality control and flux sampling problems for tower and aircraft data. J. Atmos. Ocean Tech., 14: 512-526.

Webb E.K., Pearman G.I., and Leuning R., 1980. Correction of flux measurements for density effects due to heat and water vapour transfer. Q. J. Roy. Meteor. Soc., 106(447), 85-100.

Wilczak J.M., Oncley S.P., and Stage S.A., 2001. Sonic anemometer tilt correction algorithms. Bound.-Lay. Meteorol., 99, 127-150.

Wilson J.D. and Swaters G.E., 1991. The source area influencing a measurement in the planetary boundary layer: The "footprint" and the "distribution of contact distance". Bound-Lay. Meteorol., 55(1-2), 25-46..

Appendix A – symbols, abbreviations and acronyms used in the text

For the sake of clarity, here follows a list of symbols and variables as reported in the text and used in formulas of the manuscript:

s: generic scalar
s_g: gas scalar
t: time
x(t): generic time series
x: generic variable
$\vec{x}$: generic vector
x': fluctuation of x(t)
$\bar{x}$: background signal of x(t)
f: frequency
u: first horizontal component of wind speed (m s^{-1})
v: second horizontal component of wind speed (m s^{-1})
w: vertical component of wind speed (m s^{-1})
l_{AI}: length of the average interval (s)
z: measurement height (m)
z_0: roughness length (m)
d: displacement height (m)
h_c: canopy height (m)
T_s: sonic temperature (K)
T_v: virtual temperature (K)
T: absolute air temperature (K)
T_p: potential temperature (K)
T_a: air temperature (°C)
a: absorptance
$X_{g,w}$: mole fraction of generic gas g (ppm or ppt)
$X_{g,d}$: dry mole fraction of generic gas g (ppm or ppt)
$\rho_{g,m}$: mass concentration (kg m^{-3})
p: atmospheric pressure (Pa)
p_0: reference atmospheric pressure (Pa)
e: partial pressure of water vapour (Pa)
q: specific humidity (kg kg^{-1})
M: molecular mass (kg mol^{-1})
c_p: specific air heat capacity at constant pressure (J kg^{-1} K^{-1})
c_d: specific dry air heat capacity at constant pressure (J kg^{-1} K^{-1})
c_v: specific water vapour heat capacity (J kg^{-1} K^{-1})
g: gravitational acceleration (m s^{-2})
R: universal gas constant (J mol^{-1} K^{-1})
L: Obukhov length (m)
ζ: stability parameter
Og: Ogive
F: generic for flux
H: sensible heat flux (W m^{-2})
E: evapotranspiration flux (kg m^{-2} s^{-1})
LE: latent heat flux (W m^{-2})
τ: momentum flux (kg m^{-1} s^{-2})
u_*: friction velocity (m s^{-1})

v: molar volume (m^3 mol^{-1})
λ: specific latent heat of evaporation (MJ kg^{-1})
TL: time lag (s)
TF: transfer function
f_s: sampling frequency (Hz)
n: natural frequency
f_c: cut-off frequency
f_w: normalised frequency
f_p: peak frequency
S: spectrum
$CS_{Id,F}$: ideal cospectrum of F
CF: correction factor
τ_s: filter time constant
κ: von Kármán constant
l_p: SAT path length (m)
l_s: distance between centre of sonic and gas sampling paths (m)
Φ: footprint function
η: signal measured by the EC system in footprint analysis
Q: source emission or sink strength from footprint analysis
R: surface-vegetation sample volume
<>: median operator
ε: uncertainty (random or processing)

Subscripts:
v: water vapour
d: dry air
w: wet air
a: air
g: gas
corr: corrected
meas: measured
sp: spectrally corrected
0: not corrected
rand: random
proc: processing

Acronyms:
SAT: sonic anemometer thermometer
IRGA: infra-red gas analyser
EC: eddy covariance
AI: average interval
PI: Principal Investigator
ETC: Ecosystem Thematic Centre
MSA: Monitoring Station Assembly
2D: double rotation
3D: triple rotation
PF: planar fit rotation
BA: block average
RH: relative humidity
LD: linear detrending
AF: autoregressive filtering

PDF: probability density function
MAD: median absolute deviation
FS: flux stationarity
ITS: integral time scale
FFT: Fast Fourier Transform
NRT: near-real time
NEE: net ecosystem exchange of CO_2
MDS: marginal distribution sampling
GPP: gross primary production
TER: total ecosystem respiration
IIR: infinite impulse response filter
QC: quality control
TSE: turbulent sampling error