Review response to discussion article “Towards standardized processing of eddy covariance flux measurements of carbonyl sulfide” by Kohonen et al.

Reviewer comments in black
Author response in purple
Altered text in the manuscript in italic
Changes made to the revised manuscript since the submission of the author response
Altered text in the revised manuscript (after author response to review comments)

Reviewer #1: Georg Wohlfahrt

General assessment: Kohonen et al. report on the effects of varying various post-processing steps required for eddy covariance COS flux measurements with the aim, as stated in the title, to standardize these. COS EC flux measurements are increasingly making their way into the literature as COS offers a novel means of constraining GPP and stomatal conductance. Yet, the necessary processing steps are way not as harmonized as is the case for CO2, potentially causing systematic bias between studies using different processing schemes, thus impeding synthesis activities. Overall I think this is a timely and relevant addition to the literature, which fits with the scope of the journal. I though also believe that the manuscript suffers from several issues, which will require significant changes, as detailed below.

Major comments:
(1) First, I have several formal issues with the manuscript: English style is often poor, which creates situations in which the intended meaning is not entirely clear (e.g. l. 32 the explanation of footprint limitations during stable stratification). Some of the formulations are too sloppy and thus misleading (e.g. l. 33 where “operation at high frequency” is mixed with “fast time response”). Some text is trivial or circular (e.g. l. 434-435), some of the concepts are wrong (e.g. l. 49) and some information is missing (e.g. legend of Fig. 11). Often some later, in-house studies are cited instead of the original papers. Finally, a mix of tenses is used when typically the past tense should be used to describe own results.
We will revise the manuscript and try to improve the english language, avoid circular text and repetition and only use the past tense. We cite the original papers as suggested. The caption of Fig. 10 (former Fig. 11) will be improved.
Manuscript has been revised and english language improved. Circular text and repetition has been reduced and only past tense is used in the text. All figure captions have been carefully checked to include all information.

(2) Novelty and justification of the study: In 2017 a methodological paper on COS EC flux measurement post-processing was published in the same journal (Gerdel et al.). The authors justify their paper mainly by stating that their analysis goes beyond this previous paper. While this is partially true (in particular the analyses on lag times is novel), I think the authors should follow what the title of the paper suggests and rather sell their work as contributing towards a standardization of COS EC flux post-processing routines. To this end, I suggest to synthesize, e.g. in a table, the various processing steps that were used by previously published studies as a starting point and use this as a backbone for their analysis and the resulting recommendations. This table would then summarize whether and if so how previous studies detrended their time series, how the lag time was found, how low/high-frequency response corrections were applied,

whether data were filtered for low u* (how were thresholds found) and which QC/QA was used. Following this suggestion requires at least the introductory section to be more or less completely re-written and would allow the paper to live up to what its title suggests and eventually become a reference for COS EC flux measurements.
Thank you for this suggestion. The Introduction section will be reorganized and partly rewritten. The study's objectives will be more clear. A table summarizing previous studies is a very good idea and will be implemented as Table 1 in the new version of the manuscript. We will revise the whole manuscript and extend the discussion to processing routines used in earlier studies.
The introduction section has been reorganized and partly rewritten, ending with a statement of study objectives. Table 1 summarizing processing steps used in previous studies has been added to the introduction.

(3) Vertical advection: This section is somewhat odd – the authors acknowledge that knowing the magnitude of vertical advection is meaningless unless the magnitude of horizontal advection is known as well, yet vertical advection is reported even though horizontal advection has not been quantified. Unless the authors can come up with a discussion of what their results on the sign and magnitude of vertical advection actually mean in the context of their study, I thus suggest removing the results on vertical advection and all text/material that pertains to it.
We have considered this point carefully and came to the conclusion to leave out the results regarding vertical advection. The reviewer has a good point and as we are aiming for harmonization of processing protocols - where vertical advection is not used - we decided to leave this section out of the revised manuscript.
Vertical advection was left out of the manuscript.

(4) Corrections for high-frequency flux loss: Comparing two different approaches is novel for COS, yet surprisingly none of the underlying results are shown – I suggest to expand this section.
We discuss and show the difference to the reference processing scheme and show in Table 2 the effect of spectral corrections to final fluxes. Histograms and PDFs of different spectral correction schemes will be added to the Supplement (Fig. S2) and daytime and night-time median fluxes added to Table 2 and discussed in the text. We will also add a figure on flux attenuation to Supplementary material Fig. S3 and scatter plot comparing the final fluxes to Fig. S7.
Table 2 summarizes the differences in the final fluxes with the different spectral correction methods (total median flux, daytime and night-time median fluxes). We added the two cospectral models to Fig. 4 of the revised version of the manuscript, and histograms (and PDF) of the final fluxes to Supplementary material Fig. S6, as well as a scatter plot comparing the final fluxes to Fig. S11.

(5) Changes of co-spectral peak frequency with stability: Among the results of this study is a figure comparing the changes in the co-spectral peak frequency with stability for the Horst model and this study. While interesting, this analysis and the results are not motivated in the introduction and are barely discussed. Again, unless the authors are able to come up with a discussion of what the observed differences mean for their study, I suggest removing this material (or possibly moving it into a supplement).
We will move the figure to supplementary material (Fig. S4), as suggested. Equations 15 and 16 of the former version will be moved to Methods-section and presented as equations 6 and 7 in the revised version of the manuscript.
The figure was moved to supplementary material Fig. S8 and all equations realted to high-frequency spectral corrections are reported in Sect. 2.4.3 of the revised manuscript.

(6) Gap-filling: This is an indeed novel aspect, however way underexploited by the authors. Only a single arbitrarily chosen gap-filling algorithm is tested, the authors and miss to put it to a true test and results of gap-filling (e.g. time course of estimated parameters and selected results illustrating gap-filling behaviour) are lacking.

A time series of gap-filling parameters will be shown in the supplementary material Fig. S5. We will add a diurnal plot of the measured flux compared to different gap-filling functions in Fig. 11 in the revised manuscript as well as residuals of different methods in Fig. S6 in the supplementary material. We are not going into further detail in comparing different gap-filling methods, as that would result in a whole new paper. Long-term budgets are not usually the interest of the COS community, as GPP calculations from COS often aim at understanding the CO_2 exchange dynamics rather than calculating long-term budgets. However, it is important to fill short gaps (individual 30 mins or a bit longer) to e.g. get diurnal variations. The presented gap-filling method is just one example that can be used for gap-filling COS data. We will add discussion in Section 3.6 (former Section 3.7): "*Three combinations of environmental variables (PAR, PAR and relative humidity, PAR and VPD) were tested using the gap-filling function Eq. 16. These environmental parameters were chosen because COS exchange has been found to depend on stomatal conductance, which in turn depends especially on radiation and humidity (Kooijmans et al., 2019). Development of the gap-filling parameters a, b, c and d over the measurement period is presented in the Supplementary material Fig. S5. While saturating function of PAR only captured the diurnal variation already relatively well, adding a linear dependency on VPD or RH made the diurnal pattern even closer to the measured one (Fig. 11). Therefore, the combination of saturating light response curve and linear VPD dependency was chosen. Furthermore, we chose a linear VPD dependency instead of a linear RH dependency due to smaller residuals in the former (Fig. S6).*"

