



Readiness of ICOS for Necessities of integrated Global Observations

D3.5

Protocol for non-CO₂ eddy covariance measurements, QA/QC, data processing and gap-filling



RINGO (GA no 730944)
Public Report

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730944

Deliverable: D3.5.

Author(s): Ivan Mammarella¹, Toprak Aslan¹, George Burba^{2,3}, Nick Cowan⁴, Carole Helfter⁴, Mathias Herbst⁵, Lukas Hörtnagl⁶, Andreas Ibrom⁷, Antje M. Lucas-Moffat⁵, Giacomo Nicolini⁸, Dario Papale^{8,9}, Olli Peltola¹⁰, Üllar Rannik¹, Domenico Vitale⁹, Karen Yeung⁴, Eiko Nemitz⁴

Date: 27/11/2020

Activity: RINGO WP3, Task 3.4

Lead Partner: University of Helsinki (UHEL), FI

Document Issue: 1

Dissemination Level: Public

Contact: Ivan Mammarella (ivan.mammarella@helsinki.fi)

Affiliations: ¹University of Helsinki, Helsinki, Finland; ²Environmental Division, LI-COR Biosciences, Lincoln, Nebraska, NE 68504, USA; ³Bio-Atmospheric Sciences, University of Nebraska, Lincoln, Nebraska, NE 68508, USA, ⁴UK Centre for Ecology and Hydrology, Bush Estate, Penicuik, EH26 0QB, UK; ⁵German Meteorological Service (DWD), Centre for Agrometeorological Research, Bundesallee 33, 38116 Braunschweig, Germany; ⁶Department of Environmental Systems Science, ETH Zurich, 8092 Zurich, Switzerland; ⁷Dept. Environmental Engineering, Technical University of Denmark (DTU), Lyngby, Denmark; ⁸Department for Innovation in Biological, Agro-food and Forest systems (DIBAF), University of Tuscia, Viterbo, 01100, Italy; ⁹Euro-Mediterranean Centre on Climate Change Foundation (CMCC), Lecce, 73100, Italy; ¹⁰Climate Research Programme, Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland

	Name	Partner	Date
From	Ivan Mammarella	UHEL	28.10.2020
Reviewed by	Ville Kasurinen	ICOS ERIC HO	20.11.2020
Approved by	Janne-Markus Rintala	ICOS ERIC HO	27.11.2020

Version	Date	Comments/Changes	Author/Partner

Deliverable Review Checklist

A list of checkpoints has been created to be ticked off by the Task Leader before finalizing the deliverable. These checkpoints are incorporated into the deliverable template where the Task Leader must tick off the list.

- Appearance is generally appealing and according to the RINGO template. Cover page has been updated according to the Deliverable details. ✓
- The executive summary is provided giving a short and to the point description of the deliverable. ✓
- All abbreviations are explained in a separate list. ✓
- All references are listed in a concise list. ✓
- The deliverable clearly identifies all contributions from partners and justifies the resources used. ✓

- A full spell check has been executed and is completed.

v

DISCLAIMER

This document has been produced in the context of the *project Readiness of ICOS for Necessities of integrated Global Observations (RINGO)*

The Research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730944. All Information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors view.

Amendments, comments and suggestions should be sent to the authors.

Abstract

The biosphere-atmosphere exchanges of methane (CH_4) and nitrous oxide (N_2O) are important components of carbon and nitrogen cycles of several ecosystems, like peatland, cropland, grassland, marshland and lakes. For this reason, Eddy Covariance (EC) flux measurements of these two greenhouse gases are necessary within ICOS flux tower network. Commercially available fast-response analyzers for CH_4 and N_2O have recently become more sensitive, more robust and easier to operate, giving the possibility for long term EC flux measurements. Nevertheless, the complex spatio-temporal dynamics of CH_4 and N_2O exchanges and the non-linear relationships with multiple biotic and abiotic drivers make the flux data processing and gap-filling more challenging than for CO_2 . Background fluxes of CH_4 and N_2O are often close or below the detection limit of the EC system requiring novel methodological approaches for unbiased results and long-term budgets.

In this report we extend the current ICOS measurement protocol recently published in Nemitz et al (2018), providing further guidelines, tools and recommendations in particular on the following methodological aspects: 1) Data acquisition and synchronization; 2) Overflow inlet system; 3) Raw data despiking; 4) Time lag estimation; 5) Spectral correction approaches for high frequency loss; 6) Friction velocity filtering; 7) Gap-filling; 8) Additional guidance on the conditions under which site conditions fluxes of CH_4 and N_2O are sensible to measure with EC approaches.

Results, presented in this report and in Nemitz et al (2018), will be used by the Ecosystem Thematic Center (ETC) to compile an instruction document for the Station PIs, as well as they will be implemented in data processing chain at ETC in order to produce L2 ICOS data related to CH_4 and N_2O EC fluxes.

Table of Contents

1. INTRODUCTION	5
2. DATA ACQUISITION AND SYNCHRONIZATION.....	6
2.1 Background information on the modern timing protocols for hi-speed data flows.....	6
2.2 Implications of the timing protocols for 10-20 Hz flux measurements	7
2.3 Resulting choices for accepting high-speed data flows from non-PTP CH ₄ /N ₂ O gas analysers	8
3. THE OVERFLOW INLET SYSTEM	9
3.1 Background	9
3.2 Description of the system.....	10
3.3 Determination of system-specific parameters.....	11
3.4 Flux calculations	13
3.5 Results from a field test.....	13
3.6 Recommendations	17
4. RAW DATA SPIKE DETECTION	19
4.1 Background	19
4.2 New despiking routines	19
5. TIME LAG DETECTION	23
5.1 Background	23
5.2 Description and evaluation of a new method for time lag detection	23
6. FREQUENCY RESPONSE CORRECTION METHOD.....	27
6.1 Background	27
6.2 Using power spectra versus cospectra when deriving the spectral transfer function of an EC system from measurements with low signal-to-noise ratio	28
6.3 <i>Determining the spectral transfer function and correction factor with cospectra and its interdependence with the time lag estimation</i>	34
6.4 Conclusions and recommendations	35
7. FILTERING OF 30 MIN FLUXES BASED ON FRICTION VELOCITY.....	36
8. FLUX GAP-FILLING.....	38

8.1	Background	38
8.2	N ₂ O fluxes.....	39
8.3	CH ₄ fluxes	44
9.	GUIDANCE ON FLUX MAGNITUDE AND MEASUREMENT REQUIREMENTS	46
9.1	Importance of CH ₄ exchange with N European Agricultural Soils.....	46
10.	CONCLUSIONS AND RECOMMENDATIONS	50
11.	<i>List of symbols and acronyms</i>	53
12.	<i>References</i>	56

1. INTRODUCTION

Methane (CH_4) and nitrous oxide (N_2O) are the most important non- CO_2 greenhouse gases. The knowledge of their biosphere-atmosphere exchange rates and their short- and long-term budgets are particularly relevant for several ecosystem types, like peatland, cropland, grassland, marshland and lakes. In principle, eddy-covariance (EC) flux measurements of CH_4 and N_2O at ICOS sites are mandatory for all Class-1 observations stations and further recommended for Class-2 sites. However, it may be demonstrated for individual stations that annual budgets of CH_4 and/or N_2O are not measurable by EC following the flux limits provided in the first ICOS Protocol Paper (Nemitz et al., 2018).

Where fluxes are measured by EC, the latest generation of $\text{CH}_4/\text{N}_2\text{O}$ gas analysers is now capable to measure fluxes near their background level (Peltola et al, 2014; Rannik et al, 2015), which can still add up to a considerable fraction of the annual budget. One crucial difference between the fluxes of CO_2 compared with CH_4 and N_2O is that the carbon dioxide (CO_2) flux tends to be strongly bi-directional at the diurnal scale and that the annual budget for ecosystems tends to be a relatively small difference between a large downward component during the day and a large upward component during the night, which makes it particularly important to quantify both components correctly. The flux of N_2O tends to be upward for most of the time, whilst for CH_4 both uptake (oxidation) and emission are observed, but typically change between sites or seasonally at a given site, rather than at the diurnal cycle. Thus, the annual budget of N_2O and CH_4 is much less sensitive to the accuracy of corrections that preferentially apply during day or night-time. By contrast, for N_2O in particular, the flux can be highly sporadic, and annual emission can be entirely dominated by the fluxes during a few days, e.g. related to fertiliser applications, rain episodes or freeze-thaw cycles. As a result, unlike chamber approaches, the eddy-covariance method is ideally suited for capturing the high emission events with good spatial representativeness and temporal coverage, but it may well be challenged by the small fluxes during the remainder of the year.

Within the ICOS network, there are already labelled Ecosystem Stations, where $\text{CH}_4/\text{N}_2\text{O}$ exchange have been measured in the past and the experience gained at these sites have been crucial for writing a state of art protocol including a comprehensive set of recommendations related to the instrument setup and data processing (Nemitz et al, 2018). Most of the methodological approaches introduced in the current protocol are specific for estimating small fluxes which are often close to the detection limit of the EC system. However, Nemitz et al (2018) have recognized the need to further investigate and improve methodological approaches especially related to the data processing and gap-filling.

In this report we aim to provide further guidelines for eddy-covariance flux measurements of CH₄ and N₂O, focusing in particular on the following critical aspects:

- Data acquisition and synchronization
- Overflow inlet system
- Raw data despiking
- Time lag estimation
- Spectral correction approaches for high frequency loss
- Friction velocity filtering
- Gap-filling
- Additional guidance on the conditions under which site conditions fluxes of CH₄ and N₂O are sensible to measure with EC approaches

For each of these aspects, we present novel theoretical approaches, algorithms and experimental setup in order to advance existing methods and provide state of art guidelines to the ICOS Ecosystem Community.

2. DATA ACQUISITION AND SYNCHRONIZATION

2.1 Background information on the modern timing protocols for hi-speed data flows

There are two main modern timing protocols broadly available for a range of computing devices: the older network time protocol (NTP) and the newer precision time protocol (PTP). These two protocols do not perform equally well when it comes to handling collection and alignment of high-frequency data flows. Excerpts below, from IEEE 1588 Standard (Eidson and Lee, 2002) and Rockwell Automation timing protocol guide (Matson, 2013) provide a clear and concise explanation of differences between the two timing protocols.

NTP is used for ‘application-level’ synchronization:

- Coarse level granularity
- Requirement for synchronization guarantee does not exist
- Example: time-stamping error log files
- Sync Accuracy NTP/Ethernet 100 milliseconds
- About 99% are synchronized within 1 second to the synchronization peer

PTP is used for precision synchronization:

- Mission-critical applications
- Dedicated hardware to minimize on-path issues
- High-end algorithms to eliminate network & equipment jitter
- Sync Accuracy PTP Ethernet 20-100 nanoseconds
- About 99% are synchronized within 100 nanoseconds to the synchronization peer on a network specifically designed for IEEE-1588
- Software-based PTP solution has an accuracy of about 100 microseconds, still several orders of magnitudes better than NTP

2.2 Implications of the timing protocols for 10-20 Hz flux measurements

There are two separate key issues (and a number of smaller issues) with different clocks operating on different devices. One key issue is the clock drifts - somewhat consistent time delays and related increases and decreases in the timing of the sampling. These can be linear and non-linear and can go backward and forward for a period of time, often as a result of changes in individual clock temperatures. The second key issue is inconsistent intervals between the individual sampling points, sometimes called “jitter”, often as a result of clock quality, timing protocols, electronics interferences, etc. The first issue is important for any measurements (e.g., open-path, closed-path, 3D motion removal etc.) because it leads to a de-correlation between the signals coming from the sonic anemometer with one clock and the signals coming from another device with a second clock. This issue can be significantly reduced by frequently sending GPS-PTP time via NTP timing protocol to all the NTP-enabled devices on the flux station (e.g., GPS->PTP->NTP route). The second issue is less important for closed-path devices, with smoothed and attenuated time series and effective frequencies well below 10 Hz, but it is crucial for measurements where high-frequency point-by-point data adjustments are fundamental to the flux quality (e.g., point-by-point conversion of density to dry mole fraction, point-by-point 3D motion removal from 3 wind components, etc.). In such cases, the small shifts in sampling intervals due to jitter will cause the effect opposite to the one desired - worsening the signal that was supposed to have been improved by the adjustment. Since the synchronisation errors and timing uncertainties fundamental to NTP timing protocol reach up to 100 ms, it is absolutely unacceptable for clock synchronisation when doing any point-by-point corrections at sampling rates of 10-20 Hz, and a much more accurate timing protocol (such as PTP) must be used in these cases. As a result, all the devices used in point-by-point adjustments need to be either PTP-compatible, or at least be able to provide ways for PTP timing to be implemented.

GPS-PTP based microcomputers can use Ethernet inputs from PTP-enabled devices to achieve true PTP time synchronization. They can also use RS232, RS485 or Ethernet with compatible non-PTP-enabled

devices and still achieve synchronization substantially similar to the true PTP synchronization using the method described in Fratini et al (2018), along with further details and implications of precise timing for stationary EC flux measurements.

2.3 Resulting choices for accepting high-speed data flows from non-PTP CH₄/N₂O gas analysers

If using GPS-PTP based microcomputer (such as, for example, Smartflux used already in over 50% of ICOS sites) as both a system master clock and a fast data collector/integrator, two fundamental solutions may be considered. One actual solution is available now, and one potential solution may be available in the future. Both provide significantly improved synchronization in comparison with the present communications, stand-alone NTP and various legacy timing protocols, as described above in Section 2.2.3. The first solution uses GPS->PTP->NTP time synchronisation to prevent or greatly reduce clock drifts. This solution is available and ready to use as of May 2020. It uses a combination of the hardware and software in GPS-PTP based microcomputer SmartFlux models 2 and 3 to frequently push GPS-PTP time into NTP-enabled CH₄ and N₂O gas analysers. The CH₄ and N₂O time series will still have to be merged with sonic anemometer and CO₂/H₂O gas analyser data afterwards during post-processing, based on the new precise timestamps, and then can be processed further by ICOS ETC flux processing server. Note that files should be submitted to the Carbon Portal (CP) as they are, with their respective (but now synchronised) timing, and then they will be merged by ETC. The GPS->PTP->NTP time synchronization is designed to work with selected analyser models standardized across the entire ICOS network (currently, Aerodyne – QCL, Los Gatos - GGA-911, and Picarro - G2311-f) but potentially could work with other models equipped with Ethernet access and enabled NTP timing protocol. The SmartFlux software update to at least version 2.2.49 is required to implement this solution. The new software can be downloaded from here: <https://www.licor.com/env/support/SmartFlux-2/software.html?Highlight=SmartFlux2>. Detailed instructions on implementing the update are available here: <https://www.licor.com/env/support/SmartFlux/topics/smartflux-software-update.html?Highlight=SmartFlux#UpdatetheSoftware>). The second solution requires the implementation of GPS->PTP, or stand-alone PTP timing protocol, running on the CH₄ and N₂O gas analysers themselves. In this solution, integrating the PTP-stamped data together with sonic anemometer and CO₂/H₂O gas analyser directly inside the GPS-PTP based microcomputer not only removes the clock drifts, as in the first solution, but also greatly reduces the high-frequency jitter. As a result, the processing of all data can be done in near-real-time on-site, or later on ICOS ETC flux processing server. This solution is in principle doable, provided collaboration from the CH₄ and N₂O gas analyser manufacturers and so far, this has not been one of their priorities.

Currently, none of the currently used CH₄ and N₂O gas analysers have an implemented PTP timing protocol. Software changes for the GPS->PTP route could be a major effort and will need to be assessed by different analyser manufacturers. Embedded software along with the communication and integration of data will need to be made for each specific model of the analyser, and then further tested using actual specific models. This will require a high degree of coordination between various manufacturers, and will

result in significant additional software development and extended timelines. For these reasons, the GPS->PTP synchronisation only remains a potential solution.

If not using existing GPS-PTP based microcomputers (like Smartflux or others), the researchers will need to identify or build such a device, configure it and test the entire workflow thoroughly. Specifically, one needs to build a network-wide device with GPS-PTP timing, ability to push PTP timing into current NTP-based CH₄/N₂O analysers and further on to the timed components of the system such as data acquisition cards, and ability to output data in ICOS format. Then, one needs to be able to do PTP-timed data extraction and logging of the sonic anemometer, all other gas analysers and other fast instrumentation. Finally, one needs to make sure that such a solution is compatible with BADMs and EddyPro engine used by ICOS for central processing of raw data.

3. THE OVERFLOW INLET SYSTEM

3.1 Background

With fluxes of CH₄/N₂O being often small for an extended period of time, the ICOS CH₄/N₂O EC Protocol Paper (Nemitz et al., 2018) recognised the processing of fluxes near the detection limit as a particular challenge for these compounds. When fluxes are small over an extended period of time, it is difficult to use the experimental approach to assess the inlet response time. In addition, the use of cross-covariance maximisations (CCM) to establish the time lag between the vertical wind velocity component and the gas mixing ratio can significantly bias the flux towards larger (absolute) values (Langford et al., 2015; Nemitz et al., 2018; Kohonen et al., 2020). There are several potential approaches to deal with these uncertainties regarding small fluxes. A good approach is to derive response and time lag from another compound, measured in the same instrument, which shows larger fluxes, as long as the compound shows similar properties in terms of tube wall interaction. For example, for N₂O a concurrent measurement of water vapour is not suitable, but of CO₂ likely is and CH₄ and CO would be suitable as long as their fluxes are sufficiently large. This method of time-lag by proxy gas has also been used for other compounds showing small fluxes such as VOCs, CO and COS (e.g. Cowan et al., 2018; Kohonen et al., 2020). An alternative, second approach, statistical pre-treatment of the data to remove spurious correlations and identify statistically significant peaks is explored in Section 5 below.

Here, we explore the utility of a third, experimental, approach which was raised by Nemitz et al. (2018) as a possibility: an inlet overflow system is used to periodically deliver a constant reference gas to the analyser. From this system, three characteristics of the EC setup can be derived: (a) tube/analyser transport time-lag, (b) associated time-response and (c) instrument noise. Such experimental setups appear to have been used in the literature to quantify flux detection limits (e.g.

Shurpali et al., 1993; Eugster and Merbold, 2015), but have not explicitly been characterised or tested for routine operation.

3.2 Description of the system

Two eddy-covariance systems for the measurement of fluxes of nitrous oxide (N_2O) and were operated at the Easter Bush grazed, managed grassland site (3.2065 W, 55.8655 N, 190 m a.s.l.) in Scotland, for two periods during the 2019 growing season. First, a dual-laser Quantum Cascade Laser instrument (QCL, Aerodyne) was operated, measuring N_2O , CO and H_2O with one laser, and NO_2 and H_2O with the other, at 10 Hz. Then a single-laser QCL instrument was operated to measure N_2O and carbon dioxide (CO_2) from August to November 2019. Carbon dioxide fluxes are relatively large at the site in summer (typically in the range -20 to +20 $\mu\text{mol m}^{-2} \text{s}^{-1}$); in contrast, N_2O fluxes are usually small (< 8 $\text{nmol m}^{-2} \text{s}^{-1}$) with episodic peaks of up to 100 $\text{nmol m}^{-2} \text{s}^{-1}$ (e.g. Jones et al., 2011). The eddy-covariance system was based on an ultrasonic anemometer (Windmaster Pro; Gill Instruments) operating at 20 Hz. The air was sampled from 25 cm below the centre of the anemometer's head through a 15 m-long Dekabon tube of outer diameter 6.1 mm ($\frac{1}{4}''$). The effective measurement height was 2.5 m and the lateral separation between the inlet and the centre of the anemometer's averaging volume was 9 cm (Figure 1). The inlet was fitted with a normally-closed solenoid valve to allow for injection of compressed air, near the inlet, with concentrations of N_2O and CO_2 above typical ambient values. The injection pressure of the compressed air was adjusted to obtain a small overflow at the inlet to prevent the ingress of ambient air whilst minimising pressure fluctuations. A custom LabView™ program was used to automate the actuation of the solenoid valve at user-defined intervals and to record both the measurement data (wind and gas concentration) as well as the valve status.

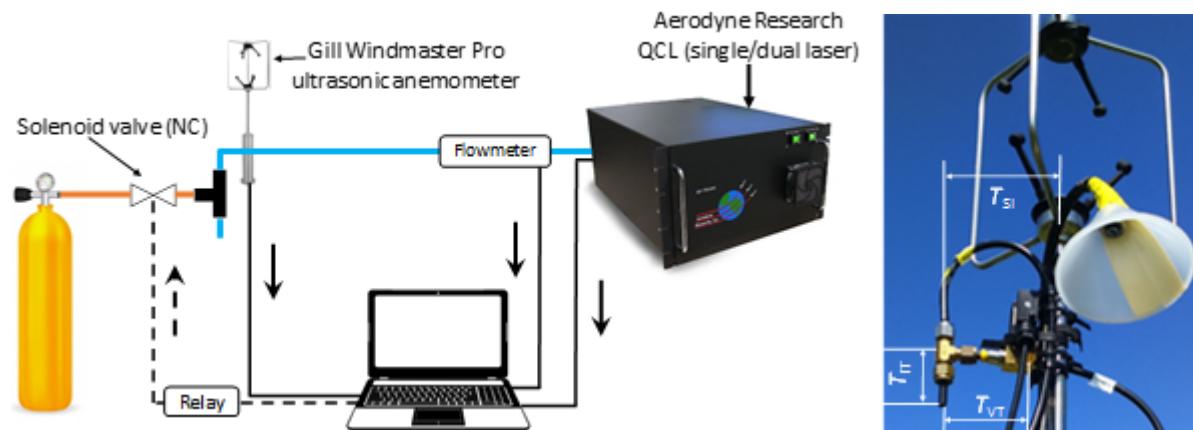


Figure 1. (a) Schematic of the setup. (b) Photograph of the inlet and the distances involved.