Time series of gap-filling parameters and residuals of different gap-filling methods are shown in Figs. S10 and S9 in the supplementary material, respectively. Diurnal variation of the different gap-filling functions in July-August are shown in Fig. 7 of the revised manuscript. The revised text in Sect. 3.6 reads "*Three combinations of environmental variables (PAR, PAR and relative humidity, PAR and VPD) were tested using the gap-filling function Eq. 17. These environmental parameters were chosen because COS exchange has been found to depend on stomatal conductance, which in turn depends especially on radiation and humidity (Kooijmans et al., 2019). Development of the gap-filling parameters a, b, c and d over the measurement period is presented in the Supplementary material Fig. S10. While saturating function of PAR only captured the diurnal variation already relatively well, adding a linear dependency on VPD or RH made the diurnal pattern even closer to the measured one (Fig. 7). Therefore, the combination of saturating light response curve and linear VPD dependency was chosen. Furthermore, we chose a linear VPD dependency instead of a linear RH dependency due to smaller residuals in the former (Fig. S9).*"

(7) Li-6262 and QCL CO2 cross-covariance maximisation: An important detail of this study is the setup, which includes a closed-path IRGA that is used to measure CO2 and H2O concentrations (not clear whether from the same tube as the QCL). These data are used to account for the drift in the computer clocks acquiring the sonic (& IRGA) and QCL data, respectively. While this nicely shows the benefit of having a complementary suite of measurements at "super-sites" such as Hyytiälä, in my view the reliance on an additional instrument is a drawback of this study as it limits the applicability of the proposed approach at other sites where no additional IRGA or an open-path model or closed-path model with a short tube is deployed. Even more so, this approach is unnecessary, as there are simpler, software-based, solutions available to keep computer clocks in

a network synchronized. In addition, by aligning data this way, the authors ruin a truly independent means of cross-comparing the QCL CO2 and H2O fluxes. I thus think it would be useful to explore the possibility of aligning the data sets in time without the help of the IRGA data. This is possible by expanding the time window in which the lag determination algorithm searches for, as with computer clocks being reset only once a day, time shifts of several seconds may result (in both directions). The reliability of this approach may then be checked by comparing against lag times and fluxes calculated with the IRGA.

We thank the Referee for this point. However, we recognize that standard CO2/H2O flux towers measuring other gases than CO2 typically use QCLs, which are nowadays providing also H2O and CO2, beside the target gas (CH4, N2O, COS, etc). In our study we have taken advantage of this setup. We acknowledge that better synchronization approach should be used already in the data logging system and a short text will be added to Section 2.2. Moreover, we have considered the Referee's suggestion of using a larger time window for COS, but it would be problematic because the covariance peak is not always clear and with a large time window there might be multiple peaks (for example related to low frequency variations). Instead, we have tested the file combination maximizing the covariance of CO2 (QCL) and w, and this resulted in a very similar outcome as the previous method. The results are shown as histograms in the supplement (Fig. S1). We rewrote this part of the manuscript as: "*The following procedure was done to combine two data files of 30 min length (of which one includes sonic anemometer and LI-6262 data and the other includes Aerodyne QCLS data): 1) the cross-covariance of the two CO_2 signals (QCLS and LI-6262) was calculated 2) the QCLS data were shifted so that the cross-covariance of the CO2 signals was maximized. Note that this will result in having the same lag time for QCL and LI-6262. The time shift was a maximum of 10 seconds, with most varying between 0 s to 2 s during one day. It is also possible to shift the time series by maximizing the covariance of CO_2 and w, which will then already account for the lag time (Fig. S1) or combine files according to their time stamps and allow a longer window in which the lag time is searched. However, in this case it is important that the lag time (and time shift) is determined from CO_2 measurements only, as using COS data might result in several covariance peaks in longer time frames due to low signal-to-noise ratios and small fluxes.*"

The reported changes listed above were implemented in the revised manuscript.

(8) Conclusions: The authors should conclude with referencing against what processing steps have been used by previous studies and put their results into perspective with these, by highlighting critical steps and the need for further harmonization.

As suggested by the Referee, we will add more discussion related to previous studies in the "Results and Discussions" chapter. Moreover, in the "Conclusions" we will highlight more clearly the critical steps and needs for further harmonization.

More discussion in relation to previous studies was added to Sect. 3 Results and Discussion. We finish the Conclusion section with fnal remarks "*Our recommendation for time lag determination (CO_2 cross-covariance) differs from the most commonly used method so far (COS cross-covariance), while experimental high frequency spectral correction has been widely applied already before. Many earlier studies have neglected the storage change flux, but we emphasize its importance in the diurnal variation of COS exchange. In addition, we encourage implementing gap-filling to future COS flux calculations for eliminating short-term gaps in data.*"

Detailed comments:
l. 1-3: reformulate to better convey intended meaning "... growing in popularity with the aim of estimating gross primary productivity at ecosystem scale, however lack standardized protocols ...“
Clarified. The sentences now read “*Carbonyl sulfide (COS) flux measurements with the eddy covariance (EC) technique are becoming popular for estimating gross primary productivity. To compare COS flux measurements across sites, we need standardized protocols for data processing.*”

l. 17: replace "due to the use“ with something more suitable, e.g. "motivated by ...“
Corrected as suggested.

l. 20: ", but in contrast to CO2, COS is destroyed...“
Corrected as suggested.

l. 22: for readers not familiar with the LRU, talking about the radiation-dependency of the LRU without introducing the concept might be highly confusing
Text about LRU removed.

l. 29: "... the assumptions underlying the EC method ...“
Corrected as suggested.
This text has been removed in the revised manuscript.

l. 29-34 and l. 35-41: these two paragraphs are in my view too general to meaningfully add to the introduction
We reduced the paragraphs into one sentence: “*To meet the assumptions underlying the EC method, the site, setup design, and instrumentation need to be considered (Aubinet et al., 2000, 2012; Nemitz et al., 2018; Sabbatini et al., 2018).*“
This sentence has been removed from the revised manuscript.

l. 32: reformulate – what you likely mean is that the measurement height should be such that the footprint remains within the ecosystem of interest even during stable stratification
This sentence will be removed as we will harmonize the Introduction section and make it more compact.
The sentence has been removed.

l. 33: EC instruments need to have a fast time response, which is different from “operation at high frequency” (a slow-response sensor does not become suitable for EC only because its data are logged at 20 Hz), as it can be shown that fast-response measurements made every few seconds do not cause a systematic bias in the EC flux
Corrected.

l. 38: not sure this sentence applies universally to all closed-path analyzers and anyway I would think this is too much detail for the introduction – suggest to remove
Sentence removed.

l. 41: the first ones to report on this issue were Ibrom et al. (2007)

This sentence will be removed as we harmonize the Introduction section and make it more compact.
The sentence was removed from the revised manuscript.

l. 43-44: a rotation into the prevailing wind direction is only one step in the coordinate rotation; typically the aim is to align the coordinate system with the prevailing streamlines (2D or 3D) or with respect to some coordinate system that was established over a longer period (e.g. planar fit)
Corrected.

l. 49: the EC flux is fine – the problem is that it may represent a poor estimate of the surface-atmosphere exchange under these conditions
Corrected.

l. 53: in fact it was the following year (1999) that John Finnigan published a commentary on the Lee (1998) paper in which he demonstrated that correcting only for vertical advection is nonsense
All text and discussion regarding vertical advection will be removed from the revised manuscript.
All text and results regarding vertical advection were removed from the manuscript.

l. 56: if environmental data are lacking too, mean diurnal variation may be used as a last resort
Added mean diurnal variation.

l. 64: if the cross-covariance is flat, then a wrong lag time will not have a large effect
We agree that in case the covariance is flat, lag time does not make a large effect. But in the case when cross-covariance is not flat but noisy, and lag time determination is difficult, it affects the flux magnitude.

l. 66: I do not get the "However, . . ." which links to the previous sentence, which does not appear to make sense here
Removed "However,".

l. 69: Gerdel et al. did study lag determination (their section 3.1) and u*-filtering (their Fig. 6)
It was meant here that different methods for lag time determination have not been studied earlier. We have changed the sentence to "*Gerdel et al. (2017) describes the issues of different detrending methods, high-frequency spectral correction, lag time determination and u* filtering. However, there has not been any study comparing different methods for lag time determination or high-frequency spectral correction in terms of their effects on COS fluxes.*"

l. 71: actually you do not discuss the "EC flux measurement setup" at all
Removed.

l. 71-73: the introduction should finish with a statement of objectives
We will revise the whole introduction section and end it with a clear statement of objectives: "*In this study, we compare different methods for detrending, lag time determination and high-frequency spectral correction. In addition, we compare two methods for storage change flux calculation, discuss the nighttime low turbulence problem in the context of COS EC measurements, introduce a method for gap-filling COS fluxes for the first time and discuss the most important sources of*

random and systematic errors. Through the evaluation of these processing steps, we aim to settle on a set of recommended protocols for COS flux calculation."

l. 79: coordinate rotation
Corrected.

l. 78-88: does this have any relevance for this study?
We think it is important for the reader to know about the measurement site and its characteristics. Also, we want to mention that the same data has been published before in a paper focusing on different aspects than methodology.

l. 92: and sonic temperature
Corrected.

l. 95: flow rate through Li-6262, same pump as QCL, tube diameter, length, is the same tube as for the QCL?
The two instruments had their own inlet tubings. This is now clarified in the text and LI-6262 tubing information added: "*All measurements were recorded at 10 Hz frequency and were made with a flow rate of approximately 10 liters per minute (LPM) for the QCLS and 14 LPM for LI-6262, respectively. The PTFE sampling tubes were 32 m and 12 m long for QCLS and LI-6262, respectively, and both had an inner diameter of 4 mm. Two PTFE filters were used upstream of the QCLS inlet to prevent any contaminants entering the analyzer sample cell: one coarse filter (0.45 µm, Whatman), followed by a finer filter (0.2 µm, Pall corporation), at approximately 50 cm distance to the analyzer inlet. The Aerodyne QCLS used an electronic pressure control system to control the pressure fluctuations in the sampling cell. The QCLS was run at 20 Torr sampling cell pressure. An Edwards XDS35i scroll pump (Edwards, England, UK) was used to pump air through the sampling cell, while LI-6262 had flow control by a LI-670 flow control unit.*"

l. 106: is this the mean? what is the standard deviation?
The standard deviation was determined from the standard deviation of cylinder air measurements. Now clarified in the text: "*The standard deviation was 19 ppt for COS mixing ratios and 1.3 ppm for CO_2 at 10 Hz measurement frequency, as calculated from the cylinder measurements.*"
The sentence revised as : "*The standard deviation calculated from the cylinder measurements was 19 ppt for COS mixing ratios and 1.3 ppm for CO_2 at 10 Hz measurement frequency.*"

l. 111-112: given the precision of typical computer clocks, this will result in clock differences up to several seconds; important to add that most likely the Li-6262 data were acquired by and thus synchronized with the sonic anemometer through it's A/D input?!
Thanks for the comment. The logging system and data flow will be described in more detail in the revised manuscript.

l. 129-131: I am not sure I understand the first step – the QCL clock may be either delayed or advanced with respect to the clock of the sonic anemometer & IRGA and I thus do not understand why you shift the QCL time series by the lag time between w and IRGA CO2? In my understanding you could start off with the second step which actually aligns both time series.
The reviewer is correct, we have removed the first step. Sorry for the misunderstanding, now clarified in the text: "*The following procedure was done to combine two data files of 30 min length*

(of which one includes sonic anemometer and LI-6262 data and the other includes Aerodyne QCLS data): 1) the cross-covariance of the two CO_2 signals (QCLS and LI-6262) was calculated 2) the QCLS data were shifted so that the cross-covariance of the CO_2 signals was maximized. Note that this will result in having the same lag time for QCL and LI-6262. The time shift was a maximum of 10 seconds, most often varying between 0 s to 2 s during one day. It is also possible to shift the time series by maximizing the covariance of CO_2 and w, which will then already account for the lag time (Fig. S1) or combine files according to their time stamps and allow a longer window in which the lag time is searched. However, in this case it is important that the lag time (and time shift) is determined from CO_2 measurements only, as COS data might result in several covariance peaks in longer time frames due to low signal-to-noise ratio and small fluxes."

l. 131: might be worth showing this as a histogram in the supplement?
Time shift (computer drift + QCL lag) is now shown in the supplement Fig. S1.

l. 133-134: shouldn't this sentence come first in this section as this likely was the initial step? Or did the procedure described in l. 126-132 use data before despiking?
File combination was indeed done before despiking.

l. 142-143: not sure that computation time is a relevant issue nowadays with regard to coordinate rotation
Removed the note about computation times.

Fig. 1: the first rotation angle is typically the one that aligns the coordinate system along the main wind direction – why would that rotation be limited to less than 10 ◦ , which would mean rejecting fluxes from 340 ◦ ?
You're right, it was the second rotation angle, as mentioned in the text. Corrected in Fig. 1 now as well.

l. 200: in my memory, the first to propose this approach were Aubinet et al. (2000, 2001)
Already credited earlier, but added the reference here as well.

l. 201: a site-specific cospectral model was already used by Wohlfahrt et al. (2005)
In this line we only refer to Aubinet et al. (2000) as they introduced the technique and to De Ligne et al. (2010) as they compare different frequency response correction methods and make recommendations, but we will add Wohlfahrt et al. (2005) reference to earlier in the text.
The whole section on high-frequency spectral correction (Sect. 2.4.3) was revised. Text mentioned here was revised as "*The analytical correction was based on scalar co-spectra defined in Horst (1997) and the experimental approach was based on the assumption that temperature co-spectrum is measured without significant error and the normalized scalar co-spectra were compared to the normalized temperature co-spectrum (Aubinet et al., 2000; Wohlfahrt et al., 2005; Mammarella et al., 2009)*."