3.3 Determination of system-specific parameters

Instrument response time and experimental transport time lag (from time point of injection to detection by the gas analyser) were determined from 30-second gas injection events, which were programmed to occur every 2 to 3 hours. Additionally, instrument noise (standard deviation of the measured gas concentration during the last 2 minutes of injection) and measurement precision (Allan variance over 4 minutes when the measured gas concentration was deemed stable) were estimated from longer 5-minute injections scheduled to occur typically 3 times a week (Figure 2).

The experimental transport time lag (T_{exp}) can be expressed as a sum of setup-specific terms:

$$T_{exp} = T_{line} + T_{analyser} + T_{VT} + T_{actuation} \quad (1)$$

Here T_{line} is the gas travel time through the sampling tube, $T_{analyser}$ is an analyser-specific response time (a priori unknown), T_{VT} is the gas travel time from valve to T-piece (Figure 1) and $T_{actuation}$ is the actuation time of the solenoid valve (unknown but assumed to be negligible).

Similarly, the theoretical “eddy-covariance” time lag (T_{theo}^{EC}) can be expressed as in Eq. 2:

$$T_{theo}^{EC} = T_{line} + T_{analyser} + T_{SI} + T_{IT} \quad (2)$$

In Eq. 2, T_{SI} is the horizontal travel time from inlet to sonic (corrected for wind direction and speed) and T_{IT} is the travel time from inlet to T-piece. Equations (1) and (2) can be reorganised and merged and expressed in terms of T_{exp} :

$$T_{theo}^{EC} \cong T_{exp} - T_{VT} + T_{SI} + T_{IT} \quad (3)$$

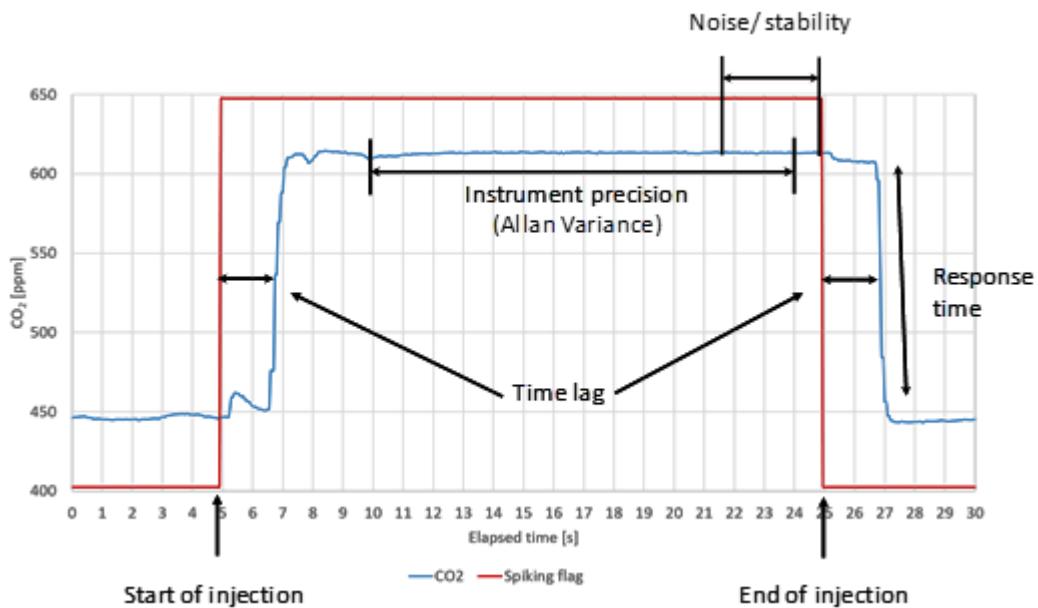


Figure 2. Example trace of CO_2 concentration recorded across one injection event.

The time response of the instrument was defined as the time constant τ of the exponential function fitted to the decaying tail of the concentration time series (Figure 3):

$$y = y_0 + Ae^{-(t-t_0)/\tau}, \quad (4)$$

where y_0 , A , t_0 and τ are fitting parameters.

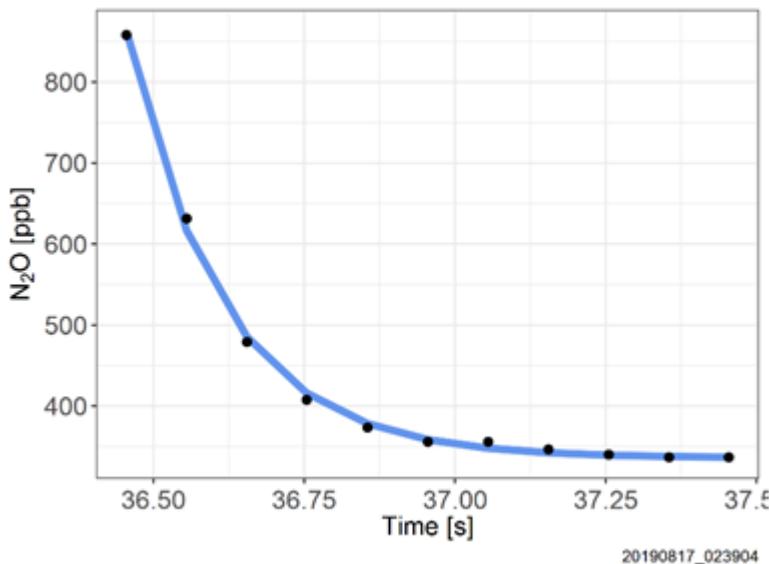


Figure 3. Exponential decay fit of the decaying part of the N_2O concentration signal.

The gas analyser time response was evaluated for each 30-s gas injection event (172 events with the dual laser QCL and 103 events with the single laser). Additionally, 100 gas injection events were carried out with the single laser analyser using compressed air with CO₂ concentration closer to ambient (raw data not shown). The non-linear regressions were applied in a pre-defined time window (36.4 s to 37.5 s, time relative to the start of the injection event). These boundaries were determined empirically for the first gas injection file by searching for the temporal range yielding the best fit results, and were then applied to all subsequent gas injection events using an automated data processing algorithm coded in R.

3.4 Flux calculations

EddyPro version 7.0.6 was used to calculate fluxes of N₂O and CO₂; time lags were determined by covariance maximisation without default. Minimum, nominal and maximum values of the time lag were left blank in the raw file description input section, forcing EddyPro to evaluate them from the physical properties of the experimental setup (flow rate, length and inner diameter of the inlet line).

The time lags for CO₂ determined by EddyPro were used to shift the high-frequency time series of N₂O concentrations and vertical wind speed with respect to one another based on the assumption that anemometer records are instantaneous and analyser records are delayed. In practice, the concentration time series of N₂O were truncated by N points (N = acquisition frequency x time lag) at the start of each hourly data file. A corresponding portion of length N points was trimmed off the end of the sonic time series to ensure that both data series were of equal length. These composite high-frequency data series were processed in EddyPro 7.0.6 with a constant time lag of 0.01 s (EddyPro does not allow a 0 s constant time lag). Finally, fluxes of N₂O were processed in EddyPro 7.0.6 using the median of CO₂ experimental time lags obtained from the gas injection events.

3.5 Results from a field test

3.5.1 Instrument noise

Instrument noise was found to be instrument-specific and varied considerably between the dual- and single-laser analysers used (Figure 4), averaging 0.23 and 0.41 ppb, respectively. Furthermore, the frequency stability of the single-laser instrument was poorer than the dual laser's (Allan variance analysis, data not shown), which was perhaps symptomatic of aging optics on the older single-laser instrument as both instrument c lasers with similar sensitivity to N₂O. The manufacturer quotes a 10-Hz RMS of 0.09 ppb for both laser types under optimum operating conditions at typical ambient concentrations (Nemitz et al., 2018). Instrument noise consists of a constant and a concentration-dependent component, with the

proportional noise component becoming dominant at high concentrations (David Nelson, Aerodyne, pers. commun.). Thus, the use of a rather large concentration of 982 ppb for the reference gas, i.e. about 3x ambient, accounts for much of the difference.

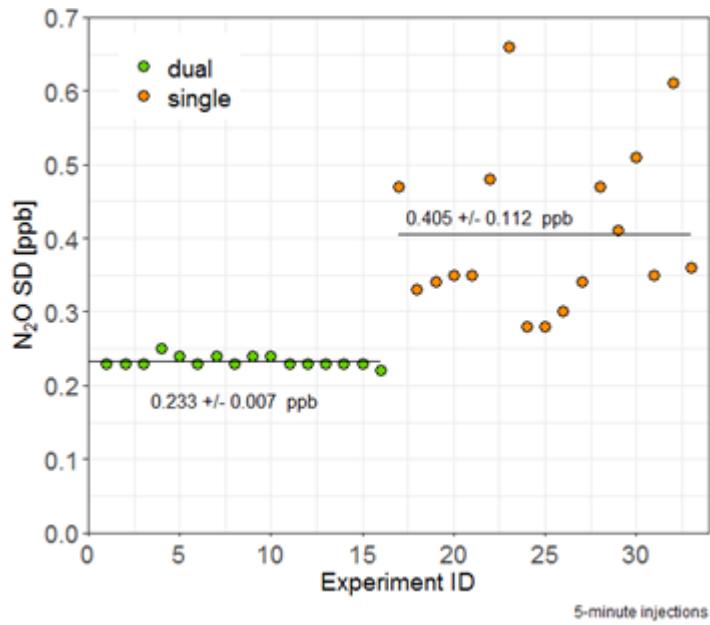


Figure 4. Comparison of instrument noise (standard deviation) for nitrous oxide concentrations measured with a single-laser and a dual-laser QCL, respectively.

3.5.2 Response time

The response time also differed between instrument setups with the dual-laser analyser's time response being twice that of the single-laser instrument (Figure 5). Furthermore, the apparent increase of the time response seen in Figure 5 was an artefact of the fixed time window used for the derivation of τ by fitting of the non-linear regression function (Eq. 4). Nevertheless, it draws attention to a change in the instrument, possibly the degradation in instrument performance (CPU limitations and tuning rate). Only the periods marked in blue in Figure 5 were used for the subsequent analysis.

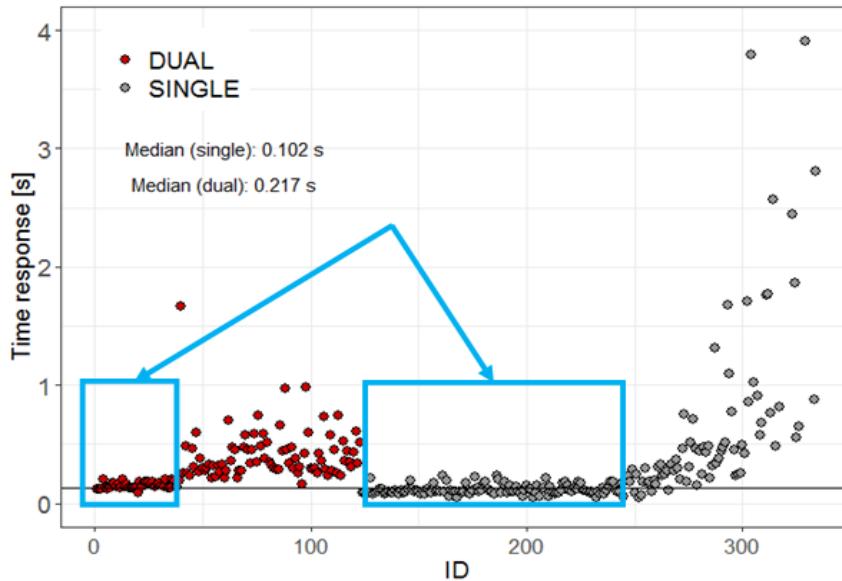


Figure 5. Consistency and apparent increase in the (e-folding) time response of the single and dual laser QCL instruments used in this study. The large change in time response is an artefact of the non-linear fitting procedure but this information is symptomatic of a degradation of the performance of the instrument over time.

3.5.3 Time lag estimates and effect on average fluxes

The time lag values for CO₂ fluxes obtained by covariance maximisation without default were in good agreement with the experimental transport time lag estimate from gas injection events (Figure 7a), but this was not the case for N₂O for which many EC time lags diverged from (Figure 7b). Stratification of the data by whether the flux exceeded the limit of detection (LoD) revealed that low signal-to-noise only explained about half of the time-lags that were clearly incorrectly quantified, both for CO₂ and N₂O. Similarly, it was found that u_* , relative humidity or the magnitude of the flux itself were also poor determinants of poor performance of the time-lag routine.

The discrepancies observed for N₂O time lags were propagated to the fluxes of that gas (Figure 7a): fluxes calculated on the basis of time lags from covariance maximisation without default, which were generally larger than the experimental time lag values, tended to overestimate the experimental time lag (TL) fluxes, in line with the biasing reported by Langford et al. (2015) and Kohonen et al. (2020). This was also the case for absolute CO₂ fluxes, albeit to a lesser extent than for N₂O. However, N₂O flux estimates from experimental TL were consistent with fluxes obtained by dynamically ascribing time lags from the CO₂ flux analyses (Figure 7c). This suggests that proxy gas (CO₂ in this instance) and overflow inlet both have the

potential to decrease the uncertainty and systematic bias which can arise from poor time lag determination in fluxes with low signal-to-noise ratios.

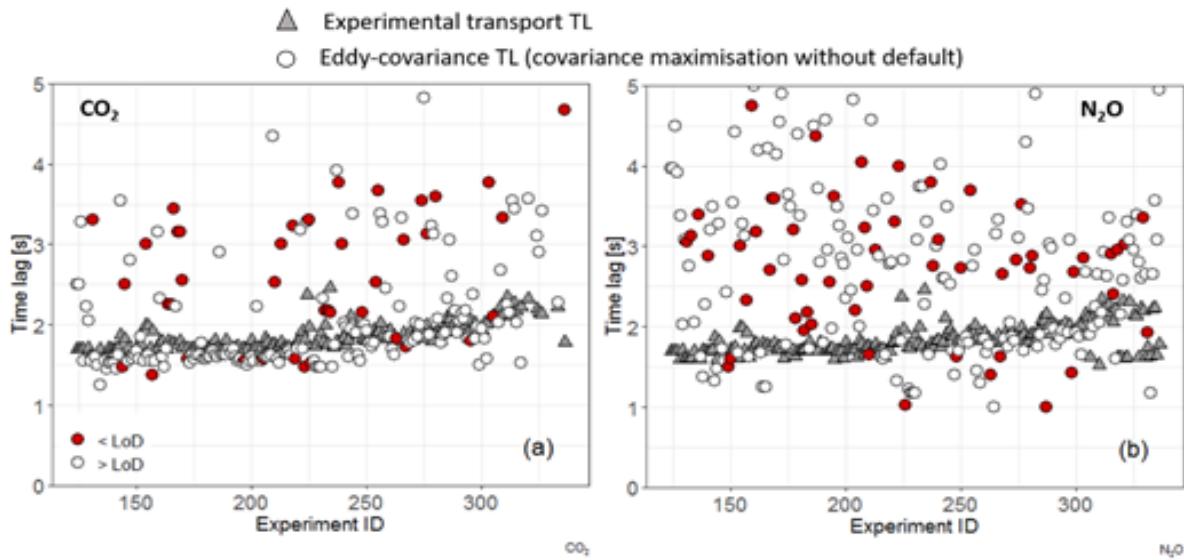


Figure 6. Comparison of experimental transport time lags (triangles) and time lags obtained from eddy-covariance analysis (circles; covariance maximisation without default) for (a) CO_2 and (b) N_2O . Solid circles denote flux values below the limit of detection.

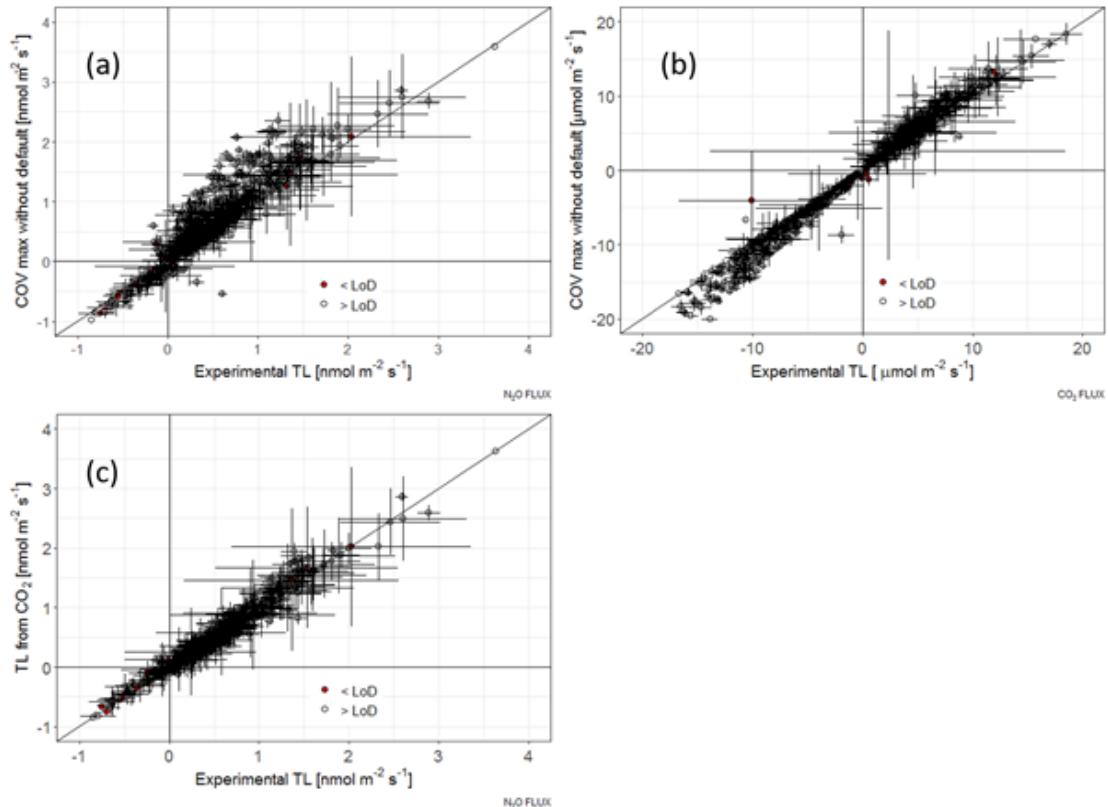


Figure 7. Comparison of fluxes calculated with time lags obtained by covariance maximisation without default with fluxes calculated with experimental time lags ((a) N₂O, (b) CO₂). (c) N₂O fluxes calculated with dynamically ascribed time lags obtained for CO₂ (covariance maximisation without default) versus fluxes from experimental time lags.

Table 1. Comparison of the average fluxes of N₂O, derived with different time-lag strategies, including (a) covariance maximisation of the N₂O data, (b) using the CO₂ time-lag as a proxy and (c) applying the time-lag experimentally derived with the overflow inlet.

Method	Mean flux ± SD [nmol m ⁻² s ⁻¹]
Covariance maximisation without default	0.525 ± 0.557
TL from CO ₂	0.408 ± 0.457
Experimental TL	0.406 ± 0.434

3.6 Recommendations

The overflow inlet proved to be a reliable way to derive instrument response times, time lags and instrument noise. There are advantages and disadvantages to this approach.

Disadvantages:

- Both time-lag and response time do not describe the response of the full system. However, for most setups the dominant aspects are reflected. The additional time-lag caused by the inlet is very small, whilst the time-lag due to sensor separation can be predicted reasonably well from first principles. The response time does not cover the effect of sensor separation and spatial averaging.
- There is no commercial standard setup available and at present each research team would need to develop its own hardware implementation. However, if the setup were to become used more widely, common implementation schemes would be shared and commercial solutions may appear.
- In order to automate the analysis of the data additional flags need to be recorded by the data acquisition system and interpreted by the analysis routines.

Advantages:

- As a measurement-based approach the overflow inlet does not make any theoretical assumptions which may be violated in real measurement situations.
- The overflow inlet arguably provides the most robust approach for quantifying and tracking instrument noise. Although software implementations for noise quantification do exist, based on the assessment of the auto-correlation function (Lenschow et al., 2000), these only work for white noise and also start to fail if noise levels are small.
- With the overflow inlet in place it is easy to use it for a number of other useful applications, such as long-term stability assessment (e.g. drift, non-white noise characterisation) and periodic automatic calibrations.

The experience with the overflow inlet test has provided some important pointers on what to take into consideration for future applications and development:

1. *Concentration of reference gas.* In this trial, we used compressed gas cylinders as reference gases rather than (certified) standards of a specified concentration. Whilst this is a very cost-effective way of providing a gas at a constant concentration, the concentration is left to chance. In our case, it was well above ambient levels for N₂O (982 ppb) and also CO₂ (4550 ppm). This provided the large step-change in concentration needed for the time-lag and time-response quantification, but it was not a perfect solution for the noise quantification. Because above a certain concentration the measurement noise becomes a function of the concentration, our measurements did not directly quantify the noise at ambient concentration, although they can be extrapolated to ambient concentrations and also allow instrument performance to be tracked with time. An alternative approach would be to modify the system to provide first a low (e.g. 200 ppb) or high (600 ppb) gas and then switch to a near-ambient standard for the characterisation. Note that the measurement of zero gases can be problematic with optical absorption instruments as it causes problems in tracking the absorption features. The use of specified reference gas concentrations adds cost to the operation of the technique because calibration standards, even if not certified, are significantly more expensive than compressed gas cylinders.
2. *Purging of reference gas line.* The concentration trace recorded during an injection (Figure 2) shows a number of repeatable features. Some are due to the gradual response of the instrument to abrupt (even small) pressure changes. A transient dip in the concentration following the rise of the signal following the reference gas injection is likely due to a small depletion of N₂O in the static reference

gas since the last injection. This might be avoided if the injection is carried out twice in short succession and only the 2nd injection used for the characterisation.