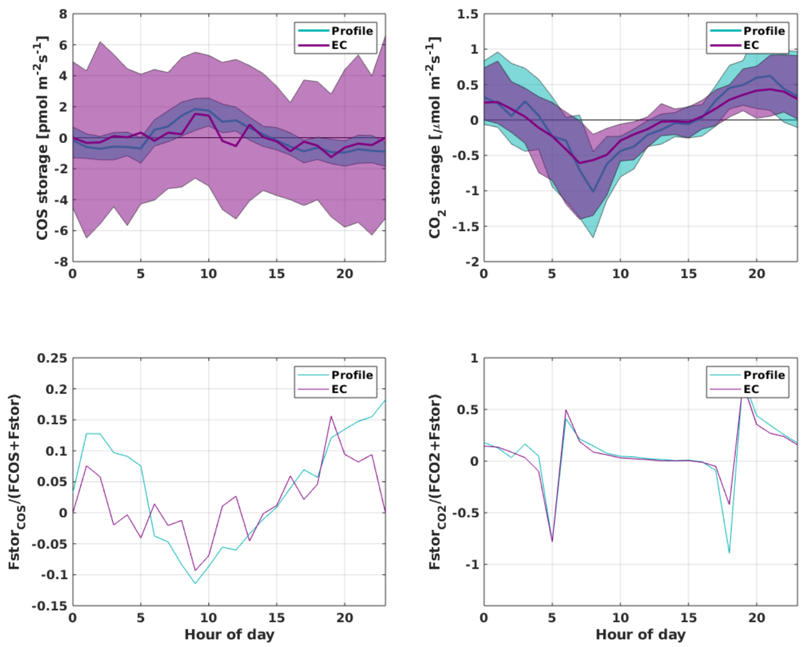
l. 213: which rotation angle?
Second rotation angle, now clarified in the text.

l. 232: but fluxes are available half-hourly – how do you come up with a half-hourly storage change estimate?

The storage change estimate from the concentration profile measurements are hourly and from EC measurements half-hourly. Concentration change used in the calculation is the change that occurred during that 30 min (cafter – cbefore), hence we get half-hourly storage estimates from the EC setup. The profile measurements were subsampled to 30 min to correct for the storage change in measured 30 min fluxes.

l. 237-238: 5 hours sounds like a really long time to reduce random noise – in fact I would expect a 5 hour moving average to even average out true storage; how exactly did you calculate the one-point storage term?
This is the same time window that was used in smoothing profile measurements (Kooijmans et al., 2017) and we chose the same time window for smoothing EC data for better comparison. If a shorter time window is used (e.g. 1 hour), the diurnal shape of the storage change flux does not change but noise of COS storage change flux increases almost three-fold (see figure below).
One-point storage term was calculated by assuming that the concentration change in time is uniform from the measurement height to ground:
FCOSstor = h*ΔC/Δt
Where h is the measurement height (in m) and ΔC/Δt is the concentration change in time (in mol m-3 s-1) at the measurement height.

Below is an example of storage change flux when 1 hour moving average is used for smoothing concentration measurements. Median diurnal variation of COS storage change flux is not different from the 5 hour moving average, but variation has increased notably.



l. 264: clearly here only Lee (1998) is to credit with this approach
All text related to vertical advection will be removed from the revised manuscript.
All text related to vertical advection were removed from the revised manuscript.

l. 270: “missing CO2 fluxes”?

Corrected.

l. 282: add units

Units added.

l. 284-286: isn't this a repetition from above?

Yes, we will incorporate all the relevant information into the methods section.

All relevant information is given in the Methods section and repetition removed in the revised manuscript.

l. 288-289: all measurements are characterised by noise to a certain degree . . .

Yes, we do agree and we have rephrased it as "*As COS measurements are often characterized by low signal-to-noise ratio, the maximization of the absolute value of cross-covariance may determine the lag time from a local maxima, as demonstrated in Fig. 2.*"

This text, along with the former Fig.2, were removed from the revised manuscript.

l. 289-290: I do see several local minima in Fig. 2, but not that any of the tested algorithms gets stuck in one of these

We noticed that there was an artefact related to lag time calculation, as the lag time results were obtained from a lag time optimization tool. In the revised manuscript, we are using the lag times calculated based on the covariance maximization.

In the previous version of the manuscript, we used lag time optimization tool that replaces the lag time with a mean lag time if lag time is detected on the window border. We have reprocessed all fluxes without using the lag time optimization tool. Instead, we are now using a different method for determining the lag time:

"*The time lag between w and COS signals was determined using the following five methods: 1) From the maximum difference of the cross-covariance of the COS mixing ratio and w ($\overline{w'\chi_{\cos}'}$) to a line between covariance values at the lag window limits (referred hereafter as COS lag). This applies also to other covariance methods explained below, and prevents the time lag to be exactly at the lag window limits.*"

l. 298-299: so what? How does that sentence relate to your results?

Thanks for the comment. The sentence will be rephrased accordingly.

l. 305: reformulate – I guess that what you mean is that with the DetLim method, the COS lag was selected in 40 % of all cases (while the CO2 lag was chosen in 60 %)

Reformulated as "*By using the DetLim lag method, the COS lag time was estimated for 40 % of cases from the* $\overline{w'\chi_{\cos}'}$ *covariance maximization, while the* CO_2 *lag was used as proxy for the COS lag in about 60 % of cases.*"

After reprocessing all data, the phrase reads: "*By using the* $DetLim_{lag}$ *method, the COS time lag was estimated for 54 % of cases from* COS_{lag} *, while the* $CO2_{lag}$ *was used as proxy for the COS time lag in about 46 % of cases.*"

l. 305: can you compare lag times of COS and CO2 both fluxes are clearly higher than the flux detection limit? Is there a systematic difference between the two (e.g. COS lag time always longer

than CO2) and if so could the DetLim Method be improved by adding this offset to the CO2 lag instead of just using the CO2 lag?
We tested this idea, but there was no systematic offset found between the lag times.

l. 309: why did you choose the cumulative flux as a metric? Wouldn't the cumulative flux potentially be affected by compensating effects (over- and underestimation during certain conditions resulting in similar cumulative fluxes)? Unlike for CO2, I also do not see much nee to calculate a daily or longer term budget for COS; I also do not see how you can get units of nmol/m2s for COS as for a cumulative flux you need to integrate over time, i.e. multiply the half-hourly flux by 1800 s and then sum these up this then yields molar units per m2 and time period over which the cumulative was calculated – same for CO2.
It is not a cumulative sum as used in budgets, thus not multiplied with 1800s and has units of nmol m-2s-1, i.e. just simply cumulative sum of all measured (NOT gap-filled!) fluxes. We have considered the reviewer's comment on compensating over- and underestimations in cumulative sum and chose to use median fluxes instead. Median fluxes are also affected by the compensating over- and underestimations but not as much to differences in missing data. We will add overall median fluxes as well as separate night-time and daytime median fluxes of all different processing schemes to Table 2.
All median fluxes (total, daytime and night-time, were implemented in Table 2 of the revised manuscript. All medians are calculated to periods when fluxes from all different processing schemes are available, to avoid systematic biases.