Overall, these tests have shown that the overflow inlet provides a robust pragmatic experimental solution to estimate time-lag and system response time for low-flux situations based on an objective method, which, unlike CCM, leads to unbiased fluxes. Because of the additional complexity and cost in implementing, operating and processing the data from the overflow inlet, the use of a laser which also measures another compound with a large flux is the favoured approach. This makes the N₂O/CO₂ and N₂O/CO₂/CO lasers particularly useful for measuring N₂O fluxes. However, at sites which already have purchased a different laser, or cannot afford purchasing and operating separate instruments for N₂O and CH₄ at sites where both fluxes can be small over long period of times, the overflow inlet would be an appropriate solution.

4. RAW DATA SPIKE DETECTION

4.1 Background

Raw, high-frequency, eddy covariance (EC) time series may be corrupted by unexpectedly impulsive peaks originating from mechanical or electronic malfunctions of one of the components constituting the EC measurement system, or from disturbances due to the presence of insects, to name a few. Spikes constitute a particular class of anomalies known as additive outlier and need to be removed to avoid biases in subsequent analyses. The size of bias introduced in flux estimates depends on the magnitude of the spikes, on their number and, in particular, on the simultaneous occurrence in both the time series needed for the covariance calculation. Moreover, their presence can have disturbing effects on estimates for such quantities as higher-statistical moments (e.g. variance, kurtosis), ordinary least squares (OLS) regression coefficients (e.g. those derived for linear trend removal) and (co-)spectral density (i.e. spikes pose a problem for the estimation of reliable spectral correction factors used to adjust for flux attenuation due to sensors separation).

4.2 New despiking routines

Despiking EC data is challenging because the complex dynamics (e.g. the presence of non-linear trend components, structural changes, abrupt changes) and the high and time-varying level of noise characterizing high-frequency time series can make it difficult to distinguish features in the time-series that are introduced by analyser signal disturbances (spikes) from those introduced by atmospheric

turbulence (“good” data). To overcome such difficulties, we evaluated the enforceability of a new despiking procedure based on robust functionals. The procedure involves *i)* a preliminary signal-from-noise separation to alleviate the confounding effects of time series dynamics, followed by *ii)* an analysis of the distribution of the noise component for the detection of those anomalies representative of spikes. To this end, we assume an additive component model as follows:

$$x(t) = \mu(t) + e(t) \quad (5a)$$

$$e(t) = \sigma(t)\varepsilon(t) + \eta(t), \quad (5b)$$

where $x(t)$ is a generic time series observed at time t , $\mu(t)$ denotes the underlying signal level, $e(t)$ denotes the noise component. To allow heteroskedasticity of the noise component, $e(t)$ is split into a stationary white noise process $\varepsilon(t)$ with zero expectation and unit variance multiplied by a time dependent standard deviation $\sigma(t)$. The time-varying variance $\sigma^2(t)$ is allowed to change slowly over time, so that we can treat it locally as (almost) constant. The spike generating mechanism is represented by $\eta(t)$ which is zero most of the time, but occasionally generates large absolute values.

For $\mu(t)$ signal extraction, we used a robust regression technique based on repeated median (RM) filtering (Siegel, 1982), which has been widely demonstrated to perform better than standard filtering methods such as running median when time series are characterized by local temporal trends dynamics (Davies et al, 2004). For the scale parameter of the noise component $\sigma(t)$, the Qn estimator by Rousseeuw and Croux (1993) was selected because of its suitability to prevent implosion of the estimates when data are characterized by extreme low variability as often encountered in EC data (Vitale et al, 2020). Both functionals are calculated in moving time windows of small to moderate length to account for the complex dynamics of the signal and the heteroskedastic behaviour of the noise component. The RM and Qn estimators were performed by means of the R functions implemented in the *robfilter* (Fried et al, 2019) and *robustbase* (Maechler et al, 2019) packages, respectively.

Once $\mu(t)$ and $\sigma(t)$ estimated, the spike detection is performed through the examination of outlier scores $z(t) = [x(t) - \mu(t)]/\sigma(t)$. Any data point in raw, high-frequency, time series with outlier scores exceeding ± 5 is detected as spike and removed.

The effectiveness of the proposed approach was demonstrated via a number of experiments on simulated data. In particular, to mimic the heteroskedastic behaviour of EC time series, synthetic time series were generated from a first-order autoregressive AR(1) model with heteroskedastic error structure. To this end, we used the Generalized AutoRegressive Conditional Heteroskedastic (GARCH) model popularized

by Bollerslev (1986). Such a model originates from econometrics (for more details, the reader can refer to the authoritative textbook by Tsay 2005) and has been applied to a wide range of time series analyses since it constitutes the standard tool to model time series exhibiting changes in variance and volatility clustering (in finance, volatility clustering refers to the observation that large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes, see Mandelbrot1997). The simulation model specification is as follows:

$$\begin{aligned}
 y(t) &= \mu(t) + e(t) \\
 \mu(t) &= \varphi y(t-1) \\
 e(t) &= \sigma(t) \varepsilon(t) \\
 \sigma^2(t) &= \omega + \beta \sigma^2(t-1) + \alpha [y(t-1) - \mu(t-1)]^2
 \end{aligned} \tag{6}$$

where $y(t)$ is the simulated time series, $\mu(t)$ is the conditional mean process generated by a stationary AR(1) process with autoregressive parameter φ set equal to 0.975; $e(t)$ is the error term composed by $\varepsilon(t)$, an innovation series normally and independently distributed with zero mean, unit variance and a finite fourth moment, and $\sigma(t)$, the square root of the conditional variance process generated by a GARCH(1,1) process with $\omega=1e-6$, $\alpha=0.1$, $\beta=0.89$. Notice that a high value of β causes $\sigma(t)$ to be highly correlated with $\sigma(t-1)$ and gives the conditional standard deviation process a relatively long-term persistence. With this parameter setting, we simulated 18000 values, as in EC raw high-frequency data sampled at 10 Hz scanning frequency within a 30-min interval.

Successively, each simulated time series was intentionally corrupted with 180 spikes (1% of the sample size). The spike locations were randomly selected, and divided into 90 single spikes, 60 spikes as double (consecutive) events, and 30 spikes as triple (consecutive) events. Once located, the corresponding time series values were multiplied by $k = 7$ in order to generate spikes of a magnitude similar to those observed in real data. Simulations were executed in the R programming environment by using tools implemented in the R package fGarch (Wuertz et al, 2019). Each scenario was permuted 99 times. Figure 8 depicts the realizations of stationary AR(1)-GARCH(1,1) processes simulated with the parameter setting as previously specified.

The skill of the proposed despiking procedure was evaluated in terms of Precision (the fraction of reported spikes that truly turn out to be spikes), Recall (the fraction of ground-truth spikes that have been reported as spikes) and F-Score (a measure that combines Precision and Recall metrics; the closer to 1 the F-Score, the greater the effectiveness of the despiking method). To aid the comparison, the metrics were calculated also for existing despiking methodologies.

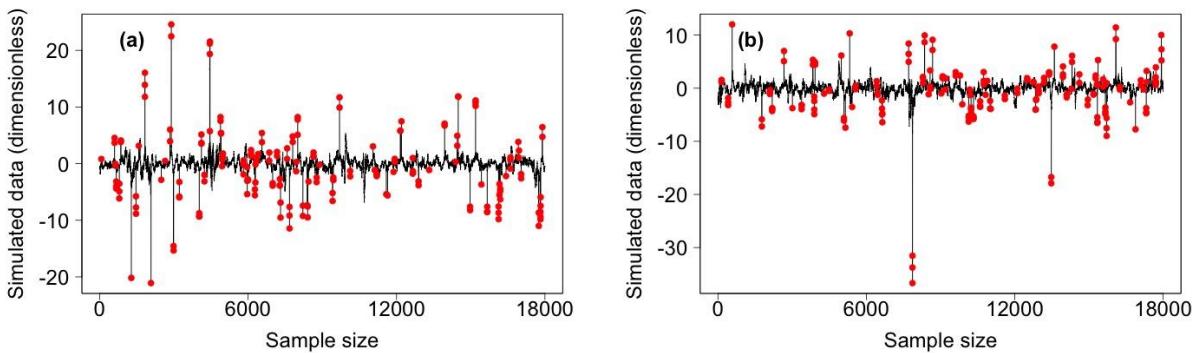


Figure 8. Illustrative examples of synthetic time series generated from AR(1)-GARCH(1,1) processes (black lines) contaminated by spikes (red points). The model parameters are $\varphi=0.975$, $\omega=1e-6$, $\alpha=0.1$, $\beta=0.89$.

As reported in Table 2, we observed superior performance of the proposed approach compared to existing methodologies with Precision, Recall and F1-Score metrics equal to 0.99, 0.81 and 0.89, respectively. The approach by Starkenburg et al (2016) showed excellent results in terms of Precision (0.99) but a very low performance in terms of Recall (0.08). Approaches by Vickers and Mahrt (1997), Metzger et al (2012) and Mauder et al (2013) exhibited moderate performance levels (all metrics resulted lower than 0.75).

Table 2. Performance evaluation metrics of despiking methods

Despiking method	Precision	Recall	F1-Score
Perfect performance	1.00	1.00	1.00
Vickers and Mahrt (1997)	0.63	0.53	0.58
Metzger et al (2012)	0.69	0.59	0.63
Mauder et al (2013)	0.74	0.33	0.46
Starkenburg et al (2016)	0.99	0.08	0.14
this work	0.99	0.81	0.89

Although promising results were achieved, further examinations are required. In particular, performances need to be evaluated over a wide range of scenarios, including CH₄ bursts related to ebullition events and/or presence of ruminants in the field. Furthermore, although not prohibitive for the processing of EC data, it would be straightforward to reduce the computational demand required by the RM and Qn estimators. An in-depth description of the above considerations will be reported in Vitale et al (in prep), while a software implementation suitable to EC data processing will be freely available in the R software package RFlux (Vitale et al, 2019)

5. TIME LAG DETECTION

5.1 Background

As already introduced in Section 3, the exact detection of time lag is essential to properly synchronize EC time series and achieve unbiased estimates of flux covariances. One of the most used procedure consists of identifying the optimal time lag in correspondence of the time lag that maximizes (in absolute terms) the covariance between vertical wind speed (w) and the atmospheric concentrations of the scalar of interest (e.g., CO₂, H₂O, CH₄, N₂O), within a plausible temporal window. However, such an approach is effective only under stationary conditions and when the signal-to-noise ratio is moderate to high. In all other circumstances, the optimal time lag detection becomes difficult to achieve. As a consequence, there is a concrete risk to introduce significant bias in flux estimation, in particular when flux exchanges are of small magnitude (Langford et al., 2015).

5.2 Description and evaluation of a new method for time lag detection

To overcome such difficulties, we evaluated the enforceability of a new time lag detection procedure based on what is known in the statistical literature of time series analysis under the name of *prewhitening* (the interested reader can refer to the leading textbook on time series analysis by Cryer and Chan (2008) to have a good exposition of the mathematics underlying the method). In summary, given two EC time series x (e.g. a scalar concentration) and y (e.g. vertical wind speed), the proposed procedure for time lag detection consists of the following steps:

- I. Determine a time series model for the x -variable, e.g. via an autoregressive model of order p ;
- II. Filter the y -variable series using the x -variable model (i.e. apply the same transfer function model for x to y), this constitutes the pre-whitening step;
- III. Estimate the cross-correlation function (CCF) between the residuals from x -variable model (step 1) and the filtered y -values (step 2);
- IV. Detect the (statistically) significant peak of the CCF as representative of time lag between x and y time series (notice that prewhitening makes it possible to approximate critical values of the CCF as $\pm z_{\alpha/2}/\sqrt{n}$ where $z_{\alpha/2}$ is the critical value of a Standard Normal distribution for a given confidence level and n is the sample size).

In the following, we show some illustrative examples of the application of the proposed procedure. To better outline the potential benefits deriving from prewhitening we performed a simulation study. The

main purpose of the simulations is to create pairs of reference time series with known covariance structure that, after being misaligned with a specific time lag, allow an objective evaluation of the ability of the procedure to detect the “true” time lag. In particular, we first simulated three bivariate systems of stationary autoregressive (AR) processes (x, y) with $\phi=0.99$ and correlation equal to 0.10 (moderate), 0.05 (low) and 0 (independent processes), respectively (for more detail see Vitale et al, 2020, Appendix C). Subsequently, y time series was shifted of 20 timesteps to simulate a time lag.

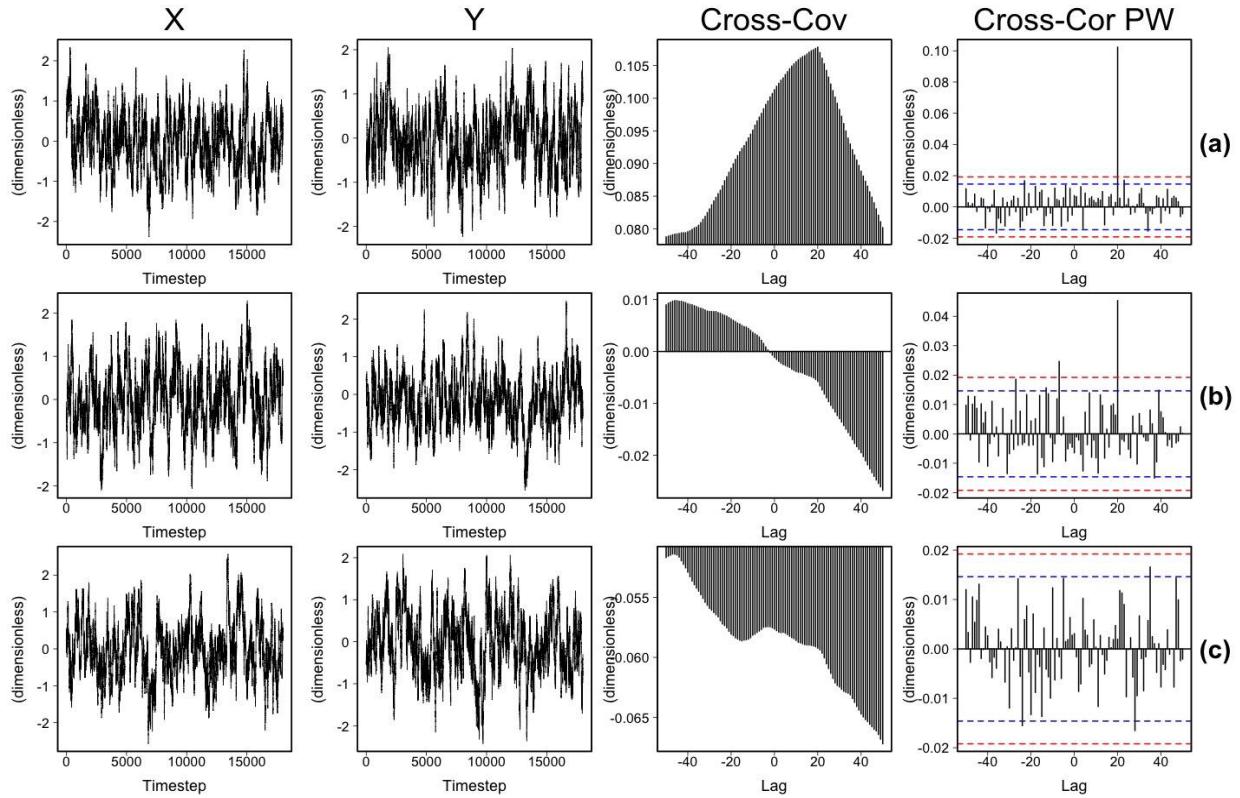


Figure 9. Three bivariate systems of simulated first-order autoregressive processes (x, y) with pre-fixed correlation structure equal to 0.10 (panel a), 0.05 (panel b) and 0 (independent processes, panel c) and their cross-covariance function and cross-correlation function after prewhitening. Blue and red dashed lines indicate the 0.05 and 0.01 significance level of the crros-correlation estimates, respectively.

The pairs of simulated AR processes are shown in panels a-c of Figure 9 with their cross-covariance (Cross-Cov) and cross-correlation functions, this last estimated after prewhitening (Cross-Cor PW) procedure. As expected, both the Cross-Cov and Cross-Cor PW exhibit a dominant peak at lag 20 in case of the bivariate system with correlation structure equal to 0.10. When the correlation structure between time series decreases, the shape of the Cross-Cov function becomes more difficult to interpret. In panels (b-c) the dominant peak of the Cross-Cov in fact disappears, making it difficult to detect the time lag by means of the standard maximum covariance procedure. In these circumstances, the risk to obtain biased flux

estimates becomes relevant. Notice that such a risk is mainly related to the degree of serial dependence and not only to the degree of signal-to-noise ratio. The higher the degree of serial dependence, the higher the probability of incurring in the so-called spurious correlations.

By prewhitening time series such a risk is completely avoided, because the cross-correlations are estimated between serially uncorrelated residuals resembling white noise processes. The evaluation of the Cross-Cor PW functions clearly shows the presence of statistically significant peaks at lag 20 between correlated time series (a-b panels of Figure 10, while no peak (at 0.01 significance level) was detected between independent time series, as expected (panel c).

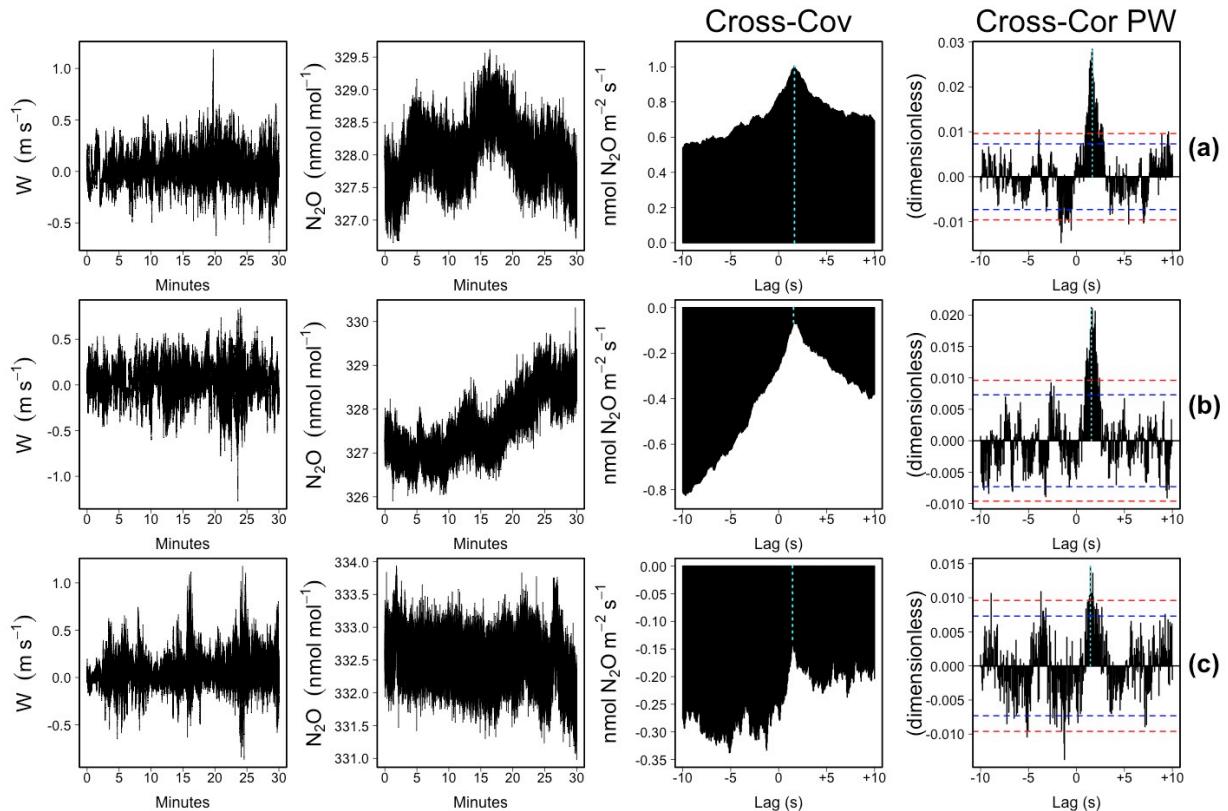


Figure 10. Illustrative examples of time lag detection by means of prewhitening on real EC data. From left to right: vertical wind speed (w), nitrous dioxide (N_2O), cross-covariance estimates, cross-correlation function after prewhitening. Blue and red horizontal dashed lines indicate the 0.05 and 0.01 significance level, respectively. Vertical cyan dashed lines denote the optimal time lag.

Illustrative examples of time lag detection by means of prewhitening on real EC data (for more details about data and EC system setup see Section 3.2) are given in Figure 11. For the time series depicted in the top panel (a), both the cross-covariance function and the cross-correlation function after prewhitening detected the optimal time lag at 1.5 s. In this case, the use of prewhitening may not be necessary since the Cross-Cov exhibits a dominant peak in absolute terms. A different situation occurs in cases of time

series depicted in the b-c panels. Here the time lag detection by means of the covariance maximization could lead to biased flux estimates since the optimal time lag should be detected when the Cross-Cov function reaches the minimum, not the maximum (in absolute terms). Time lag detected by means of the cross-correlation function after prewhitening reveals instead another story. In both cases and as expected for this EC system, the cross-correlation function exhibits a dominant peak (statistically significant at 0.01 level) at 1.5 s. Notice that the statistical significance of correlations is evaluated independently from the flux magnitude, which in these cases are close to zero.

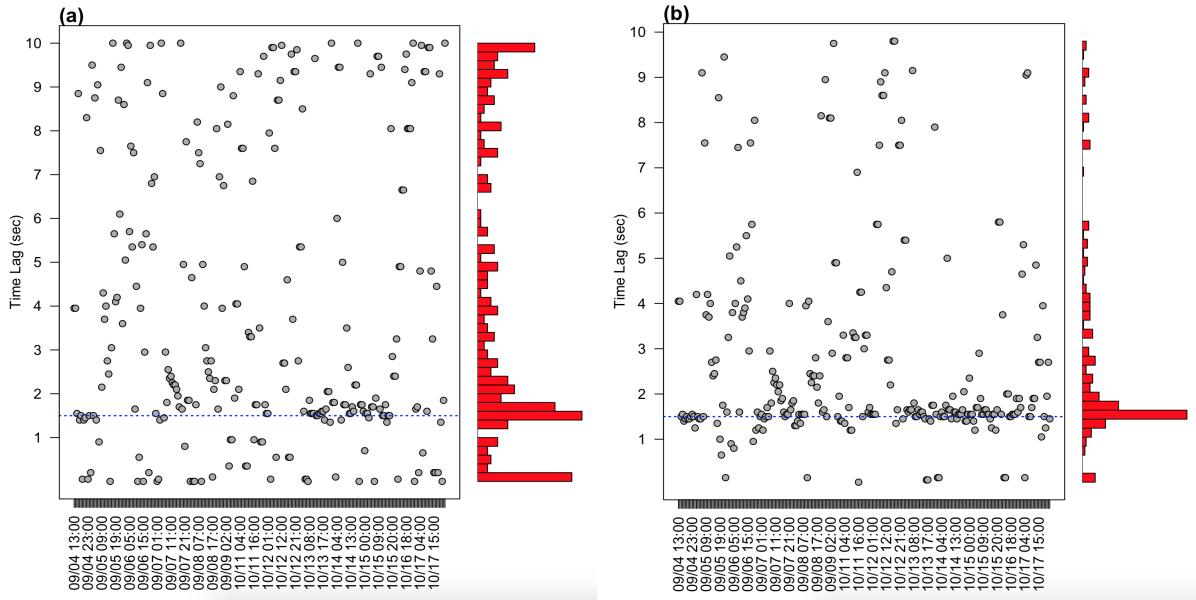


Figure 11. Time lag detected by means of cross-covariance maximization (CCM) approach (panel a) and prewhitening (panel b) on real EC data. For both procedures the optimal time lag was detected in a plausible temporal window of 0-10 sec. Dashed horizontal blue line indicate the experimental time lag. For more details about data and the method used to calculate the experimental time lag see Section 3.

A comparison of the time lag detected by pre-whitening procedure with those identified by the widely used approach based on Cross-Covariance Maximization (CCM) is depicted in Figure 11. Both the procedures were performed by detecting the optimal time lag in a plausible temporal window ranging from 0 to 10 seconds. Results indicate that time lags detected by pre-whitening are more stable compared to those obtained by CCM. Furthermore, the most frequent value of time lag detected was in strict agreement with the experimental time lag calculated for this EC system (see Section 3). Large departures from the experimental value occur mostly in the case of low magnitude flux derived from raw, high-frequency, time series affected by strong non-stationary conditions (e.g. abrupt changes in the mean levels, severe heteroskedastic behaviour).

An in-depth description of the prewhitening procedure suitable for the automatic and fully-data driven ingestion of large EC raw data will be reported in Vitale et al (in prep), while a software implementation will be freely available in the R software package RFlux (Vitale et al, 2019).

6. FREQUENCY RESPONSE CORRECTION METHOD

6.1 Background

Turbulent flux measurements rely on the fast detection of atmospheric signals. While wind measurements can and need to be done *in situ*, gas concentrations are often measured in analysers that have their analytic unit enclosed (closed-path analysers). In such systems, the air sample is transported from the position of the sonic anemometer to this analytic unit via a sampling system. Both, the sampling system and the analyser dampen the atmospheric signal in the high frequency domain. While the fixed sampling interval to, usually, 30 minutes acts as a high-pass filter, the attenuation by the closed-path sensors acts as a low-pass filter. The turbulence statistics, i.e. variances and covariances, calculated with the attenuated signals must be corrected for both low-pass and high-pass filtering to represent the true atmospheric situation. This correction is called frequency response correction (or spectral correction) and it is usually performed based on a priori knowledge of the system transfer function and the unattenuated cospectrum. In this section, we focus on the spectral correction of the low-pass filter effects, only.

The low-pass filter flux correction factor (CF) is defined as

$$CF = \frac{\int_{f_1}^{f_2} Co_{WT} df}{\int_{f_1}^{f_2} Co_{WT} H_{LPF} df} \quad (7)$$

with Co_{WT} the normalized cospectrum of kinematic heat flux $\bar{w'}T'$, which is assumed to be unattenuated, H_{LPF} the low pass filter transfer function, f is the frequency in Hz, f_1 and f_2 are the integration limits set by the length of the averaging period and the Nyquist frequency. Spectral correction is part of all current raw data post processing software, but there are different approaches to derive the spectral features of the eddy covariance systems and to apply the correction to variances and covariances. In this section, we present results from investigations on two aspects related to spectral correction and flux calculation in non-CO₂ GHG eddy covariance flux measurements.

6.2 Using power spectra versus cospectra when deriving the spectral transfer function of an EC system from measurements with low signal-to-noise ratio

The way an EC system affects the signal is represented by its spectral transfer function (H_{LPF}), which is the relative magnitude of the attenuated signal as a function of natural frequency (f). The H_{LPF} can be described analytically, if all quantitative effects of the components of the system are known, which is generally not the case with closed-path EC systems. Alternatively, H_{LPF} can be empirically estimated with experimental approaches or by comparison with an unattenuated atmospheric signals, the latter being the most commonly used approach in turbulence data processing. Previous work has suggested using co-spectra (Aubinet et al., 1999, Moncrieff et al., 1997, Mammarella et al., 2009) or power spectra (Ibrom et al., 2007; Fratini et al., 2012; Sabbatini et al., 2018) for this comparison, a choice which has hence not been systematically examined. An important condition for the applicability of either of the approaches is the signal to noise ratio (SNR), i.e. the magnitude of the flux compared to its detection limit. If the SNR is low, the noise compromises the use of power spectra, as the spectral domain where the turbulent signal is small, especially when the sensor is a strongly low-pass filtered, is contaminated by noise, while the attenuated reference signal is not. In order to compensate for this effect, the noise can be subtracted from the power spectrum (Figure 11).

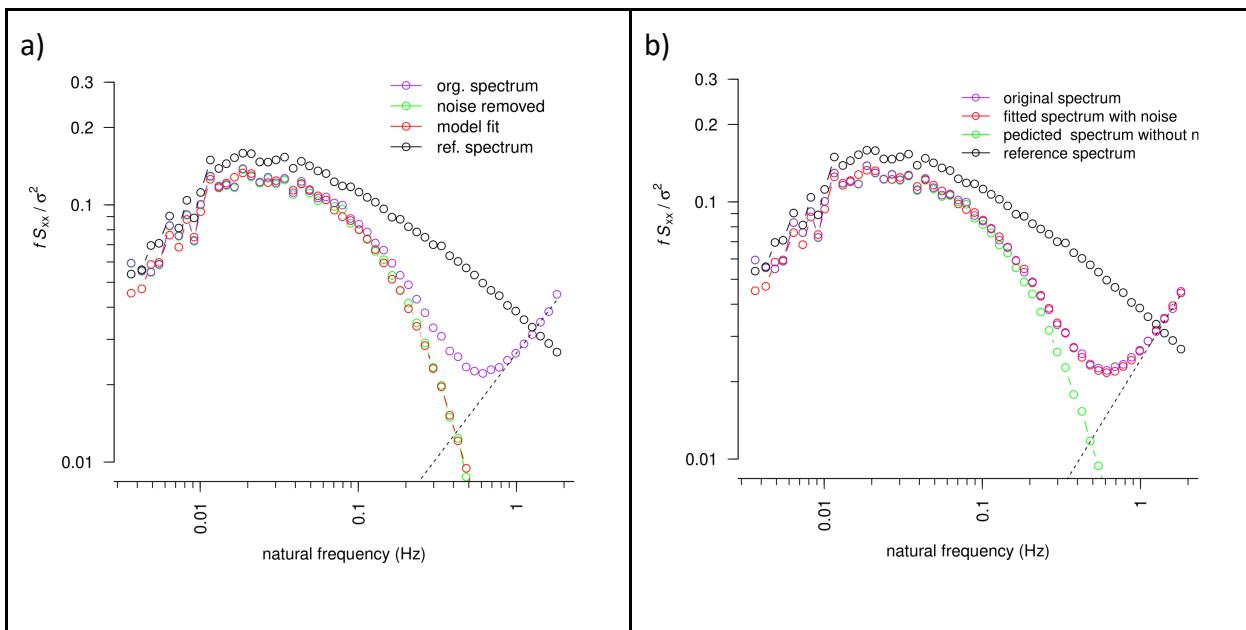


Figure 11. Example of a noisy power spectrum (CO₂, LI-6262 gas analyser, LI-COR, USA, green) and a sonic temperature power spectrum (black, Solent R2, Gill Lymington, UK) above a beech forest (June 1996, Sorø, DK-Sor) used to determine the parameters of a first-order infinite impulse response filter model (red), (a)

after Ibrom et al. (2007), (b) after Aslan et al. (2020). The dashed black line represents the effect of the white noise that is also apparent in the original spectrum.

While the example in Figure 11a shows the noise effects very clearly, there are situations, where the noise cannot be distinguished as clearly. Since the noise does not correlate with the vertical wind component, it is tempting to use co-spectra instead of power spectra for determining H_{LPF} . However, there are additional complications, especially the uncertainty in determining the time lag between the vertical wind speed and the gas mixing ratio measurements. Desynchronization between the two data pairs for calculating co-spectra can have an effect that is very similar to, if not indistinguishable from spectral attenuation. From this, the objective of this analysis is to investigate when which of the two approaches should be applied rather than generally prioritise one over the other.

For a systematic analysis, artificial data were generated based on a set of unattenuated turbulent sonic temperature (T_s) time series. The time series were low-pass filtered in time domain with increasing filter constants and white noise was added to compensate for the loss of variance. By this, a fully factorial set of test data was available. To consider the imperfection of a single throw of the random noise generator, the numerical experiment was repeated 100 times and all data were analysed. For details, see Aslan et al. (2020). The spectral transfer function (H_{LPF}) was then estimated with the power spectral approach (PSA) and the co-spectral approach (CSA). As we work with an infinite impulse recursive filter (IIR), H_{LPF} for power spectra is equal to the Lorentzian transfer function H (Ibrom et al., 2007):

$$H = \frac{1}{1 + \left(\frac{f_c}{f}\right)^2} = \frac{1}{1 + (2\pi\tau f)^2}, \quad (8)$$

where f_c is the cut-off frequency in Hz and τ is the filter time constant in s.

Based on the results, we developed the approach by Ibrom et al. (2007) further. Their two-step approach removed the noise first and then the pure attenuated spectrum was used to fit the spectral model (Figure 11a). We instead used the measured attenuated spectrum and describe the noise in the regression model:

$$f \frac{S_c}{\sigma_c^2} = f \frac{S_r}{\sigma_r^2} F_n H + fN, \quad (9)$$

where index c stands for the attenuated scalar and index r for the reference spectrum, i.e. usually the sonic temperature power spectrum (Aslan et al, 2020). F_n is the normalization factor (Ibrom et al, 2007) and N stands for a noise term. For white noise, N is constant, i.e., when multiplied with f and expressed as a function of f , a function with a slope equal to 1 with a y axis intercept of N (see stipulated line in Figure 11). Multiplication with f on both sites is necessary to separate the pure noise effect (high-frequency domain) from the turbulent signal (dominating the low and medium frequency domain).

We then used thus two alternative PSA for the estimation of H : PSA_{107} following Ibrom et al. (2007) or the improved approach PSA_{A20} following Aslan et al (2020).

Alternatively, we also estimated H_{LPF} with three different co-spectral approaches (CSAs). These CSA were using Horst's suggested $H_{LPF} = H$ (eq. (7), "CSA_H", Horst, 1997), using $H_{LPF} = \sqrt{H}$, without (CSA_{sqrt(H)}) and with shifted w time series (CSA_{sqrt(H),sync}), i.e. shifting the w time series to achieve maximum covariance between w and c .

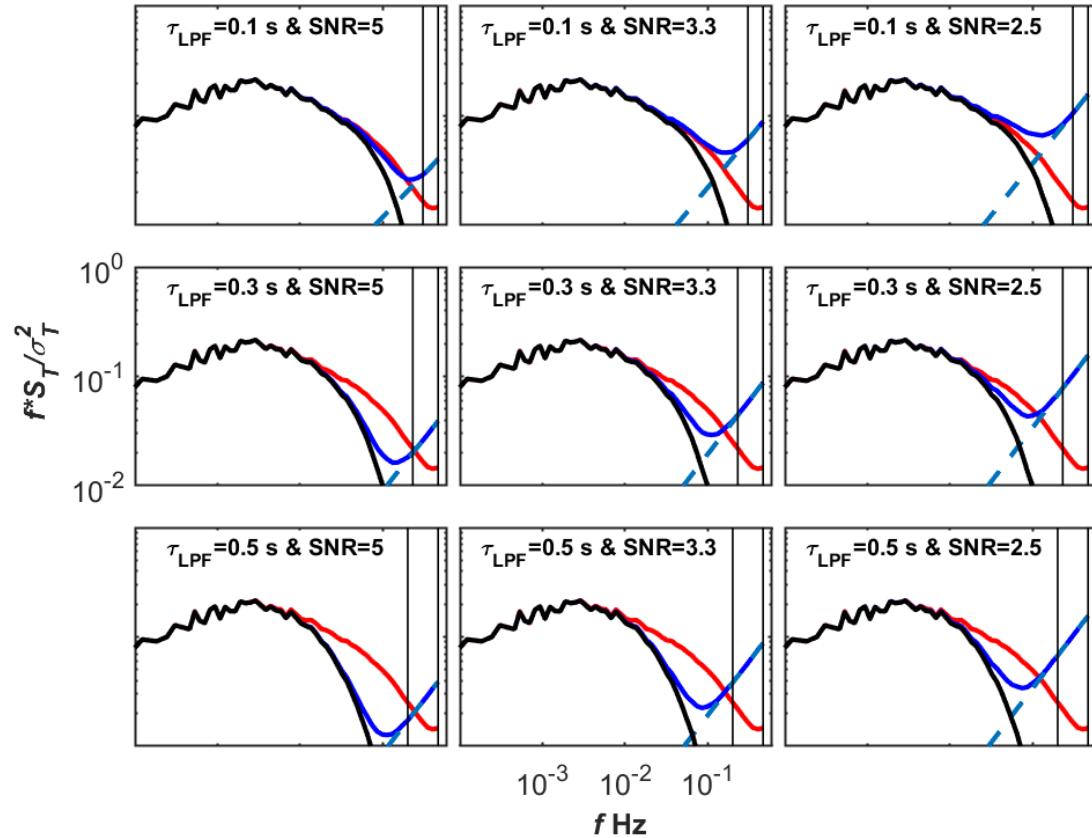


Figure 12. Normalised ensemble power spectra ($n=70$) of artificial sonic temperature time series with increasing attenuation (τ) and signal to noise ratio (SNR) (Aslan et al., 2020). The red curves show the reference spectra, the solid blue curves the measured spectra, the dashed blue lines the fitted noise, and the black curves show the measured spectra after noise removal. The vertical lines mark the frequency range used for fitting for noise removal. The lower thresholds of the frequency range are 3, 2.3 and 2 Hz for the attenuation levels of 0.1, 0.3 and 0.5 s, respectively.

Figure 12 illustrates the effect of different degrees of low-pass filtering ($\tau = 0.1, 0.3,$ and 0.5 s) and SNR (5, 3.3 and 2.5) conditions on ensemble-averaged power spectra of sonic temperature (T) from 70 half-hour raw data sets. These are shown on a double-logarithmic scale, where f is natural frequency, S_T is spectral density, σ_T^2 is the variance of the un-attenuated T series. The measurement height was 3 m and the average wind speed was 2.1 ms^{-1} . Shown are the power spectra of raw measured sonic temperature (red lines), and spectra of low-pass filtered and noise superimposed sonic temperature before (dark blue) and after (black) noise removal with the method described in Ibrom et al. (2007). The blue dotted lines are the curves fitted to the high-frequency end of blue line, then extended towards lower frequencies. The vertical lines mark the frequency range used for fitting for noise removal (i.e., 3, 2.3 and 2 Hz for the attenuation levels of 0.1, 0.3 and 0.5 s, respectively).

For comparison, some co-spectra are shown in Figure 13. Contrary to the power spectra, the noise does not appear, despite an even lower SNR. Because the noise does not correlate with the vertical wind speed (w), co-spectral analysis acts as a filter for noise.

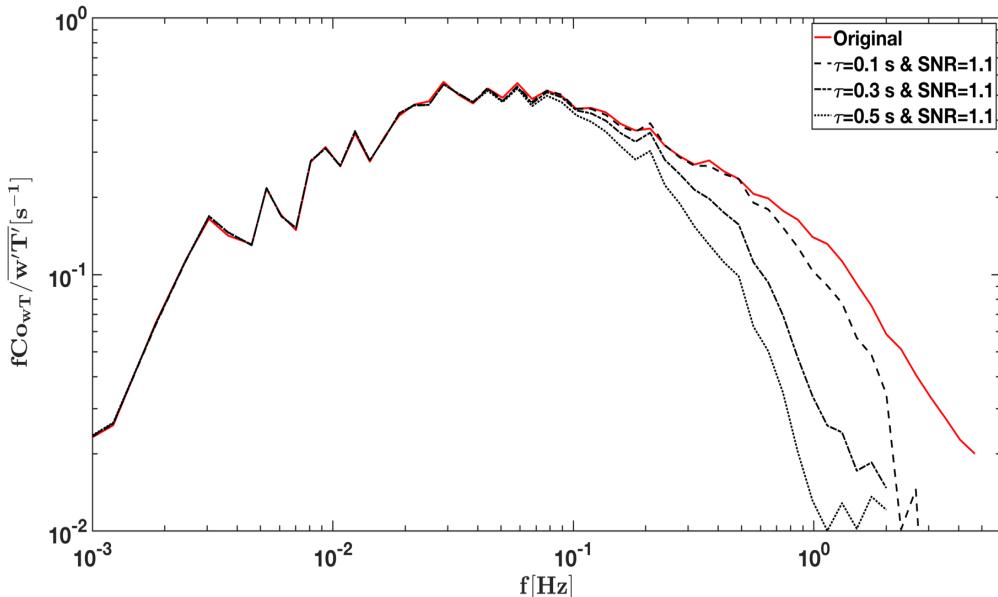


Figure 13. Normalised ensemble co-spectra for the original and differently attenuated time series. White noise was added at a signal to noise ratio (SNR) of 1.1.

The results of this analysis are summarised in Figures 14 and 15 for PSA and CSA approaches, respectively. The upper panel of Figure 14 shows the variability of the estimated filter time constant τ when applying the five different τ (panels) and adding noise with different SNR (x axis). The PSA₁₀₇ generally approaches the expected τ , but fails slightly at low low-pass filtering intensity (small τ) and at high SNR. According to our analysis, the estimation of the correct H seems to be perfectly solved with the new approach (Figure 14 b). In all cases, PSA_{A20} yields the correct τ values with only little deviance. It is not unlikely that the

deviance might be caused by the uncertainty of the white noise generator, rather than by the $\text{PSA}_{\text{A}20}$ approach. The difference in the results is caused by the assumptions made over the frequency response of the noise. In the $\text{PSA}_{\text{I}07}$ approach, the dependence of the noise of frequency (i.e. the slopes of the dashed blue lines in Figure 12) are fitted to the data at the high-frequency end of the spectrum, whilst in the improved $\text{PSA}_{\text{A}20}$ the slope is fixed to +1 through Eq. (8). At low noise conditions the $\text{PSA}_{\text{I}07}$ approach derives a slope < +1 which results in erroneous additional noise removal at lower frequency.

Integrating the noise into the regression model of Eq. (8), solves this issue and also overcomes problems in deriving the noise slope in the first place, sometimes accounted with the $\text{PSA}_{\text{I}07}$ approach. The spectral domains to estimate the noise and H is now found automatically and with large robustness.

Figure 15 show the results from three different CSA. Using the same H for co-spectra as for power spectra, as suggested by Horst (1997) yields systematically too low τ estimates. Using \sqrt{H} , instead, strongly overestimates τ . Only if using the lag-time shifted time series in combination with \sqrt{H} , the correct τ is approached, however, compared to the improved PSA approach ($\text{PSA}_{\text{A}20}$) with larger bias and less precision.