Table 1: instead of repeating the values for the reference three times, list it only once?!
Reference fluxes will be added to the table caption and removed from the table.
Rerefence fluxes were added to the table caption and removed from Table 2.

l. 315: how do you know this difference is large on an annual scale? What is the basis for this recommendation?
Changed the sentence to "*This difference might become important in the annual scale, and we recommend using the detection limit method in lag time determination of small fluxes, as in Nemitz et al. (2018)*" for clarity.
After reprossing all data and considering also the previous results, we landed on recommending the CO2 lag instead: "*This difference might become important in the annual scale, and as the most commonly used covariance maximization method does not produce a clear time lag distribution for* $DetLim_{lag}$ *or* COS_{lag}*, we recommend using the* $CO2_{lag}$ *for COS fluxes, as in most ecosystems the* CO_2 *cross-covariance with w is more clear than the cross-covariance of COS and w signals.*"

l. 335-336: "low-frequency corrections"
Corrected as suggested.

l. 336-337: combine both sentences
Combined sentences.

l. 337-338: this is kind of trivial
While this recommendation may be stating the obvious, it is not always carried out in field practice and therefore is worth emphasizing. We decided to keep the sentence, as it adds to the discussion.

l. 344: what is the “normal” CO2 cospectrum? Fig. 7a shows the COS cospectrum, 7b the COS power spectrum – this sentence does not make sense

Corrected the figure reference to 7a and 7c.

The correct figure reference is Fig. 4a and c in the revised manuscript.

Fig. 7: are the (co)spectra in any way filtered for stability or really averaged over the entire month? I suggest to remove the sub-grid lines in all four panels as otherwise these are too busy and become blurry

All the spectra are filtered for stability. Added a sentence to the (co)spectra figure’s caption “*All data are filtered for stabilities -2 < z/L < -0.0625 and COS data only accepted when the covariance was higher than three times the random error due to instrument noise (Eq. 1).*” Gridlines will be removed.

Gridlines were removed. We added the model cospectra from the experimental and Horst (1997) methods to the figure and the caption now reads: “*Cospectrum and power spectrum for COS (a and b, respectively) and CO_2 (c and d) in July 2015. All data were filtered by the stability condition -2 < ζ < -0.0625 and COS data were only accepted when the covariance was higher than three times the random error due to instrument noise (Eq. 1). The cospectrum models by experimental method and Horst (1997), that were used in the high-frequency spectral correction, are shown in grey continuous and dashed lines, respectively.*”

l. 349: in Fig. 7 you are using normalized frequency, thus no units

Changed the text from “3 s” to “*normalized frequency higher than 3*”

l. 350: while the attenuation is clearly visible for CO2, it does not show well for COS, whose cospectrum mostly overlaps with the sensible heat cospectrum

Cospectrum is now plotted for only those times when COS flux surpasses the random noise. While there is still quite a lot of noise in the high frequency end of COS cospectrum, the attenuation is now more visible. Probably the instrument random noise was overlapping with sensible heat cospectrum by chance.

l. 350-351: how was this calculated?

Thanks for the comment. We will add more details in Section 2.4.3. In practice the response time τ_s was determined by fitting a sigmoidal function $1/(1+(2\pi f \tau_s)^2)$ to the ratio between ensemble averaged CO2 and T cospectra.

The modified text regarding response time calculation now reads:

“*In both approaches (analytical and experimental), the time constant τ_s was empirically estimated by fitting the transfer function $T_{ws}(f)$ to the normalized ratio of cospectral densities*

$$T_{ws} = \frac{N_\theta Co_{ws}(f)}{N_s Co_{w\theta}(f)} \qquad (10)$$

where N_θ and N_s are normalization factors and Co_{ws} and $Co_{w\theta}$ the scalar and temperature cospectra, respectively. The estimated time constant was 0.68 s for the Aerodyne QCLS and 0.35 s for LI-6262.”

l. 353: indeed this is as expected . . . remove?

Removed the sentence.

l. 355-357: I think it would be instructive to show at least one characteristic example comparing the experimental and analytical frequency response correction approaches as otherwise this remains a “black box” for the reader

The two correction methods are now fully compared by adding median daytime and night-time fluxes to Table 2, histogram of fluxes to Supplementary material Fig. S2 and stability categorized flux attenuation versus wind speed to Supplementary material Fig. S3. Otherwise, we do not understand what the referee means by saying “to show at least one characteristic example”.

We have added the two different cospectral models to Fig. 4 of the revised manuscript. In addition, a histogram of fluxes is shown in Fig. S6 and flux attenuation factor versus wind speed in Fig. S7 in the supplementary material.

l. 358: while I understand that the experimental frequency response correction approach is part of the standard against which the comparison is made, I think (i) that the no frequency response correction scenario is useless as we know that this leads to a bias and (ii) instead it would be useful to compare the magnitude of the correction between the analytical and experimental approach in order to understand how much of a difference it makes whether one or the other correction is used

The “no correction” option demonstrates how much high frequency corrections affect the final fluxes and is thus left to the analysis. Magnitude of the correction is demonstrated in Table 2 as total median fluxes and daytime and night-time median fluxes. Histogram of the resulting COS fluxes is added to Supplementary material Fig. S2 and flux attenuation versus wind speed to Fig. S3. Final fluxes are also compared as a scatter plot in Fig. S7.

Histogram of the resulting COS fluxes is added to Supplementary material Fig. S6 and flux attenuation versus wind speed to Fig. S7. Final fluxes are also compared as a scatter plot in Fig. S11.

l. 364-369: this section comes a bit as a surprise as it never has been mentioned as a goal before to do this kind of comparison and also lacks a proper discussion – without discussion I rather suggest to remove Fig. 8 and the corresponding text; one point of discussion might be how much the difference in the cospectral reference models contributes to the differences between the analytical and experimental approach; btw., the results in Fig. 8 are qualitatively consistent with Fig. 11 of Wohlfahrt et al. (2005)

As this was not a very important part of our analysis, we have now moved the figure to supplementary material (Fig. S4), as suggested, and removed the corresponding text. Equations 15 and 16 of the former version were moved to Methods-section and are now presented as equations 6 and 7 in the present version of the manuscript.

Former Fig. 8 was moved to supplementary material Fig. S8 and all text to Methods section.

l. 371-377: here I have the impression that you not describing the actual magnitude of the correction, but rather what we might expect based on one example shown in Fig. 5?!

This correction is based on theoretical transfer functions by Rannik & Vesala (1999), and the range of magnitude of this correction is maximum 15 %. The correction is not related to setup or site specific, it is a general and theoretical correction, and expected to be similar to Rannik & Vesala (1999) study. It does not make sense to only correct for low frequency loss in this study to check the actual magnitude of the correction. We will remove this section from the revised manuscript and move Fig. 5 to supplementary material.

Discussion on low frequency spectral correction was left out of the revised manuscript.

l. 372-374: this expectation would only be justified if the algorithm used for correcting for flux loss at lower frequencies would “know” of the noise, which I expect it does not

Exactly, it does not know of the noise and that’s why cannot make large corrections. We have decided to leave out the result section on low frequency spectral corrections.

Low frequency spectral correction was left out of the revised manuscript.

Figure 9: to what period do the data shown refer to?