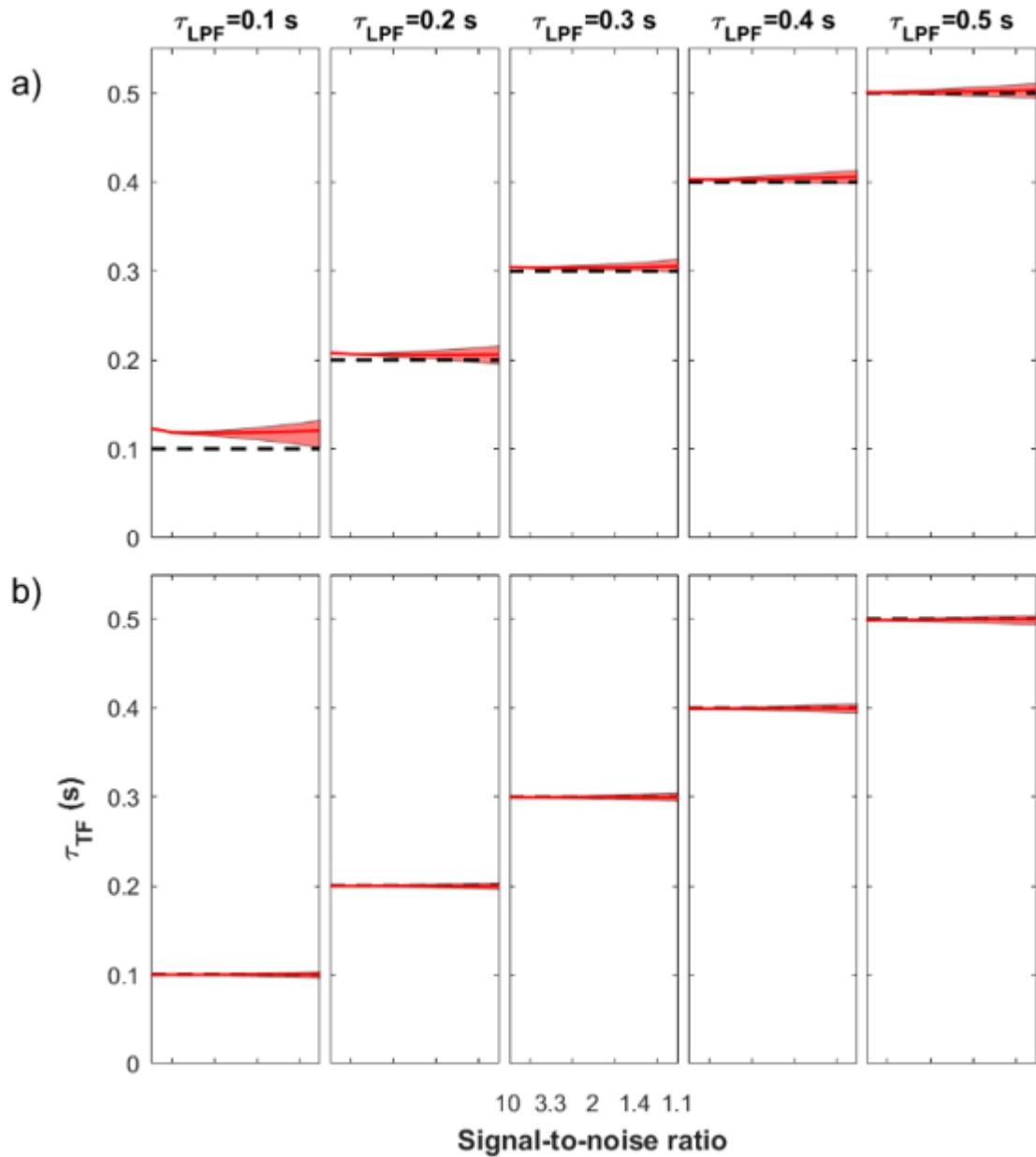


Figure 14. Comparison of the estimated spectral filter time constant (τ) based on a) power spectra after Ibrom et al. (2007) and b) using the improved power spectral approach (Aslan et al. (2020)). The black stipulated lines characterise the values of τ which were used to generate the attenuated time series.

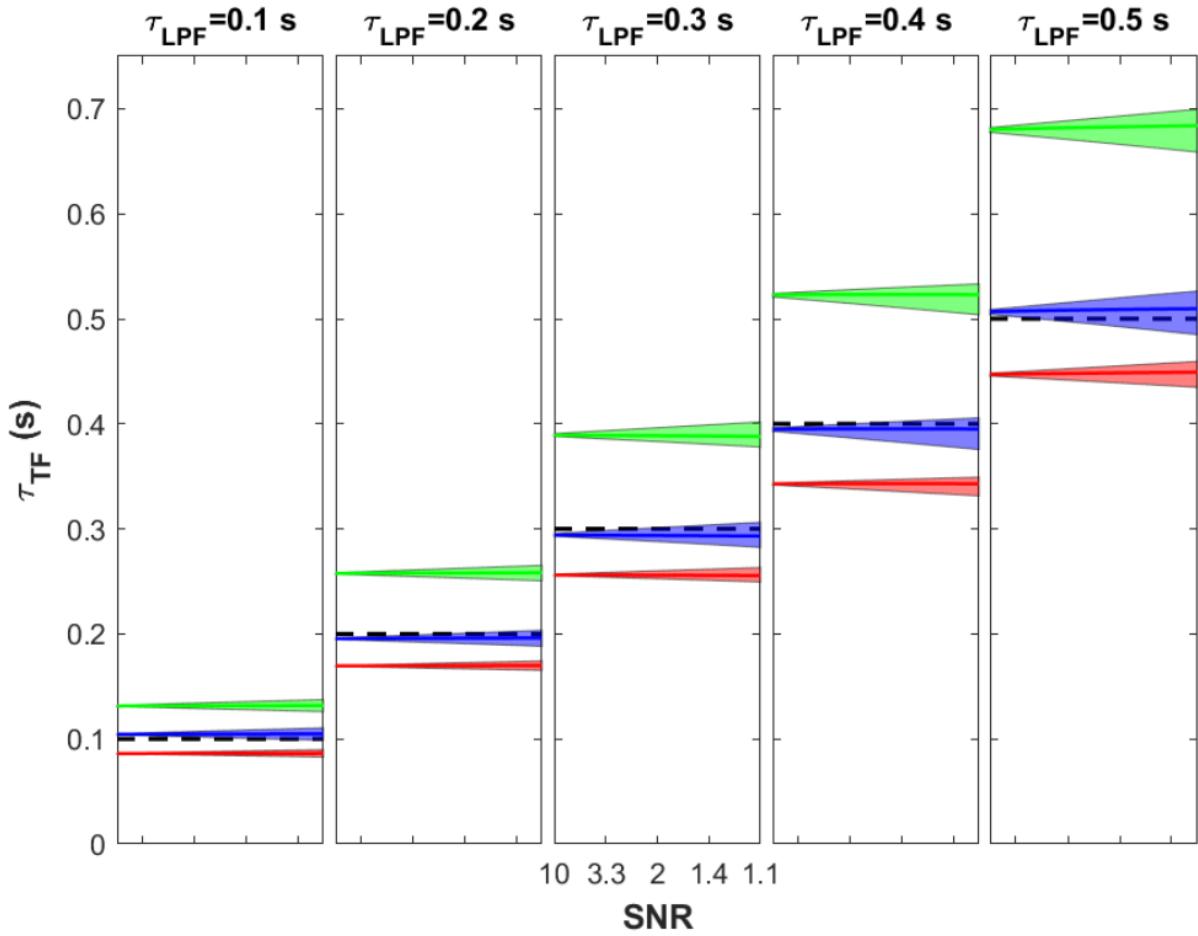


Figure 15. Comparison of the estimated spectral filter time constant (τ) based on cospectra with CSA_H , $CSA_{\sqrt{H}}$ and $CSA_{\sqrt{H}}, sync$ in red, green and blue, respectively. The black stipulated lines characterise the values of τ which were used to generate the attenuated time series (Aslan et al., 2020).

The resulting uncertainty of annual flux estimates caused by the choice of the method to estimate H have been investigated using data from the Siikaneva fen site in Finland (Peltola et al., 2013). We showed that the CH_4 fluxes corrected using the PSA_{107} based time constants were overestimated between 0 and 12% (on average 4%), while the PSA_{A20} approach showed no bias. On the other hand, fluxes corrected using the CSA_H based time constant were underestimated up to 8%, while the $CSA_{\sqrt{H}}$ overestimates the fluxes up to 22%, the bias of which was further improved with $CSA_{\sqrt{H}}, sync$, varying between $\pm 2\%$.

6.3 Determining the spectral transfer function and correction factor with cospectra and its interdependence with the time lag estimation

The results of the CSA from the above study pointed to fundamental challenges when working with cospectra from attenuated time series, with and without sensor noise. First of all, literature suggests two

different ways to deal with the spectral transfer function (H_{LPF}). While a majority of publications suggest using $H_{LPF} = \sqrt{H}$ for co-spectra (Moore 1986, Moncrieff et al. 1997, Ibrom et al. 2007, Fratini et al., 2012), Horst (1997) suggested using $H_{LPF} = H$ for both power spectra and co-spectra, which was used by Eugster and Senn (1995), Mammarella et al. (2016), Aubinet et al (2000), Aubinet et al (2012) and Hunt et al. (2016) who successfully motivated the respective change of EddyPro, an open-source eddy covariance post-processing software, sponsored by LI-COR (Lincoln, USA) and developed by G. Fratini and others. On the other hand, all these publications have neglected the phase shift caused by low-pass filtering (Ibrom et al., 2007b).

The existence of a time lag due to low-pass filtering Δt_{LPF} complicates the correct processing in two ways. The questions arise, first, whether the time series of the scalar and the vertical wind speed has to be shifted by Δt_{LPF} or not prior to calculating cospectra, and second, when the time series is shifted, what does this mean for the amplitude attenuation? So far, the above approaches do not explicitly mention how phase shift effects were considered.

In the case co-spectra are used for the estimation of the system response time and related flux correction factor, we provide a new CSA approach which specifically accounts for the interaction between low-pass filtering induced phase shift and high frequency attenuation (Peltola et al., 2020). In particular, we show that if the CCM between the scalar and w is used, the correct transfer function H_{LPF} is H^*Hp , where $Hp = \cos \phi - \omega \tau \sin \phi$ is the part related to the phase shift $\phi = \omega \Delta t_{LPF}$ and ω is the radial frequency. The new transfer function provides the correct estimation of τ , Δt_{LPF} and the flux correction factor. Note that when using only H as done in Aubinet et al (2000) and Mammarella et al(2016), the system response time is underestimated, but the flux correction factor is very close to the ones given by using H^*Hp . We further demonstrated that H^*Hp is well approximated by \sqrt{H} . This is consistent with results by Aslan et al (2020) and practically means that if the correct estimate of τ (e.g. as given either by PSA_{A20}, by H^*Hp or by the overflow inlet system) is used and the flux is calculated by using CCM, the flux correction factor has to be calculated either taking H^*Hp or \sqrt{H} as low pass filter transfer function H_{LPF} in eq. 7. In this case analytic formulas for CF as proposed by Massman (2000) and Horst et al (1997) likely give biased results. For more details, please refer to Peltola et al. (2020).

6.4 Conclusions and recommendations

The analysis showed that subtle differences in four published approaches to estimate the spectral transfer function (H_{LPF}) from measured power or co-spectra yielded considerably deviating and biased results. Generally, the approaches using power spectra resulted in estimates closer to the real H , with an improved

approach that includes both H and noise in one fitting approach yielding the best results. Considerable, systematic differences were found between the results from three approaches based on co-spectra, where even the choice of the shape of the H differs in literature. These differences point to a lack of fundamental understanding on how to interpret co-spectra between vertical wind speeds and attenuated scalars.

As a final remark, when using power spectra, where some noise is evident on the attenuated scalar measurement, i.e. practically any closed-path eddy covariance system, the new fitting approach $\text{PSA}_{\text{A}20}$ is recommended as the most accurate, precise and robust method to empirically estimate the time constant of H (Aslan et al., 2020). Then CF is calculated following the Fratini et al (2012) approach, e.g. using $H_{LPF} = \sqrt{H}$ in Eq. 7. On the other hand, we recommend to use H^*H_p (or the approximation \sqrt{H}) for the correct estimates of τ and CF when the cross-covariance maximization and cospectra are used (Peltola et al., 2020).

7. FILTERING OF 30 MIN FLUXES BASED ON FRICTION VELOCITY

The impact of friction velocity (u_*) filtering is potentially large at high-flux sites, especially fertilized agricultural sites where episodic non-CO₂ emission bursts are often recorded after fertilizer application or rain events. Flux data availability around such high-emission events needs to be as complete as possible for an accurate gap-filling of flux data gaps. At low-flux sites and in the absence of episodic peak emissions, the impact of u_* filtering is generally expected to be low. At such sites, flux values are close to the detection limit and often fluctuate around zero throughout the year, i.e. even many data points lost due to u_* filtering can be gap-filled with relatively high confidence.

Therefore, it is advisable to keep data loss due to u_* filtering to a minimum, in particular for agricultural sites. The current approach in FLUXNET is to calculate the final yearly u_* threshold for NEE as the maximum of 4 respective seasonal thresholds (Pastorello et al., 2020). This approach has the disadvantage that measurements that can be considered good data with regard to the respective seasonal thresholds are rejected because u_* was below the yearly threshold. In addition, the computation of the yearly threshold is potentially driven by a single season that is of lesser interest in regard to fluxes. For example, at the grassland site Früebüel in Switzerland ([CH-Fru](#)) the calculation of seasonal thresholds for 2018 by means of the MPT method (100 bootstrap runs, thresholds are given as the 50% quantile of the bootstrapped uncertainty distribution, Wutzler et al. 2018) was found at 0.17 m s⁻¹ for the time period between September and November (autumn), but at 0.04 m s⁻¹ between June and August (summer) of

the same year. Following the FLUXNET approach and applying the autumn threshold to summer fluxes results in a loss of 37% of recorded NEE values, compared to a loss of only 4% when the summer threshold is applied (Figure 16). It is important to note that this data loss comes in addition to data rejected by the steady state and integral turbulence characteristics tests (Foken and Wichura, 1996). Generally, seasonal u_* thresholds fluctuate over the course of a year, and yearly thresholds as the maximum of these seasons are therefore potentially significantly different from respective seasonal thresholds (Figure 17).

To minimize data loss due to the u_* filter it is therefore advisable to use seasonal thresholds detected for the CO₂ flux instead of one single threshold for the complete year. The robustness of seasonal u_* thresholds could be increased by pooling seasonal data from adjacent years before calculations. For example, to calculate the summer threshold for 2018 for a given site, the threshold computation would be run using the pooled summer data 2017, 2018 and 2019. This approach is somewhat similar to the calculation of the variable u_* threshold (VUT) described for the FLUXNET2015 dataset, where a yearly threshold is extracted from the joint population of the thresholds found for the respective year and the thresholds found for adjacent years (Pastorello et al., 2020).

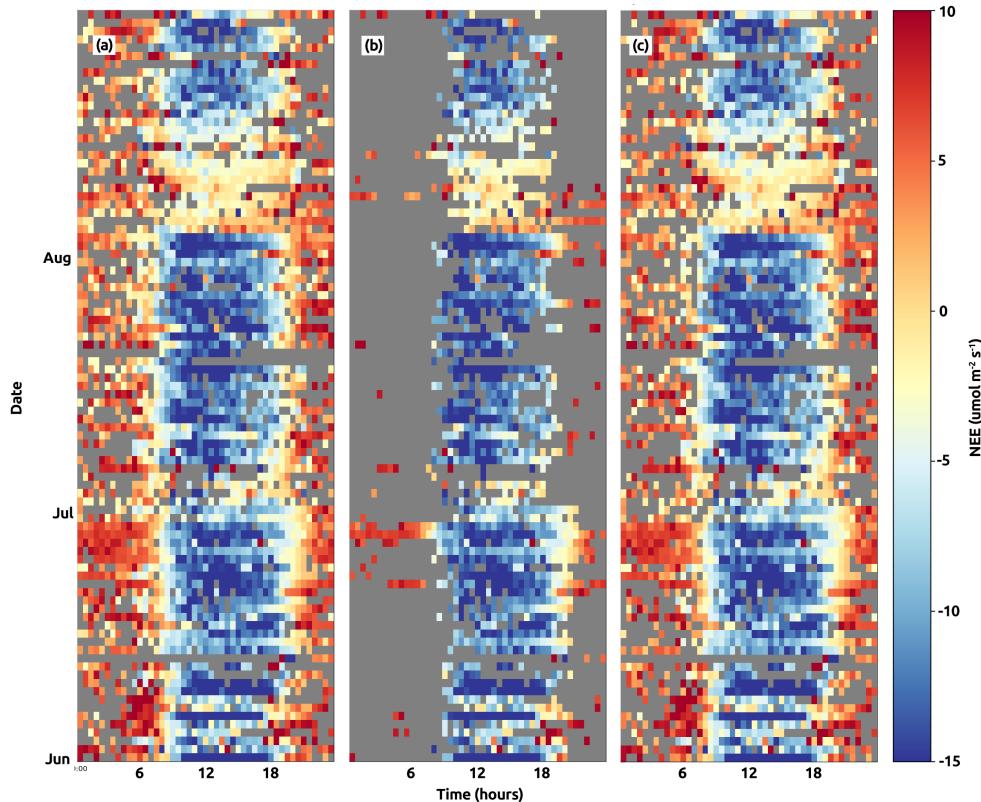


Figure 16. Impact of applying a yearly u_* threshold, calculated as the maximum of 4 seasonal u_* thresholds, on summer NEE data at CH-Fru in 2018. (a) measured data after removal of data that failed the steady state and integral turbulence characteristics tests (Foken and Wichura, 1996), (b) measured data after additionally rejecting all values where $u_* < 0.17 \text{ m s}^{-1}$ (autumn threshold calculated from

September-November data, corresponds to the maximum of seasonal thresholds in 2018), (c) measured data after rejecting all values where $u_* < 0.04 \text{ m s}^{-1}$ (summer threshold calculated from June-August data).

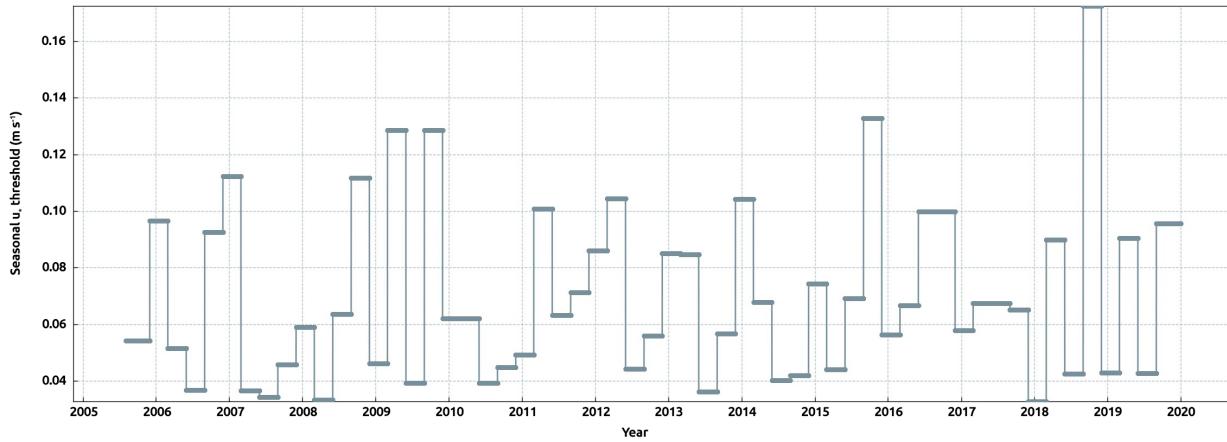


Figure 17. Seasonal u_* thresholds at CH-Fru between 2005 and 2019. Thresholds are given as the 50% quantile of the bootstrapped uncertainty distribution from 100 bootstrap runs.

8. FLUX GAP-FILLING

8.1 Background

Gap-filling techniques are well established and standardized for CO₂ fluxes (Moffat et al., 2007, Papale et al., 2006), but not for the other trace gases. Gap-filling N₂O fluxes is challenging due to their dependence on multiple drivers as well as their spatial and temporal heterogeneity (Butterbach-Bahl et al., 2013; Cowan et al., 2015). Especially at agricultural sites, N₂O fluxes are often characterized by sporadic high-emission events, while fluxes in between those events remain low and often below the limit-of-detection of the applied analyser, which in turn further exacerbates the identification of viable input parameters for gap-filling methods. Even at agricultural sites such as intensively managed grasslands, where high N₂O quantities can be emitted during and after management events such as fertilizer application and ploughing, fluxes are typically low during the rest of the year (Hörtnagl et al., 2018; Merbold et al., 2014). It is therefore unlikely to be able to define a single set of drivers that can reliably be applied to gap-fill N₂O fluxes across sites and ecosystems. However, in recent years the random forest method has emerged as a useful method not only for gap-filling time series data, but also for feature selection in modelling approaches (e.g., Kim et al., 2019; Kursa and Rudnicki, 2010). In this chapter, we tested random forests for gap-filling and feature selection, and then used selected features in the widely applied MDS gap-filling method (Reichstein et al., 2005).

For CH₄ emissions, since the drivers of the underlying processes are more complex than those of CO₂ fluxes, artificial neural networks seem to be promising with an overview of studies found in Knox et al. (2016).

Here, we tested different gap-filling approaches from simple interpolation methods to machine learning algorithms for N₂O and/or CH₄.