To the whole measurement period. Clarified in the figure caption “*...during the measurement period 26 June to 2 November 2015.*”

l. 400-401: again, this is well established

We agree with the Reviewer, yet storage change flux is neglected in most of COS EC studies (Table 1). Thus, we feel it is important to emphasize that for diurnal variation it cannot often be neglected. We revised the sentence to emphasize this point: “*In conclusion, the storage change fluxes are not relevant for budget calculations – as expected – and have not been widely applied in previous COS studies (Table 1). However, storage change fluxes are important in the diurnal scale to account for the delayed capture of fluxes by the EC system under low turbulence conditions.*”

Sentence now reads: “*In conclusion, the storage change fluxes are not relevant for budget calculations – as expected – and have not been widely applied in previous COS studies (Table 1), even though storage change flux measurements are mandatory in places where the EC system is placed at a height of 4 m or above according to the ICOS protocol for EC flux measurements (Montagnani et al., 2018). In addition, storage change fluxes are important in the diurnal scale to account for the delayed capture of fluxes by the EC system under low turbulence conditions.*”

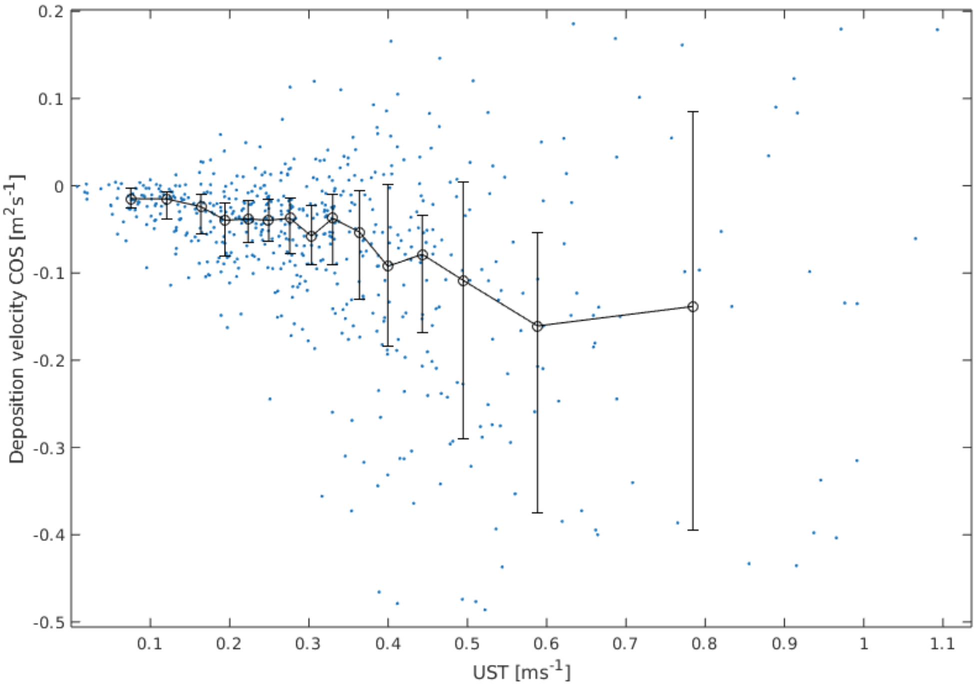
l. 418: I would not call data erroneous – they simply do not fulfil the assumptions underlying our simplified model of surface-atmosphere exchange

Modified the sentence to “*...reliable tool to filter out data that is not representative of the surface-atmosphere exchange under low turbulence (Aubinet et al., 2010).*”

l. 421-423: with this reasoning, wouldn’t it make sense to get rid of the concentration dependency by using the deposition velocity, i.e. the flux normalized with the concentration, instead of the flux itself? Or perform the analysis with data stratified by COS concentration to minimize the issue?

Thank you for this suggestion! We made the u* plot using the deposition velocity (EC flux + storage change flux normalized with concentration gradient), but the u* dependency did not disappear. We added the following text to the manuscript: “*However, we did not see u_* dependency disappearing even with a concentration gradient-normalized flux, so the u_* filtering is applied here normally to overcome the EC measurement limitations under low turbulence conditions.*”

The sentence now reads: “*However, as we did not see u_* dependency disappearing even with a concentration gradient-normalized flux, the u_* filtering is applied here normally to overcome the EC measurement limitations under low turbulence conditions.*”



l. 428-437: knowing that the u*-filtering needs to be applied on the sum of the storage and EC flux, why do you still give the numbers for the EC flux without storage?

There are quite many COS studies that use *u** filtering even though they neglect (or don't mention) the storage change flux: Asaf et al. 2013, Billesbach et al. 2014, Maseyk et al. 2014, Commane et al. 2015, Yang et al. 2018 (and not sure how it was done in Spielmann et al. 2019, as it is not mentioned). So even though this is somewhat trivial for EC measurements, it has not been implemented widely in COS studies, and is worth mentioning here.

l. 434-435: circular argument - without applying the storage correction it is pretty clear that the storage flux is clear that the storage flux is ignored

Reformatted the sentence to "*If fluxes are not corrected for storage before deriving the u* threshold, there is a risk of flux overestimation due to double accounting. The flux data filtered for low turbulence would be gap-filled, thereby accounting for storage by the canopy, but then accounted for again when the storage is released and measured by the EC system during the flushing hours in the morning (Papale et al. 2006).*"

Figure 11: also shows the vertical advection without being mentioned in the figure legend?!

Vertical advection was mentioned in the figure legend, but was missing from the caption. We have decided to leave out the section regarding vertical advection, according to the reviewer's suggestions, and thus removed also from this figure.

l. 442-444: is this a useful comparison? The cumulative COS flux must be less negative if missing data (which are generally negative) are not gap-filled! In order to evaluate the skill of the gap-filling algorithm the authors need to create artificial (but realistic) gaps in their time series (see CO2 flux literature to that end) and then compare the gap-filled against the measured (during the artificially created gaps) fluxes!

This comparison is just demonstrating how much change there is when using gap-filled versus non gap-filled fluxes in a cumulative sum, and the magnitude of course depends on the number of gaps in the data. We will add a diurnal plot of the gap-filling algorithm and measured fluxes (Fig. 11), a plot of the residuals of different gap-filling algorithms (Fig. S6) and a time series of the gap-filling parameters (Fig. S5) to the revised manuscript and supplementary material.
Diurnal variation of the measured COS flux and three different gap-filling algorithms are shown in Fig. 7 of the revised manuscript. Residuals of different gap-filling algorithms and time series of the gap-filling parameters are shown in Figs. S9 and S10, respectively.

l. 474-475: why do you recommend the experimental high-frequency correction approach? Is it more accurate? If so this needs to be demonstrated! What about the performance of the experimental approach in situations with low/noisy sensible heat cospectra and what about the effect of the QCL on the ratio central to the experimental approach?
We did not use single 30 min T cospectra for the model cospectra. Only the averaged cospectra (when the cospectra was good) were used for creating the model (see Sabbatini et al., 2018 for reference). A site specific cospectral model is suggested to be used instead of analytical ones in several studies (Aubinet et al. 2000, De Ligne et al. 2010), and especially in the new eddy covariance data processing protocols (Sabbatini et al., 2018; Nemitz et al., 2018). We will provide both Horst (1997) and experimental cospectral models in the cospectrum figure, together with the mean cospectrum.
The changes mentioned above were implemented in the revised manuscript.

Reviewer #2

Erkkilä (Kohonen) et al present a detailed and valuable analysis of the impact of various eddy covariance data processing options on the calculated ecosystem uptake of carbonyl sulfide (COS). They attempt to quantify the flux uncertainty deriving from the data processing, and they make recommendation for some of the options. The methods are sound and the manuscript is fairly well written and easy to follow. I do have some concerns about the analysis and interpretation of the results. In addition to the “major comments” of referee Wohlfahrt, with which I agree, I believe the paper could be made stronger by addressing the issues below.