- Interpolation methods:
 - SLI – simple linear interpolation
 - SRAM – simple running arithmetic mean
 - SRMED – simple running median
 - MDC – Mean diurnal course after Falge et al. (2001)
 - MDA - Mean diurnal averaging after Reichstein et al. (2005)
- Look-up tables:
 - LUT – look-up tables binned to one to three variables V1 to V3 in a time window of ±3 or ±7 days after Falge et al. (2001)
 - MDS – a LUT combined with MDA in marginal distribution sampling (MDS), a gap filling scheme adopted for CO₂ fluxes after Reichstein et al. (2005)
- Machine learning algorithms:
 - RF – random forests (Breiman 2001)
 - ANN – artificial neural networks based on backpropagation (Moffat et al., 2010)

8.2 N₂O fluxes

Gap-filling approaches were tested mainly for N₂O fluxes measured at the grassland site Chamau ([CH-Cha](#)) between 2013 and 2015. The site is intensively managed with up to six cuts per year, multiple fertilizer applications and occasional grazing (Fuchs et al., 2018; Merbold et al., 2014). Data coverage for N₂O fluxes during the measurement period was high (50,341 of 52,560 half-hours available, data coverage 95.8%). After quality control, 30,390 measured half-hours (57.8%) were available.

Among the tested gap-filling approaches were (1) marginal distribution sampling (MDS) following the approach by Reichstein et al. (2005), (2) random forest (RF), an ensemble learning method based on decision trees (Breiman, 2001), (3) simple linear interpolation (SLI), (4) simple running arithmetic mean (SRAM) and (5) simple running median (SRMED).

Gaps in the data series were filled using different approaches, with several methods arriving at similar overall N budgets (Table 3). The resulting N₂O budget for the complete 3-year time period and applying RF was found at 15.23 kg N₂O-N ha⁻¹. In RF, a large number of decision trees operate as an ensemble: the most "voted for" class prediction becomes the model prediction. This approach worked well not only for gap-filling N₂O data (even in the presence of emission peaks, see Figure 18 results from 100 decision trees/estimators and Figure 19 for a comparison of measured and predicted fluxes), but also for model feature selection from large auxiliary datasets. During our tests, RF identified the variable *time since slurry application* (tsSA) as the most important model feature (model importance: 0.16), which highlights the importance to collect and share management data. Other important factors were: soil water content (SWC) and air temperature (TA; both 0.11), soil temperature (TS; 0.10), shortwave incoming radiation (SW_IN; 0.09) and vapor pressure deficit (VPD; 0.08).

The three variables with the highest RF model importance were selected to generate the lookup table for the MDS method, with the exception that TS was used instead of TA due to its more direct role in the production and consumption of N₂O (Butterbach-Bahl et al., 2013). The MDS algorithm was tested using the following settings: in MDS(1), similar meteorological conditions were defined by dividing SWC and TS into 20 evenly distributed lookup bins, which corresponds to similarity limits of SWC ±1.6 % and TS ±1.4°C. Conditions for days since slurry application were considered similar in a time window of tsSA ±2.4 days. Across the 3-year period, the MDS(1) budget amounted to 15.03 kg N₂O-N ha⁻¹ (Table 3). Similar results were found for MDS(2), using the same settings as MDS(1) with the exception that the lookup window for tsSA was slightly larger at ±5 days (14.94 kg N₂O-N ha⁻¹). Since VPD scored high during the RF feature selection and is used by default in MDS gap-filling of CO₂ fluxes, it was included in MDS(3) in combination with the soil parameters SWC and TS. Similar to SWC and TS, similarity limits for VPD were calculated so that measured data are evenly divided into 20 similarity windows. VPD conditions were considered similar in the range of ± 211 Pa. Over 3 years, MDS(3) yielded the most similar result to RF (15.30 kg N₂O-N ha⁻¹). To complete the MDS tests, the typical CO₂ drivers SW_IN, TA and VPD were applied to gap-fill N₂O fluxes, using CO₂-typical similarity ranges of ±50 W m⁻², ±2.5°C and ±50 Pa, respectively, in MDS(4). The resulting budget showed slightly higher N₂O emissions (15.53 kg N₂O-N ha⁻¹) than MDS(1)-(3) (Table 3). To some extent, the application of MDS with CO₂ drivers is counter-intuitive since the three input drivers had relatively low predictive power for observed N₂O exchange rates. However, since MDS applies look-up tables it seems to make sense that it nevertheless yields annual budgets that fall into expected ranges.

In addition to RF and MDS, simpler methods were tested to identify the potential range of budgets calculated with the various approaches. Budgets calculated with the straight-forward simple linear

interpolation (SLI, with no gap size limit) yielded budgets similar to RF and MDS ($15.18 \text{ kg N}_2\text{O-N ha}^{-1}$; Table 3), but is not recommended due to its high volatility to the length and distribution of gaps. Therefore, the risk to significantly over- or underestimate the true budget is high. However, SLI is potentially useful in connection with other gap-filling methods, e.g. to fill the smallest gaps in the dataset (e.g. 1-2 missing half-hourly values) before proceeding with other methods. In our test setup, SLI yielded virtually the same result for the grassland, mainly because (1) the available data around fertilization events was relatively complete, (2) there were no major gaps throughout the 3-year period, (3) high fluxes were underestimated and (4) low fluxes were overestimated.

The simple running arithmetic mean (SRAM) in a time window of $\pm 2.5\text{d}$ (centered), short enough to capture peak dynamics, produced results similar to RF when the gap distribution of the dataset after quality control was approx. evenly distributed over the year and no major gaps (longer than five days, i.e. ± 2.5 days) were present in the dataset ($15.17 \text{ kg N}_2\text{O-N ha}^{-1}$). Gap-filling with the simple running median (SRMED, centered time window) yielded the lowest estimates of all methods due to the insensitivity of the median to extreme values.

Table 3. Results of selected gap-filling runs for N₂O fluxes 2013 – 2015 measured at CH-Cha.

Method	Drivers	Time window (days)	Min values	Emissions (kg N ₂ O-N ha ⁻¹)			
				2013	2014	2015	2013 – 2015
RF	P, PAR, SW_IN, SWC, TA, TS, VPD, tsP, tsSA, SA	sliding	-	5.99	5.85	3.39	15.23
MDS(1)	SWC, TS, tsSA (2.4d)	increasing, sliding	-	5.91	5.72	3.40	15.03
MDS(2)	SWC, TS, tsSA (5d)	increasing, sliding	-	5.90	5.68	3.36	14.94
MDS(3)	SWC, TS, VPD	increasing, sliding	-	5.92	5.93	3.44	15.30
MDS(4)	SW_IN, TA, VPD	increasing, sliding	-	5.97	6.03	3.53	15.53
SLI	-	-	-	5.92	5.68	3.58	15.18
SRAM	-	20	96	6.15	6.18	3.38	15.72
	-	2.5	12	5.91	5.92	3.33	15.17
SRMED	-	20	96	5.19	5.55	2.76	13.50
	-	10	48	5.26	5.56	2.85	13.67
	-	5	24	5.29	5.57	2.89	13.75
	-	2.5	12	5.30	5.58	2.91	13.79

Conditions within a time window were considered similar using the following settings: (1) SWC $\pm 1.6\%$, TS $\pm 1.4^\circ\text{C}$, tsSA ± 2.4 days (2) SWC $\pm 1.6\%$, TS $\pm 1.4^\circ\text{C}$, tsSA ± 5 days; (3) SWC $\pm 1.6\%$, TS $\pm 1.4^\circ\text{C}$, VPD ± 211 Pa; (4) SW_IN ± 50 W m⁻², TA $\pm 2.5^\circ\text{C}$, VPD ± 50 Pa, typically used for gap-filling CO₂ fluxes.

Abbreviations: Methods: RF: random forest, MDS: marginal distribution sampling, SLI: simple linear interpolation, SRAM: simple running arithmetic mean, SRMED: simple running median. Drivers: P: precipitation, PAR: photosynthetically active radiation, SW_IN: incoming short-wave radiation, SWC: soil water content (5 cm), TA: air temperature, TS: soil temperature (5 cm), VPD: vapor pressure deficit, tsP: time since precipitation, tsSA: time since slurry application, SA: slurry application.

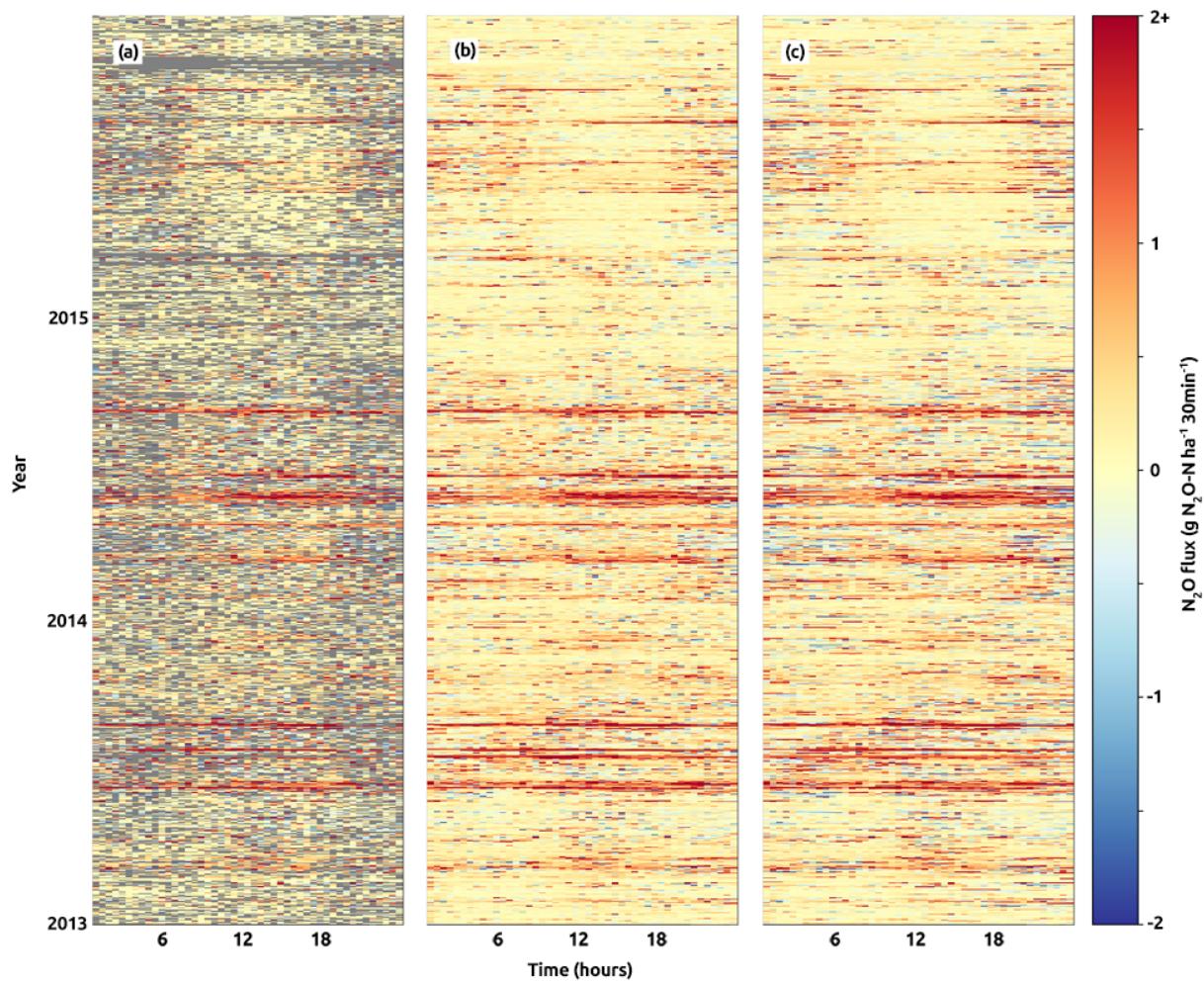


Figure 18. Comparison of half-hourly N_2O fluxes between 2013 and 2015 at the intensively managed grassland site CH-Cha. Shown are (a) fluxes after quality control, (b) fluxes after gap-filling using random forest, (c) fluxes gap-filled using the MDS method (Reichstein et al., 2005) in similarity classes of $\text{SWC} \pm 1.6\%$, $\text{TS} \pm 1.4^\circ\text{C}$ and $\text{tsSA} \pm 2.4$ days. Emission peaks were observed after the application of slurry to the grassland (red colours).

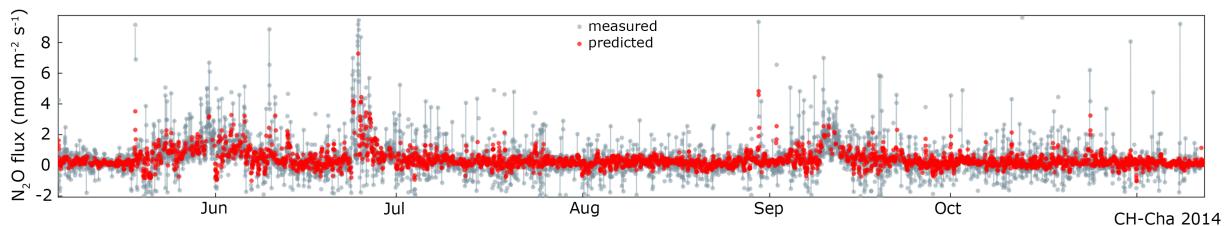


Figure 19. Gap-filled (random forest, red) and measured (after quality control, grey-blue) half-hourly N_2O fluxes measured at CH-Cha. Emission peaks were observed after the application of slurry to the grassland. Slurry application dates: 26 May 2013, 25 Jun 2014, 9 Sep 2014.

Gap-filling methods were also tested for N₂O fluxes from the ICOS Class 1 site Davos ([CH-Dav](#)). CH-Dav is characterized by weak N₂O emission throughout the year and with fluxes often below the detection limit of the gas analyser. Between January 2016 and May 2017, a total of 20,252 out of 24,875 half-hourly N₂O fluxes were recorded (data coverage 81.4%), of which 9,612 (38.6%) were available after quality control. Budgets calculated for the whole measurement period were similar across methods, with 0.31 kg N₂O-N ha⁻¹ (RF), 0.32 kg N₂O-N ha⁻¹ (MDS) and 0.35 kg N₂O-N ha⁻¹ (SRMED, time window 40 days with a minimum of 192 available values). Due to the high frequency of data gaps, SLI did not produce a realistic budget.

Conclusions

Recently, the random forest method was already successfully tested on CH₄ data (Kim et al., 2019). At the moment, no such comparison exists for N₂O eddy covariance fluxes, mainly due to the lack of multi-year eddy covariance datasets that would allow for testing across multiple sites and ecosystems. Due to its dynamic approach, RF is expected to perform well for N₂O fluxes for a wide range of different conditions because of its ability to account for the dependence of N₂O fluxes on multiple drivers. The comparison presented here stops short of testing the performance of the different approaches on datasets with artificial gaps. These performance tests constitute the next logical step in finding the “best” method for gap-filling, but eddy covariance datasets spanning multiple years of N₂O fluxes are urgently needed. For the time being, the MDS approach using the soil parameters SWC and TS in combination with a third parameter (either tsSA, if available, or VPD) seems to produce robust results, at least for relatively complete datasets such as the CH-Cha data. However, the MDS performance will decrease significantly in the presence of longer gaps due to a lack of lookup values during the respective time period. In such a scenario, the RF method has the potential to provide reasonable gap-filling results, since the algorithm can be trained on datasets spanning multiple years.

8.3 CH₄ fluxes

Besides ANNs (which have shown to be suitable for gap-filling of methane), we also tested simpler gap filling techniques that are easier to implement. As a test dataset, three months of half-hourly CH₄ fluxes from the wetland site Skjern in Denmark were used (Herbst et al., 2011). To determine the performance of the gap filling techniques, artificial gaps with a length of full days were superimposed on the dataset and filled. A total of 999 bootstrapping samples was taken and the analysis carried out separately for all half-hours in a day (“full-time”), only daylight data (“day-time”), and only “night-time” data (Figure 20)

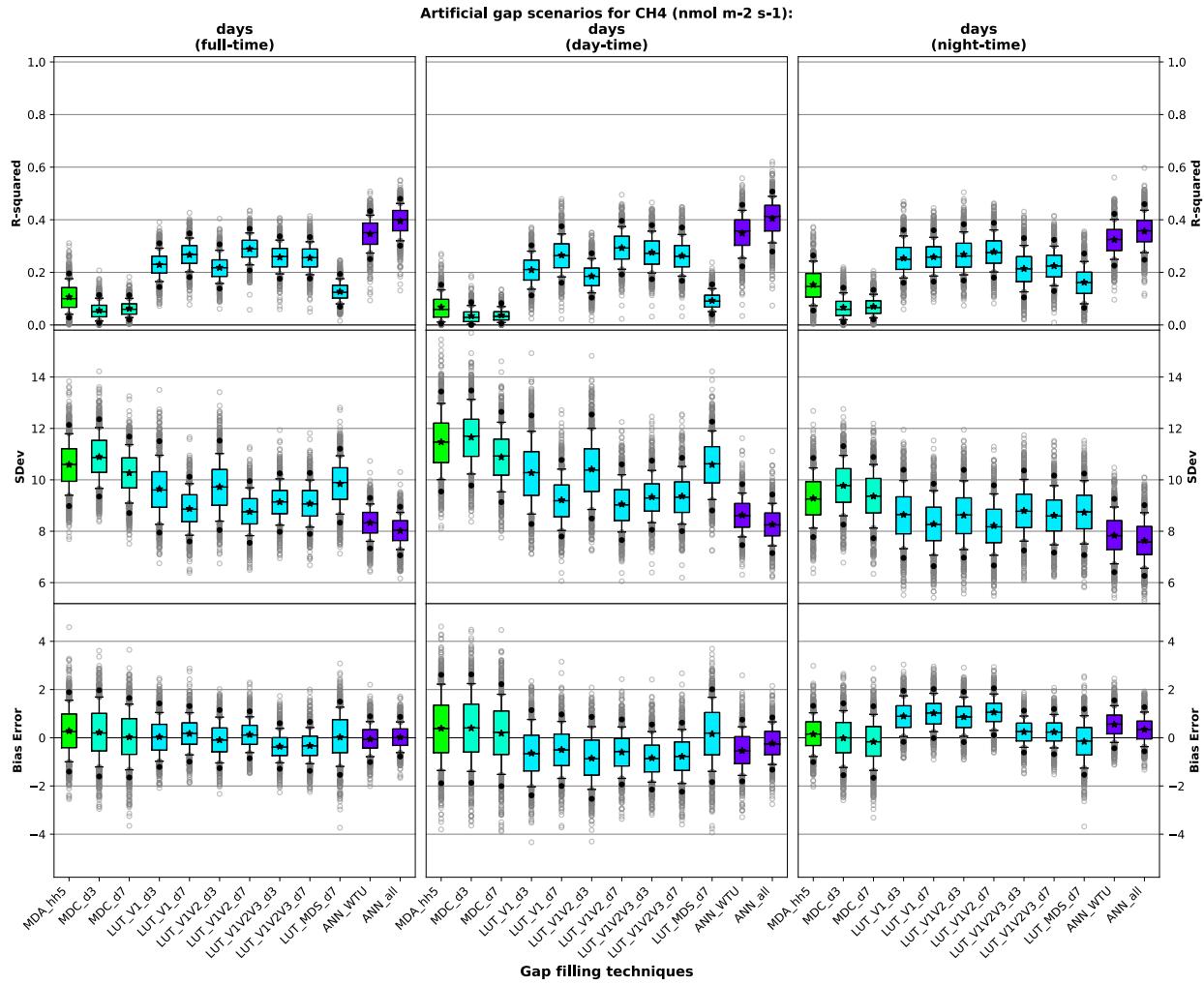


Figure 20: Performance measures for the various gap filling techniques for full-time (left), day-time (middle), and night-time (right). The boxplots are composed of the median (solid line), mean (star symbol), lower and upper quartile bounds (box), 10th and 90th percentiles (marks), and all outliers (dots) from the 999 bootstrapping results.

The performance for the three diurnal interpolation methods on the daily gaps was quite low with mean coefficient of determination (R^2) below 20%, standard deviations (SDev) larger than 10 nmol m⁻²s⁻¹ and large but centered bias errors. The look-up table from the MDS algorithm with fluxes binned to light, air temperature, and vapour pressure deficit did not work very well either, since these variables are only weakly correlated with CH₄ fluxes. To obtain a better performance of LUTs, we characterized the main drivers of the dataset using the ANN approach by Moffat et. al, 2010. The variables which showed the highest correlations with the CH₄ fluxes for this dataset were: 1) water table depth, 2) soil temperature at 20 cm depth, and 3) friction velocity. The look-up tables were tested separately for one_V1, two_V1V2, or three_V1V2V3 dependent variables. All of them had higher R^2 (20% - 30%) and smaller bias errors.

As expected, the ANNs worked best with a mean R^2 of up to 40%, SDev around $8 \text{ nmol m}^{-2}\text{s}^{-1}$ and the smallest bias for the all-day dataset. The ANNs trained on all available variables had a slightly higher performance than the one trained on the same three variables as the LUT_V1V2V3. There is a comprehensive study of different machine learning algorithms by Irvin et al. (currently in preparation) which will provide further insight in this matter for a wider range of datasets and ecosystems.

Since the CH_4 emissions at Skjern did not exhibit a strong diurnal cycle, the basic patterns of gap filling performances were similar for the three subsets of data in Figure 20. However, it is interesting to note that the LUTs and also the ANNs showed negative bias during daytime (underestimation) and positive bias during nighttime (overestimation) with neutral biases in total. This is probably due to the differences in footprint and turbulence of day versus night: The footprint is centered on methane emitting wetland during day-time but may include parts of adjacent agricultural fields without CH_4 emissions during night-time. Adding friction velocity as the third variable centered the bias at night-time. (LUT_MDS is automatically centered in this respect, since light is one of the variables used for binning which will split the data into night-time and day-time automatically by the look-up algorithm.)

9. GUIDANCE ON FLUX MAGNITUDE AND MEASUREMENT REQUIREMENTS

The ICOS Protocol (Nemitz et al., 2018) sets out the first set of criteria to help identify at which sites eddy-covariance fluxes of N_2O and/or CH_4 are required to fulfil the requirements for Class 1 operation. Partly in support of ICOS and RINGO, a meta-analysis of CH_4 fluxes from agricultural soils was undertaken based on a large existing pool of measurements for soils of the British Isles.

9.1 Importance of CH_4 exchange with N European Agricultural Soils

A meta-analysis of fluxes measured from agricultural soils across the British Isles (UK and Ireland) was carried out with the intention of establishing a better knowledge of the magnitude of emissions and uptake of CH_4 from these soil types. A total of 45,350 manual static chamber measurements from 15 different sites across the British Isles were collated (Table 4). The mean of all fluxes measured was $0.51 \pm 0.12 \text{ nmol m}^{-2} \text{ s}^{-1}$, although the majority (68%) of individual measurements are negative. The maximum CH_4 flux recorded from an individual measurement was $3054 \text{ nmol m}^{-2} \text{ s}^{-1}$, while the largest negative flux (uptake) was $-74.7 \text{ nmol m}^{-2} \text{ s}^{-1}$. Although real, observations of this magnitude are rare in the data and attributed to “hotspot” areas where soil conditions are markedly different to general field conditions (i.e. water content, organic matter content, mechanical agitation, etc.). The positive tail of the distribution of observed flux magnitude stretches widely, following a log-normal pattern; however, the vast majority of

measurements are close to the detection limits of the flux chamber methods (lower = 0.173, upper = 0.519 nmol m⁻² s⁻¹). A total of 66% of all measurement points fall below the magnitude of the upper detection limit of the method with a 95% quantile range of -1.57 to 1.84 nmol m⁻² s⁻¹ (Figure 21).

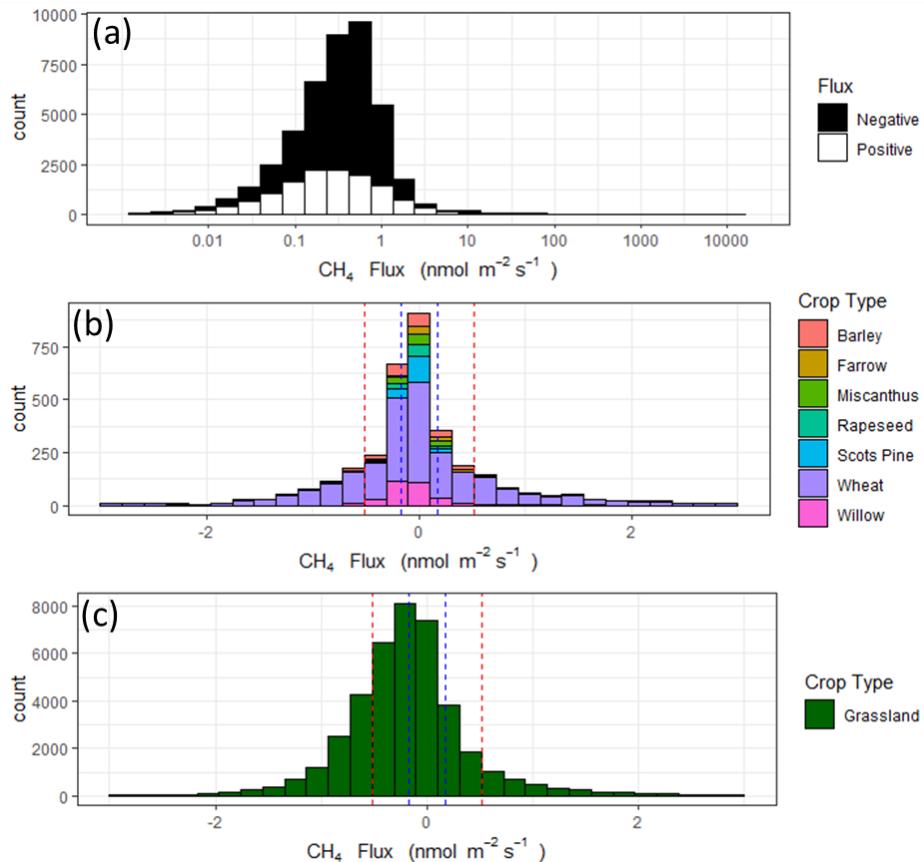


Figure 21 (a) Histogram of all CH₄ fluxes carried out using static chambers on a log-normal scale. (b) Histogram of all CH₄ fluxes measured in arable studies using static chambers. (c) Histogram of all CH₄ fluxes measured in grassland studies using static chambers. Histograms 1b and 1c are limited to range of -3 to 3 nmol m⁻² s⁻¹, where the vast majority of observations occurred. The lower (blue) and upper (red) detection limits of the flux chamber methods used across the studies are included.

Splitting the data by field site reveals that there are differences between the different field types, with uptake more likely to be observed at arable sites than grasslands (Table 4); however, these differences are not statistically different (t-test, p = 0.33) and none of the sites reports uptake of CH₄ greater than the limit of detection of the chamber methodology used to calculate fluxes. In studies where animal manures or tillage events occurred, there was usually a relatively large increase in CH₄ emissions reported. Assuming a global warming potential of 28 over 100 years (IPCC 2014, Fifth Assessment Report (AR5)), we present the estimated annual flux in units of g-CO₂eq m⁻². Although methane uptake is shown to be statistically significant from zero in some cases, fluxes of CH₄ during these periods remain small in terms of global warming potential (GWP), with the largest annual uptake observed at any site to be -1.59 g CO₂eq m⁻² year⁻¹.

$\text{m}^{-2} \text{ yr}^{-1}$. Where emission events such as manure application or tillage occur, there is a significant increase in short term emissions, lasting weeks to months. In these conditions, fluxes increase exponentially for short periods of time and chamber measurements are able to observe this behaviour.

Table 4. A summary of the origins of the CH_4 flux data presented in this study with the associated mean CH_4 reported. A GWP multiplier of 28 is used (IPCC 2014, AR5).

Site	n	Crop Type	Mean Flux (nmol $\text{m}^{-2} \text{ s}^{-1}$)	Annual GWP (g $\text{CO}_{2\text{eq}} \text{ m}^{-2}$)	Related Publication
Arable					
Norfolk	56	Farrow	1.31	13.88	Unpublished
Salisbury	65	Farrow	1.02	10.81	Unpublished
Surrey	33	Farrow	1.39	14.73	Unpublished
Lincolnshire (a)	114	Miscanthus	-0.04	-0.42	(Drewer et al., 2011)
Lincolnshire (b)	110	Rapeseed	-0.07	-0.74	(Drewer et al., 2011)
Lincolnshire (c)	109	Wheat	-0.09	-0.95	(Drewer et al., 2011)
Lincolnshire (d)	114	Willow	-0.11	-1.17	(Drewer et al., 2011)
East Grange (a)	227	Barley	-0.03	-0.32	(Drewer et al., 2017a)
East Grange (b)	184	Scots Pine	-0.03	-0.32	(Drewer et al., 2017a)
East Grange (c)	210	Willow	-0.03	-0.32	(Drewer et al., 2017a)
Boghall	1792	Wheat	0.05	0.58	(Bell et al., 2016)
North Wyke (a)	500	Wheat	0.16	1.70 ^(*)	(Sánchez-Rodríguez et al., 2018)
Grasslands					
Cow Park	83	Grassland	0.02	0.21	(Jones et al., 2005)
Norfolk	55	No Grazing	1.27	13.46	Unpublished
Salisbury	61	No Grazing	0.08	0.85	Unpublished
Surrey	36	No Grazing	0.59	6.25	Unpublished
Easter Bush (a)	704	Sheep Grazing	0.08	0.85	(Skiba et al., 2013)
HouseO'Muir	80	Grassland	-0.33	-3.50	Unpublished
Easter Bush (b)	701	Grassland	2.84	30.09 ^(**)	(Drewer et al., 2017b)
East Grange	184	Grassland	-0.03	-0.32	(Drewer et al., 2017a)
Crichton	434	Silage Crop	0.39	4.13	Unpublished
Boghall Glen	560	No Grazing	1.58	16.74 ^(*)	Unpublished
Kirkton	580	No Grazing	3.01	31.89 ^(*)	Unpublished
Easter Bush (c)	650	Silage Crop	0.02	0.21	(Cowan et al., 2019)
Easter Bush (d)	811	Sheep Grazed	0.19	2.01	Unpublished
North Wyke (b)	704	Silage Crop	2.67	28.29 ^(*)	(Carswell et al., 2019)
Upper Joiner	544	Silage Crop	0.82	8.69	(Cowan et al., 2019)
Johnstone Castle (a)	1278	No Grazing	-0.15	-1.59 ^(*)	(Maire et al., 2018)
Johnstone Castle (b)	2479	Cattle Grazing	-0.11	-1.17	Unpublished
Johnstone Castle (c)	31892	Grassland	0.50	5.31 ^(*)	Unpublished

(*) Experiment included application of manure or organic fertiliser: (**) Experiment included tillage event

To put the impact of CH₄ emissions and uptake from agricultural soils in context with the other GHGs, we present the data from the Easter Bush grassland field (Midlothian, Scotland) where a variety of GHG measurements have occurred over several years (Jones et al., 2017; Cowan et al., 2018; Cowan et al., 2020a; Cowan et al., 2020b), (Table 5). Excluding the study in which the tillage event occurred, the mean CH₄ flux for the Easter Bush field site is 0.1 nmol m⁻² s⁻¹, which is close to the magnitude of CH₄ measured from sites when organic fertiliser application and tillage events are ignored. Using these estimates, the relative contribution of CH₄ to emissions from the soil are negligible when compared to uncertainties that would be expected in CO₂ and N₂O emission estimates. The contribution to the overall CO_{2eq} emissions from the site makes up only 0.4 % of the total budget, i.e. only fractionally more than those estimated for carbon monoxide (CO, secondary GHG) which contributes an estimated 0.3% of the total budget. At sites where uptake is expected, a similar comparison could be expected in which uptake of CH₄ would be statistically insignificant when compared to the uptake and emission of other GHGs. However, at sites where manure application and tillage occur on a regular basis (usually arable), emissions of CH₄ are likely to be of the same magnitude in terms of GWP as N₂O (approximately 30 g CO_{2eq} m⁻² yr⁻¹).

Table 5. The annual GHG budget for the soils and grass crop for the Easter Bush grassland field (Midlothian, Scotland) are presented, using GWP values provided from the IPCC 2014, Fifth Assessment Report (AR5).

GHG	Annual Emission (g C m ⁻²)/ (g N m ⁻²)	Typical Emission (g C m ⁻²)/ (g N m ⁻²)	AR5 GWP ratio	Annual Flux (g CO _{2eq} m ⁻²)	Annual GWP (%)
CO ₂	-605 to 72	-217.9	1	-217.9	88
N ₂ O	0.64 to 5.42	2.75	265	26.02	11
CH ₄	0.01 – 1.07	0.04	28	1.06	0.4
CO	0.35 to 0.38	0.37	~2	~0.74	0.3

From this meta-analysis it can be concluded that where known CH₄ emitting events such tillage or organic fertiliser application are not carried out, agricultural soils in the British Isles (UK & Ireland) or subject to similar N European climatic conditions are highly unlikely to exceed the threshold at which CH₄ fluxes are (a) robustly measurable by EC approaches and (b) make a significantly large contribution to the total site GHG budget. Thus, a blanket statement that for Class 1 compliance at N European stations on agricultural sites only N₂O but not CH₄ EC flux measurements are required seems appropriate.

10. CONCLUSIONS AND RECOMMENDATIONS

1. *Data synchronisation*: the LICOR Smartflux update now provides the capability to act as an NPT time server allowing gas analysers to synchronise themselves to its PTP timing. Guidance needs to be developed on how to set up the different analysers to make best use of this capability. Data acquisition systems not based on the Smartflux module need to reproduce this capability. PTP->NPT accuracy, if implemented well, is deemed adequate for closed path sensors. Separate data streams will need to be submitted to the Carbon Portal for the final data matching to be done by the ETC.
2. *Spike detection*: a new algorithm has been developed that performs better than previous algorithms. This is not just applicable to CH₄ and N₂O, but also to CO₂. It is therefore recommended to implement this algorithm into the EddyPro engine to be available to the ETC as well as end users. The algorithm now needs to be tested further on real flux data, especially on CH₄ flux data that contains real extreme events, e.g. from ebullition events and ruminants.
3. *Time-lag determination*: two approaches were investigated to improve the time-lag quantification when time-series show low SNR, an experimental and an improved numerical method:
 - a. The experimental method (reference gas overflow inlet) performed robustly and at the same time provides an estimate of (i) the inlet/analyser response time and (ii) instrument noise. However, implementation and data logging require some effort from site operators as no off-the-shelf implementations are currently available. The incorporation of the data streams into the ETC data system needs to be discussed.
 - b. First tests showed that the improved numerical method, based on prewhitening (PW), can detect a robust time-lag in more conditions than standard cross-covariance maximisation and it can also provide a measure of confidence for the time-lag. However, the method cannot capture the time-lag in all situations. The time lag detection strongly depends on the model adopted to filtering data. In some circumstances, in particular when the covariance is strongly unstable for instance, it would be appropriate to increase the model complexity to improve the effectiveness of the pre-whitening procedure.

Based on the information to date it is recommended:

- Where at all possible, use an analyser that also measures a compound showing a large flux for most of the time (e.g. CO₂, or CH₄ at wetland sites) so a proxy gas is available, especially for the processing of N₂O fluxes.
- To implement the prewhitening method into EddyPro and use it.

- To do further test to decide on a confidence threshold level above which the PW time-lag is deemed to be robust.
 - Use the PW time-lag if confidence is high, otherwise use the time-lag from the proxy gas if available, otherwise interpolate time-lags.
 - The ETC needs to flag stations at which time-lags vary with time and for which too little data from PW or proxy gas are available to robustly derive time-lags. In these situations, the options should be given to (i) either switch to a different analyser that measures a (different?) proxy gas or (ii) implement the reference gas overflow inlet.
4. *Quantification of time-response.* Deriving the system response time from the power spectral density of the gas mixing ratio is sensitive to how instrument noise is dealt with. The current approach of Ibrom et al. (2007) was found to remove real signal in certain condition and a new approach, based on combining the fit to the noise and the response function and constraining the noise to be white, eliminated this problem and improved the ease of the fit (Aslan et al, 2020). This approach should extend the applicability of the power-spectral approach (PSA) to spectra measured with lower signal-to-noise ratio, i.e. to smaller fluxes and noisier instruments. It is recommended to update the fitting procedure in EddyPro. In addition, the RINGO work has re-highlighted the interconnectivity between the co-spectral attenuation and the time-lag quantification. When using the co-spectral approach (CSA) to derive the response time and the gas mixing ratio and w time series are shifted according to CCM approach, it is recommended to use H^*H_p (or the approximation \sqrt{H}) for the correct estimates of response time and flux correction factor (Peltola et al., 2020).
- In general, the system time-response tends to be less variable than the time-lag and this makes it easier to extrapolate the time-response from high-flux periods to extended periods of low fluxes during which a response time cannot be derived. On the other hand, if periods with relatively high signal-to-noise ratio are not available, the recommendation is to estimate the response time from a proxy gas measured in the same analyser (e.g. CO₂) for which the spectral transfer function is well defined and easy to fit.
5. *Analyser noise quantification.* Gas analysers for N₂O and CH₄ need to work at their optimum to achieve the flux detection limits needed to monitor background fluxes and their performance therefore needs to be monitored continuously. The method to derive analyser noise from the autocorrelation function by Lenschow et al. (2000) should be implemented into EddyPro and routinely applied by the ETC. Automated feedback to the station managers would be useful if an increase in the noise level above a certain threshold is observed. This approach only quantifies random noise, but not structured noise

which can be generated, e.g., by temperature fluctuations. Thus, the environmental conditions of the analyser need to be temperature controlled as described by Nemitz et al. (2018).

6. *u_* filtering.* Under low turbulence conditions, typically characterized by low u_* values, the measured fluxes do not represent the surface exchange and other not quantifiable flux components (advection) may dominate. In case of CH₄ and N₂O flux measurements, current approach follows the recommendation given by Nemitz et al (2018) on performing u_* filtering on the basis of a u_* threshold from CO₂, for which it can be estimated, to characterize site-specific conditions of limiting turbulence. However, for those sites, like wetlands characterized by extent water surfaces/ponds within the flux tower footprint, it should be checked if turbulence is a driver of diffusive gas fluxes, and in this case u_* filtering should be avoided. In addition, based on the results presented for a grassland site CO₂ fluxes, it was proposed to assess whether current FLUXNET2015 approach (Pastorello et al, 2015) could be further developed by using seasonal u_* thresholds instead of one single threshold for the complete year. This would probably minimize data loss due to the u_* filtering. The robustness of seasonal u_* thresholds could be increased by pooling seasonal data from adjacent years before calculations.
7. *Gap-filling.* Gaps in CH₄ and N₂O flux measurements require reliable gap-filling algorithms to accurately estimate site annual budgets and seasonal dynamics. Such task is particularly challenging for CH₄ and N₂O fluxes given their high spatial and temporal variability and non-linear responses to multiple ecosystem drivers. Here we have tested several gap-filling approaches on N₂O and CH₄ fluxes measured in a grassland and wetland site, respectively. For N₂O flux dataset, the Random Forest and the Marginal Distribution Sampling methods had the better performance comparing to other simpler approaches. For the wetland site the ANN algorithm with water table depth, soil temperature at 20 cm depth, and friction velocity as drivers gave the best performance. Current efforts are focusing on systematically investigating all these approaches across a range of sites to provide best practices for gap-filling CH₄ and N₂O exchange (Irvin et al, in preparation).
8. *Update on site characteristics under which CH₄/N₂O flux measurements by EC are deemed appropriate and thus mandatory for Class-I sites.* The ICOS Protocol (Nemitz et al., 2018) sets out the first set of criteria to help identify at which sites eddy-covariance fluxes of N₂O and/or CH₄ are required to fulfil the requirements for Class 1 operation. Partly in support of ICOS and RINGO, a meta-analysis of CH₄ fluxes from agricultural soils was undertaken based on a large existing pool of measurements for soils of the British Isles. Accordingly, it was found that on

agricultural sites only N₂O but not CH₄ EC flux measurements are required, unless field practices promoting CH₄ emissions, like tillage or organic fertiliser application, are carried out.

11. List of symbols and acronyms

Symbols [Units]

$C_{o_{WT}}$	Normalized cospectrum of kinematic heat flux [-]
Δt_{LPF}	Time-lag due to low-pass filtering [s]
f	Natural frequency [Hz]
f_c	Cut-off frequency [Hz]
H_{LPF}	Spectral transfer function [-]
H	Lorentzian transfer function [-]
H_p	Part of transfer function describing phase shift [-]
N	Noise term
N	Acquisition frequency x time lag
σ_T^2	Variance of un-attenuated T [K ²]
S_T	Spectral density of sonic temperature
SW_IN	Incoming shortwave radiation [W m ⁻²]
SWC	Soil water content [%]
$tsSA$	Variable time since slurry application [day]
tsP	Time since precipitation
TS	Soil temperature [°C]
TA	Air temperature [°C]
T_{exp}	Experimental transport time lag [s]
T_{line}	Gas travel time through the sampling tube [s]
$T_{analyser}$	Analyser-specific response time [s]
T_{VT}	Gas travel time from valve to T-piece [s]
$T_{actuation}$	Actuation time of the solenoid valve [s]
T_{theo}^{EC}	Theoretical “eddy-covariance” time lag [s]
T_{SI}	Horizontal travel time from inlet to sonic [s]
T_{IT}	Travel time from inlet to T-piece [s]

T_s	Un-attenuated turbulent sonic temperature [K]
τ	Response time [s]
τ_{LPF}	Response time of first order filter [s]
u_*	Friction Velocity [m s^{-1}]
VPD	Water pressure deficit [Pa]
$\overline{w'T'}$	Kinematic heat flux [K m s^{-1}]

Acronyms

AR	Autoregressive
a.s.l.	Above sea level
ANN	Artificial Neural Network
C	Carbon
CH_4	Methane
CO	Carbon monoxide
COS	Carbonyl sulphide
CO_2	Carbon dioxide
$\text{CO}_{2\text{eq}}$	Carbon dioxide equivalent
CP	ICOS Carbon Portal
CCM	Cross-covariance maximisation
CCF	Cross-correlation function
CF	Flux correction factor
CSA	Co-spectral approach
CSA_H	Co-spectral approach implementing H
$\text{CSA}_{\sqrt{H}}$	Co-spectral approach implementing \sqrt{H}
$\text{CSA}_{\sqrt{H}, \text{sync}}$	Co-spectral approach implementing \sqrt{H} with shifted w time series
CH-Fru	Grassland site Früebüel in Switzerland
CH-Cha	Grassland site Chamau
CH-Dav	ICOS Class 1 site Davos
CPU	Central processing unit
Cross-Cor	cross-correlation
Cross-Cov	cross-covariance
EC	Eddy Covariance
EddyPro	Eddy covariance post-processing software

ETC	Ecosystem Thematic Centre
GARCH	Generalized AutoRegressive Conditional Heteroskedastic
GHG	Greenhouse Gases
GPS	Global Positioning System
GWP	Global Warming Potential
IC	Inorganic Carbon
ICOS	Integrated Carbon Observatory System
IIR	Infinite impulse response filter
LoD	Limit of detection
LUT	Look-up tables
MDA	Mean diurnal averaging after Reichstein et al. (2005)
MDC	Mean diurnal course after Falge et al. (2001)
MDS	Marginal distribution sampling
MPT	Moving Point Test
N	Nitrogen
NTP	Network Time Protocol
N ₂ O	Nitrous Oxide
OLS	Ordinary least squares
PAR	Photosynthetically active radiation
PSA	Power spectral approach
PSA _{I07}	Power spectral approach developed by Ibrom et al. (2007b)
PSA _{A20}	Power spectral approach developed by Aslan et al. (2020)
PTP	Precision Time Protocol
PW	Prewhitening
RF	Random Forest
RM	Repeated median
SLI	Simple linear interpolation
SNR	Signal-to-noise ratio
SRAM	Simple running arithmetic mean
SRMED	Simple running median
SA	Slurry application
SDev	Standard deviation
TL	Time lag

QCL	Quantum Cascade Laser
QA/QC	Quality assurance/ quality control
VOCs	Volatile Organic Compounds
3D	Three-dimensional
QCL	Quantum Cascade Laser
VUT	Variable u_* threshold

12. References

Aslan, T., Peltola, O., Ibrom, A., Nemitz, E., Rannik, Ü and Mammarella, I., 2020. On the high frequency response correction of eddy covariance fluxes. Part 1: an experimental approach for analyzing noisy measurements of small fluxes. To be submitted to *Atmos. Meas. Tech Discuss*

Aubinet, M., Grelle, A., Ibrom, A., Rannik, Ü., Moncrieff, J., Foken, T., ... & Clement, R. (1999). Estimates of the annual net carbon and water exchange of forests: the EUROFLUX methodology. In *Advances in ecological research* (Vol. 30, pp. 113-175). Academic Press.