Specific Scientific Comments
1. I disagree with the idea that the “processing uncertainty” reported here is actually an uncertainty in the calculated fluxes. Instead, it is a metric of the sensitivity of the calculated fluxes to different processing choices. Some of those choices are clearly better than others, and it doesn’t make sense to calculate the fluxes in a way that is known to be pretty good and in another way that is known to be pretty bad and then say that the difference between the two ways is the uncertainty in the flux. In particular, the following data processing choices are obviously bad: (a) the COS lag and RM lag methods, (b) the RF 30s detrending method, (c) omitting high-frequency correction, (d) omitting the storage flux, (e) determining a u* filter threshold before including the storage flux, (f) omitting gap filling (for cumulative sums). None of those methods should be included when assessing methodological uncertainty, as there is no uncertainty about the fact that those methods should not be used. Thus the “processing uncertainties” presented are misleading, in that they give an inflated impression of the real processing uncertainty in the EC method. Moreover, I think the total “processing uncertainty” in Fig. 12 is of no use even as a sensitivity metric, as it blends

sensitivity to choices that are unclear with sensitivity to other choices that are very clear. (Similarly, the total uncertainty defined in Section 2.4.6 is of limited use because it blurs the distinction between stochastic half-hourly noise, which can be averaged out, and long-term systematic bias, which cannot.) So I would present instead (and show in Fig. 12) the flux sensitivities to the various individual processing choices. Then if the authors want to identify which processing choices are genuinely debatable and use their sensitivities to calculate a more meaningful overall processing uncertainty, they can do that. And then if they want to compare the magnitude of the systematic processing uncertainty (i.e. potential bias) to that of the random flux noise, they can do that too (But why? Over what noise averaging period is such a comparison meaningful?).
The uncertainty estimates presented here are based on the well established eddy covariance data processing protocols, presented in Sabbatini et al., 2018. We agree that processing choices listed in the comments are not as good as others. In this method, we calculate the fluxes using block averaging with planar fitting, block averaging with 2D wind rotation, linear detrending with planar fitting and linear detrending with 2D wind rotation. These are all very widely used processing schemes and not thought as obviously bad. The bad choices listed in the comment are not used for estimating processing uncertainty. The estimate of processing uncertainty is based on calculating the lowest and highest possible fluxes that come out from different (reliable) processing schemes and thus tells how much variation there can be in fluxes due to processing choices. The method is explained in detail in Section 2.4.6. However, as random error is dominating the total uncertainty, even a loose definition of the processing uncertainty would not inflate the total uncertainty to a large degree.

2. The distributions of lag times in Fig. 3b and 3d are concerning. Why the spike at 2.7 s, in the wing of a broad peak centered on 3.2 s? The spike seems to suggest that the true lag was actually always 2.7 s, while the other retrieved lags were in error, perhaps due to some stochastic noise artifact. After all, if the true lag really were varying stochastically as suggested by the broad peak, then why would there be a preponderance of times when it was exactly 2.7 s? Was there perhaps a change in the experimental setup at some point during the measurement period? After seeing Figs. 3 and 4, I'm actually inclined to think that using constant lag is the most advisable option for these data. The authors instead recommend the DetLim method but do not justify that recommendation. In particular, it's unclear why the lag determined from CO2 should ever be any worse than that determined from COS (except when the CO2 flux crosses zero), given that both gases are measured by the same instrument and the CO2 almost always has a higher signal to noise ratio.
Thank you for pointing this out. We checked the lag time issue and the spike at 2.7 s and winged shape were due to an artefact caused by a lag time optimization tool used in the final flux calculation. In the revised manuscript we will show the lag times determined directly from the maximum covariance, which doesn't have this artefact. Using CO2 lag time is probably not any worse from DetLim lag, and we found a more clear lag time distribution for CO2 lag than DetLim lag. The DetLim lag method was established in the eddy covariance data processing protocol by Nemitz et al., 2018 for small fluxes, but to our knowledge the method is implemented here for the first time. We will add a more clear statement in the revised manuscript.
In the previous version of the manuscript, we used lag time optimization tool that replaces the lag time with a mean lag time if lag time is detected on the window border, and this caused the weird behaviour in the lag time distributions. We have now reprocessed all fluxes without using the lag time optimization tool. Instead, we are now using a different method for determining the lag time:
"*The time lag between w and COS signals was determined using the following five methods:*
1) From the maximum difference of the cross-covariance of the COS mixing ratio and w ($\overline{w'\chi_{\cos}'}$ *) to a line between covariance values at the lag window limits (referred hereafter as*

COS lag). This applies also to other covariance methods explained below, and prevents the time lag to be exactly at the lag window limits."

After analysing data from this method and comparing to the results of the previous processing method, we have decided to recommend the $CO2_{lag}$ instead of the $DetLim_{lag}$. The reviewer has a good point saying "In particular, it's unclear why the lag determined from CO2 should ever be any worse than that determined from COS (except when the CO2 flux crosses zero), given that both gases are measured by the same instrument and the CO2 almost always has a higher signal to noise ratio." especially, when the two molecules are chemically very similar. We have taken this into account in the revised manuscript and rewrote the recommendation on lag time as "*This difference might become important in the annual scale, and as the most commonly used covariance maximization method does not produce a clear time lag distribution for $DetLim_{lag}$ or COS_{lag}, we recommend using the $CO2_{lag}$ for COS fluxes, as in most ecosystems the CO_2 cross-covariance with w is more clear than the cross-covariance of COS and w signals.*"

3. Surprisingly, and despite the statement on lines 344 and 350, Fig. 7a seems to show that the COS cospectra don't seem to have any high-frequency signal loss, unlike the CO2 cospectra. I can think no reason why that should be the case unless the large high-frequency instrument noise for COS is synchronizing by chance with high-frequency fluctuations in w. Given that the COS cospectra seem to match well with the temperature cospectra, it doesn't seem to make sense to use the CO2 cospectral correction (based on the mismatch between the CO2 and temperature cospectra) for
COS. Unless perhaps Fig. 7a mistakenly shows COS cospectra after correction?
We think that it was indeed by chance that the instrument noise was synchronizing with high frequency fluctuations. We have now made the same plot only for COS fluxes that surpass the random noise (covariance at least three times higher than the random noise) and the flux attenuation is more evident and closer to that of CO2.

Technical Corrections
- line 2: "the recent development" should be "recent developments"
Corrected

- line 21: "not being" should be "is not"
Corrected

- line 22: "for radiation-dependency" should be "for the radiation-dependency"
Corrected
This sentence was removed from the revised manuscript.

- lines 60-61: The word "respectively" doesn't make sense here, as there's nothing for the analyzers to be respective to. I recommend changing "at 10 Hz from Aerodyne Research (Billerica, MA, USA) and Los Gatos Research (San Jose, CA, USA), respectively" to "at 10 Hz, one from Aerodyne Research (Billerica, MA, USA) and one from Los Gatos Research (San Jose, CA, USA)".
We will revise the introduction based on reviewer comments, and have left this part out of the text to harmonize and make the introduction more compact.
The sentence was left out of the revised manuscript.