Aubinet, M., Vesala, T., & Papale, D. (Eds.). (2012). *Eddy covariance: a practical guide to measurement and data analysis*. Springer Science & Business Media.

Bell, M. J., Hinton, N. J., Cloy, J. M., Topp, C. F. E., Rees, R. M., Williams, J. R., ... & Chadwick, D. R. (2016). How do emission rates and emission factors for nitrous oxide and ammonia vary with manure type and time of application in a Scottish farmland?. *Geoderma*, 264, 81-93.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3), 307-327.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32.

Butterbach-Bahl, K., Baggs, E. M., Dannenmann, M., Kiese, R., & Zechmeister-Boltenstern, S. (2013). Nitrous oxide emissions from soils: how well do we understand the processes and their controls? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1621), 20130122. <https://doi.org/10.1098/rstb.2013.0122>

Carswell, A., Shaw, R., Hunt, J., Sánchez-Rodríguez, A.R., Saunders, K., Cotton, J., Hill, P.W., Chadwick, D.R., Jones, D.L., Misselbrook, T.H. (2019). Assessing the benefits and wider costs of different N fertilisers for grassland agriculture. *Archives of Agronomy and Soil Science* 65, 625–639. <https://doi.org/10.1080/03650340.2018.1519251>

Cowan, N., Levy, P., Maire, J., Coyle, M., Leeson, S.R., Famulari, D., Carozzi, M., Nemitz, E., Skiba, U. (2020a). An evaluation of four years of nitrous oxide fluxes after application of ammonium nitrate and urea fertilisers measured using the eddy covariance method. *Agricultural and Forest Meteorology* 280, 107812. <https://doi.org/10.1016/j.agrformet.2019.107812>

Cowan, N., Carnell, E., Skiba, U., Dragosits, U., Drewer, J., Levy, P. (2020b). Nitrous oxide emission factors of mineral fertilisers in the UK and Ireland: A Bayesian analysis of 20 years of experimental data. *Environment International* 135, 105366. <https://doi.org/10.1016/j.envint.2019.105366>

Cowan, N., Levy, P., Moring, A., Simmons, I., Bache, C., Stephens, A., Marinheiro, J., Brichet, J., Song, L., Pickard, A., McNeill, C., McDonald, R., Maire, J., Loubet, B., Voylokov, P., Sutton, M., Skiba, U. (2019). Nitrogen use efficiency and N₂O and NH₃; losses attributed to three fertiliser types applied to an intensively managed silage crop. *Biogeosciences* 16, 4731–4745. <https://doi.org/10.5194/bg-16-4731-2019>

Cowan, N., Helfter, C., Langford, B., Coyle, M., Levy, P., Moxley, J., Simmons, I., Leeson, S., Nemitz, E., Skiba, U. (2018). Seasonal fluxes of carbon monoxide from an intensively grazed grassland in Scotland. *Atmospheric Environment* 194, 170–178. <https://doi.org/10.1016/j.atmosenv.2018.09.039>

Cowan, N. J., Norman, P., Famulari, D., Levy, P. E., Reay, D. S., & Skiba, U. M. (2015). Spatial variability and hotspots of soil N₂O fluxes from intensively grazed grassland. *Biogeosciences*, 12(5), 1585–1596. <https://doi.org/10.5194/bg-12-1585-2015>

Cryer, J. D., & Chan, K. S. (2008). *Time series analysis: with applications in R*. Springer Science & Business Media.

Davies, P.L., Fried, R., Gather, U. (2004). Robust signal extraction for online monitoring data. *Journal of Statistical Planning and Inference*, 122(1-2): 65–78, DOI 10.1016/j.jspi.2003.06.012.

Drewer, J., Finch, J. W., Lloyd, C. R., Baggs, E. M., & Skiba, U. (2012). How do soil emissions of N₂O, CH₄ and CO₂ from perennial bioenergy crops differ from arable annual crops? *Global Change Biol Bioenergy*, 4(4), 408-419.

Drewer, J., Yamulki, S., Leeson, S. R., Anderson, M., Perks, M. P., Skiba, U. M., & McNamara, N. P. (2017a). Difference in soil methane (CH₄) and nitrous oxide (N₂O) fluxes from bioenergy crops SRC willow and SRF Scots pine compared with adjacent arable and fallow in a temperate climate. *BioEnergy Research*, 10(2), 575-582.

Drewer, J., Leduning, M., Kerdraon-Byrne, D., Sayer, E., Sentien, J., & Skiba, U. (2017b). The impact of land-use change from forest to oil palm on soil greenhouse gas and volatile organic compound fluxes in Malaysian Borneo. *EGU GA*, 6495.

Eidson, J. and Lee, K. (2002, November). IEEE 1588 standard for a precision clock synchronization protocol for networked measurement and control systems. In Sensors for Industry Conference, 2002. 2nd ISA/IEEE (pp. 98-105). Ieee.

Eugster, W., & Senn, W. (1995). A cospectral correction model for measurement of turbulent NO₂ flux. *Boundary-Layer Meteorology*, 74(4), 321-340.

Eugster, W., & Merbold, L. (2015). Eddy covariance for quantifying trace gas fluxes from soils. *Soil*, 1(1), 187.

Falge, E., Baldocchi, D., Olson, R. J., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Ta Lai, C., Law, B. E., Meyers, T., Moncrieff, J., Moors, E., Munger, J. W., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K. and Wofsy, S. (2001). "Gap filling strategies for defensible annual sums of net ecosystem exchange." *Agricultural and Forest Meteorology* 107: 43-69.

Foken, T., and Wichura, B. (1996). Tools for quality assessment of surface-based flux measurements. *Agricultural and Forest Meteorology*, 78(1–2), 83–105. [https://doi.org/10.1016/0168-1923\(95\)02248-1](https://doi.org/10.1016/0168-1923(95)02248-1)

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

Fratini, G., Sabbatini, S., Ediger, K., Riensche, B., Burba, G., Nicolini, G., Vitale, D. and Papale, D. (2018). Eddy covariance flux errors due to random and systematic timing errors during data acquisition. *Biogeosciences*, 15(17), pp.5473-5487.

Fuchs, K., Hörtnagl, L., Buchmann, N., Eugster, W., Snow, V., & Merbold, L. (2018). Management matters: testing a mitigation strategy for nitrous oxide emissions using legumes on intensively managed grassland. *Biogeosciences*, 15(18), 5519–5543. <https://doi.org/10.5194/bg-15-5519-2018>

Fried, R., Schettlinger, K., Borowsky, M. (2019). Robust Time Series Filters. URL <https://CRAN.R-project.org/package=robfilter>, R package version 4.1.1.

Herbst, M., Friberg, T., Ringgaard, R. and Soegaard, H. (2011). "Interpreting the variations in atmospheric methane fluxes observed above a restored wetland." *Agricultural and Forest Meteorology*, 151(7): 841-853.

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.

Hörtnagl, L., Barthel, M., Buchmann, N., Eugster, W., Butterbach-Bahl, K., Díaz-Pinés, E., Zeeman, M., Klumpp, K., Kiese, R., Bahn, M., Hammerle, A., Lu, H., Ladreiter-Knauss, T., Burri, S., & Merbold, L. (2018).

Greenhouse gas fluxes over managed grasslands in Central Europe. *Global Change Biology*, 24(5), 1843–1872. <https://doi.org/10.1111/gcb.14079>

Hunt, J. E., Laubach, J., Barthel, M., Fraser, A., & Phillips, R. L. (2016). Carbon budgets for an irrigated intensively grazed dairy pasture and an unirrigated winter-grazed pasture. *Biogeosciences*, 13(10), 2927-2944.

Jones, S.K., Rees, R.M., Skiba, U.M., Ball, B.C. (2005). Greenhouse gas emissions from a managed grassland. *Global and Planetary Change*, 47, 201–211. <https://doi.org/10.1016/j.gloplacha.2004.10.011>

Jones, S. K., Famulari, D., Di Marco, C. F., Nemitz, E., Skiba, U. M., Rees, R. M., & Sutton, M. A. (2011). Nitrous oxide emissions from managed grassland: a comparison of eddy covariance and static chamber measurements. *Atmospheric Measurement Techniques*, 4(10), 2179-2194.

Jones, S. K., Helfter, C., Anderson, M., Coyle, M., Campbell, C., Famulari, D., ... & Kiese, R. (2017). The nitrogen, carbon and greenhouse gas budget of a grazed, cut and fertilised temperate grassland. *Biogeosciences*, 14(8), 2060-2088.

Ibrom, A., Dellwik, E., Flyvbjerg, H., Jensen, N. O., & Pilegaard, K. (2007a). Strong low-pass filtering effects on water vapour flux measurements with closed-path eddy correlation systems. *Agricultural and Forest Meteorology*, 147(3-4), 140-156.

Ibrom, A., Dellwik, E., Larsen, S. E., & Pilegaard, K. I. M. (2007b). On the use of the Webb–Pearman–Leuning theory for closed-path eddy correlation measurements. *Tellus B: Chemical and Physical Meteorology*, 59(5), 937-946.

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Kim, Y., Johnson, M. S., Knox, S. H., Black, T. A., Dalmagro, H. J., Kang, M., Kim, J., & Baldocchi, D. (2019). Gap-filling approaches for eddy covariance methane fluxes: A comparison of three machine learning algorithms and a traditional method with principal component analysis. *Global Change Biology*, gcb.14845. <https://doi.org/10.1111/gcb.14845>

Knox, S. H., Matthes, J. H., Sturtevant, C., Oikawa, P. Y., Verfaillie, J. and Baldocchi, D. (2016). "Biophysical controls on interannual variability in ecosystem-scale CO₂ and CH₄ exchange in a California rice paddy." *J Geophys Res G: Biogeosci* 121(3): 978-1001.

Kohonen, K.-M., Kolari, P., Kooijmans, L. M. J., Chen, H., Seibt, U., Sun, W., and Mammarella, I. (2020). Towards standardized processing of eddy covariance flux measurements of carbonyl sulfide, *Atmos. Meas. Tech.*, 13, 3957–3975, <https://doi.org/10.5194/amt-13-3957-2020>.

Kursa, M. B., & Rudnicki, W. R. (2010). Feature Selection with the Boruta Package. *Journal of Statistical Software*, 36(11). <https://doi.org/10.18637/jss.v036.i11>

Langford, B., Acton, J., Ammann, C., Valach, A., & Nemitz, E. (2015). Eddy-covariance data with low signal-to-noise ratio: time-lag determination, uncertainties and limit of detection. *Atmospheric Measurement Techniques*, 8, 4197-4213.

Lenschow, D. H., Wulfmeyer, V., & Senff, C. (2000). Measuring second-through fourth-order moments in noisy data. *Journal of Atmospheric and Oceanic technology*, 17(10), 1330-1347.

Maechler, M., Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., Verbeke, T., Koller, M., Conceicao, E.L.T., Anna di Palma, M. (2019). robustbase: Basic Robust Statistics. URL <http://robustbase.r-forge.r-project.org/>, R package version 0.93-5.

Maire, J., Gibson-Poole, S., Cowan, N., Reay, D.S., Richards, K.G., Skiba, U., Rees, R.M., Lanigan, G.J. (2018). Identifying Urine Patches on Intensively Managed Grassland Using Aerial Imagery Captured From Remotely Piloted Aircraft Systems. *Frontiers in Sustainable Food Systems* 2. <https://doi.org/10.3389/fsufs.2018.00010>

Mammarella, I., Launiainen, S., Gronholm, T., Keronen, P., Pumpanen, J., Rannik, Ü., & Vesala, T. (2009). Relative humidity effect on the high-frequency attenuation of water vapor flux measured by a closed-path eddy covariance system. *Journal of Atmospheric and Oceanic Technology*, 26(9), 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, doi:10.5194/amt-9-4915-2016

Mandelbrot BB. (1997). The variation of certain speculative prices. In: Fractals and scaling in finance, Springer, New York, NY, pp 371–418, DOI 10.1007/978-1-4757-2763-0\ 14

Massman, W. J. (2000). A simple method for estimating frequency response corrections for eddy covariance systems. *Agricultural and Forest Meteorology*, 104(3), 185-198.

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

Matson, J. (2013). Choosing the correct Time Synchronization Protocol and incorporating the 1756-TIME module into your Application. Publication ENET-WP030A-EN-E, Rockwell Automation. https://literature.rockwellautomation.com/idc/groups/literature/documents/wp/enet-wp030_en-e.pdf

Merbold, L., Eugster, W., Stieger, J., Zahniser, M., Nelson, D., & Buchmann, N. (2014). Greenhouse gas budget (CO_2 , CH_4 and N_2O) of intensively managed grassland following restoration. *Global Change Biology*, 20(6), 1913–1928. <https://doi.org/10.1111/gcb.12518>

Metzger, S., Junkermann, W., Mauder, M., Beyrich, F., Butterbach-Bahl, K., Schmid, H., Foken, T. (2012). Eddy-covariance flux measurements with a weight-shift microlight aircraft. *Atmospheric Measurement Techniques*, 5(7): 1699–1717, DOI 10.5194/amt-5-1699-2012.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G., Beckstein, C., Braswell, B. H., Churkina, G., R., D. A., Falge, E., Gove, J. H., Heimann, M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A. and Stauch, V. J. (2007). "Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes." *Agricultural and Forest Meteorology* 147: 209-232.

Moffat, A. M., Beckstein, C., Churkina, G., Mund, M. and Heimann, M. (2010). "Characterization of ecosystem responses to climatic controls using artificial neural networks." *Global Change Biology*, 16: 2737–2749.

Moncrieff, J. B., Massheder, J. M., De Bruin, H., Elbers, J., Friberg, T., Heusinkveld, B., ... & Verhoef, A. (1997). A system to measure surface fluxes of momentum, sensible heat, water vapour and carbon dioxide. *Journal of Hydrology*, 188, 589-611.

Moore, C. J. (1986). Frequency response corrections for eddy correlation systems. *Boundary-Layer Meteorology*, 37(1-2), 17-35.

Nemitz, E., Mammarella, I., Ibrom, A., Aurela, M., Burba, G. G., Dengel, S., ... & Hortnagl, L. (2018). Standardisation of eddy-covariance flux measurements of methane and nitrous oxide. *International Agrophysics*, 32(4).

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T. and Yakir, D. (2006). "Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: algorithms and uncertainty estimation." *Biogeosciences*, 3(4): 571-583.

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y. W., ... & Isaac, P. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific data*, 7(1), 1-27.

Peltola, O., Mammarella, I., Haapanala, S., Burba, G., & Vesala, T. (2013). Field intercomparison of four methane gas analysers suitable for eddy covariance flux measurements. *Biogeosciences*, 10(6), 3749-3765

Peltola, O., Hensen, A., Helfter, C., Belelli Marchesini, L., Bosveld, F. C., van den Bulk, W. C. M., Elbers, J. A., Haapanala, S., Holst, J., Laurila, T., Lindroth, A., Nemitz, E., Röckmann, T., Vermeulen, A. T., and Mammarella, I. (2014). Evaluating the performance of commonly used gas analysers for methane eddy covariance flux measurements: the InGOS inter-comparison field experiment, *Biogeosciences*, 11, 3163-3186, doi:10.5194/bg-11-3163-2014

Peltola, O., Aslan, T., Ibrom, A., Nemitz, E., Rannik, Ü and Mammarella, I. (2020). On the high frequency response correction of eddy covariance fluxes. Part 2: the experimental approach and the interdependence with the time lag estimation. To be submitted to *Atmos. Meas. Tech Discuss*

Rannik, Ü., Haapanala, S., Shurpali, N. J., Mammarella, I., Lind, S., Hyvönen, N., Peltola, O., Zahniser, M., Martikainen, P. J., and Vesala, T. (2015). Intercomparison of fast response commercial gas analysers for nitrous oxide flux measurements under field conditions, *Biogeosciences*, 12, 415-432, doi:10.5194/bg-12-415-2015

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránkova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D. and Valentini, R. (2005). On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Global Change Biology* 11(9): 1424-1439.

Rousseeuw, P.J., Croux, C. (1993). Alternatives to the median absolute deviation. *J Am Stat Assoc*, 88(424): 1273–1283, DOI 10.1080/01621459.1993.10476408.

Sánchez-Rodríguez, A.R., Carswell, A.M., Shaw, R., Hunt, J., Saunders, K., Cotton, J., Chadwick, D.R., Jones, D.L., Misselbrook, T.H. (2018). Advanced Processing of Food Waste Based Digestate for Mitigating Nitrogen Losses in a Winter Wheat Crop. *Front. Sustain. Food Syst.*, 2, 35. <https://doi.org/10.3389/fsufs.2018.00035>

Shurpali, N. J., Verma, S. B., Clement, R. J., and Billesbach, D. P.: Seasonal distribution of methane flux in a Minnesota peatland measured by eddy correlation, *Journal of Geophysical Research: Atmospheres*, 98, 20649-20655, 10.1029/93jd02181, 1993.

Siegel, A.F. (1982). Robust regression using repeated medians. *Biometrika*, 69(1): 242–244, DOI 10.1093/biomet/69.1.242.

Skiba, U., Jones, S. K., Drever, J., Helfter, C., Anderson, M., Dinsmore, K., ... & Sutton, M. A. (2013). Comparison of soil greenhouse gas fluxes from extensive and intensive grazing in a temperate maritime climate. *Biogeosciences*, 10(2), 1231-1241.

Starkenburg, D., Metzger, S., Fochesatto, G.J., Alfieri, J.G., Gens, R., Prakash, A., Cristobal, J. (2016). Assessment of despiking methods for turbulence data in micrometeorology. *Journal of Atmospheric and Oceanic Technology*, 33(9): 2001–2013, DOI 10.1175/JTECH-D-15-0154.1.

Tsay, R. S. (2005). *Analysis of financial time series* (Vol. 543). John Wiley & Sons.

Vickers, D., Mahrt, L. (1997). Quality control and flux sampling problems for tower and aircraft data. *J Atmos Ocean Technol*, 14(3): 512–526, DOI 10.1175/1520-0426(1997)014<0512:QCAFSP>2.0.CO;2.

Vitale, D., Fratini, G., Bilancia, M., Nicolini, G., Sabbatini, S., Papale, D. (2020). A robust data cleaning for eddy covariance flux measurements, *Biogeosciences*, 17(6): 1367-1391 DOI: 10.5194/bg-17-1367-2020.

Vitale, D., Papale, D., and ICOS-ETC Team, (2019). RFlux: An R Package for Processing and Cleaning Eddy Covariance Flux Measurements. URL: <https://github.com/icos-etc/RFlux>.

Vitale et al (in prep). A performance evaluation of despiking algorithms for eddy covariance data.

Vitale et al (in prep). A pre-whitening procedure for time lag detection between eddy covariance measurements.

Wuertz, D., Setz, T., Chalabi, Y., Boudt, C., Chausse, P., Miklovac, M. (2019). fGarch: Rmetrics – Autoregressive Conditional Heteroskedastic Modelling. R package version 3042.83.1, <https://CRAN.R-project.org/package=fGarch>.

Wutzler, T., Lucas-Moffat, A., Migliavacca, M., Knauer, J., Sickel, K., Šigut, L., Menzer, O., & Reichstein, M. (2018). Basic and extensible post-processing of eddy covariance flux data with REddyProc. *Biogeosciences*, 15(16), 5015–5030. <https://doi.org/10.5194/bg-15-5015-2018>