- line 64: I would delete "a basis of EC measurements"
Deleted

- line 112: "a home-made" should be just "home-made"

Corrected

- line 149: “to some extent of weather changes” should be “to some extent weather changes”
Corrected as suggested

- line 156: “different” is superfluous, and so I would delete it
Deleted

- line 178: “others” should be “other reasons”
Corrected

- line 344: “compare Fig. 7a and 7b” should be “compare Fig. 7a and 7c”
Corrected
Corrected as Fig. 4a and c in the revised manuscript.

- lines 423 ff: “The uâĹŮ filtering is applied to conform the. . . does not make sense and I’m not sure exactly what you are trying to say here.
The whole revised sentence/chapter reads “*The u* filtering is applied to conform the assumption that fluxes do not go down under low turbulence conditions, as is the case for respiration of* CO_2*, but which does not necessarily apply to COS uptake. The u* filtering may therefore bias COS fluxes due to false assumptions. However, we did not see u* dependency disappearing even with a concentration gradient-normalized flux, so the u* filtering is applied here normally to overcome the EC measurement limitations under low turbulence conditions.*”
The sentence is revised as “*Thus, the assumption that fluxes do not go down under low turbulence conditions, as is the case for respiration of* CO_2 *, does not necessarily apply to COS uptake. The* u_* *filtering may therefore bias COS fluxes due to false assumptions. However, as we did not see* u_* *dependency disappearing even with a concentration gradient-normalized flux, the* u_* *filtering is applied here normally to overcome the EC measurement limitations under low turbulence conditions.*”

Towards standardized processing of eddy covariance flux measurements of carbonyl sulfide

Kukka-Maaria Kohonen[1], Pasi Kolari[1], Linda M.J. Kooijmans[2], Huilin Chen[3], Ulli Seibt[4], Wu Sun[4,5], and Ivan Mammarella[1]

[1]Institute for Atmospheric and Earth System Research (INAR)/ Physics, Faculty of Science, University of Helsinki, Helsinki, Finland
[2]Meteorology and Air Quality group, Wageningen University and Research, Wageningen, the Netherlands.
[3]Centre for Isotope Research (CIO), University of Groningen, Groningen, the Netherlands
[4]Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, California, USA
[5]Present address: Department of Global Ecology, Carnegie Institution for Science, Stanford, California, USA

Correspondence: Kukka-Maaria Kohonen (kukka-maaria.kohonen@helsinki.fi)

Abstract. Carbonyl sulfide (COS) flux measurements with the eddy covariance (EC) technique are ~~growing in popularity with the recent development in using COS to estimate gross photosynthesis at the ecosystem scale. Flux data intercomparison would benefit from~~ becoming popular for estimating gross primary productivity. To compare COS flux measurements across sites, we need standardized protocols for ~~COS flux~~ data processing. In this study, we analyze how various data processing steps affect the ~~final~~ calculated flux and provide a method for gap-filling COS fluxes. Different methods for determining the ~~lag time~~ time lag between COS mixing ratio and the vertical wind velocity (w) resulted in a maximum of ~~12~~ 15.9 % difference in the ~~cumulative COS flux~~ median COS flux over the whole measurement period. Due to limited COS measurement precision, small COS fluxes (below approximately 3 pmol m^{-2} s^{-1}) could not be detected when the ~~lag time~~ time lag was determined from maximizing the covariance between COS and w. We recommend using ~~a combination of COS and~~ carbon dioxide (CO_2) ~~lag times~~ time lag in determining the COS flux, ~~depending on the flux magnitude compared to the detection limit of each averaging period. Different high frequency spectral corrections had a maximum effect of 10 % on~~ due to higher signal-to-noise ratio of CO_2 measurements. The difference between two high-frequency spectral corrections was 2.7 % in COS flux calculations, whereas omitting the high-frequency spectral correction resulted in a 14.2 % lower median flux and different detrending methods ~~only 1.2~~ caused a spread of 6.2 %. Relative total uncertainty was more than five times higher for low COS fluxes (lower than ±3 pmol m^{-2}s^{-1}) than for low CO_2 fluxes (lower than ±1.5 μmol m^{-2}s^{-1}), indicating a low signal-to-noise ratio of COS fluxes. Due to similarities in ecosystem COS and CO_2 exchange ~~, and~~ we recommend applying storage change flux correction and friction velocity filtering normally, but due to the low signal-to-noise ratio of COS fluxes~~that is similar to methane~~, we recommend ~~a combination of~~ using CO_2 ~~and methane flux processing protocols for COS EC~~ data for time lag and high-frequency corrections of COS fluxes.

1 Introduction

Carbonyl sulfide (COS) is the most abundant sulfur compound in the atmosphere, with tropospheric mixing ratios around 500 ppt (Montzka et al., 2007). During the last decade, studies on COS have grown in number, mainly ~~due to~~ motivated by the use of COS exchange as a tracer for photosynthetic carbon uptake (also known as gross primary ~~production~~productivity, GPP) ~~(Sandoval-Soto et al., 2005; Blonquist et al., 2011; Asaf et al., 2013)~~(Sandoval-Soto et al., 2005; Blonquist et al., 2011; Asaf et al., 2013; . COS shares the same diffusional pathway in ~~the~~ leaves as carbon dioxide (CO_2), ~~and~~ but in contrast to CO_2, COS is destroyed completely by hydrolysis and ~~not being~~ is not emitted. This ~~uni-directional~~ one-way flux makes it a promising proxy for GPP ~~, particularly in light of recent studies that account for radiation-dependency of COS to CO_2 uptake rates (Yang et al., 2018; Kooijmans et al., 2019)~~(Asaf et al., 2013; Yang et al., 2018; Kooijmans et al., 2019).

Eddy covariance (EC) measurements are the backbone of gas flux measurements at the ecosystem scale (Aubinet et al., 2012). Protocols for instrument setup, monitoring and data processing have been ~~established especially~~ recently harmonized for CO_2 (Rebmann et al., 2018; Sabbatini et al., 2018) ~~but recently also for other greenhouse gases, such as~~ as well as for methane (CH_4) and nitrous oxide (N_2O) (Nemitz et al., 2018) ~~, due to the growing network of~~ within the Integrated Carbon Observation System (ICOS) flux stations (Franz et al., 2018).

~~To meet the assumptions for EC measurements, the experimental setup needs to be carefully designed. The site should be homogeneous, so that the footprint area solely consists of the ecosystem type of interest, and the measurement height should be chosen such that it is well above any obstacles (such as canopy top). Moreover, measurements should be made in the well-mixed surface layer but still low enough that the stable atmosphere would not extend the footprint away from the desired location (Aubinet et al., 2012). Both the gas analyzer and sonic anemometer need to operate at a high frequency, usually 10 or 20 Hz, to capture the fast turbulent fluctuations.~~

~~Gas analyzers used for EC measurements can be divided into open-path analyzers, where gas analysis happens in the open air, and closed-path analyzers, where the gas sample is pumped into an enclosed sample cell through a sampling tube. Closed-path analyzers need filters in the sampling line to prevent clogging of sample tubing and contamination of instrument components. One coarser filter is recommended to be placed close to the inlet while a finer filter should be placed close to the analyzer (Nemitz et al., 2018). The instrumentation needs regular checks to maintain data quality and to prevent long measurement gaps. Other considerations include, e.g., regular cleaning or changing of the sampling lines to prevent adsorption or desorption of water vapor on tubing walls (Mammarella et al., 2009).~~

~~Standardized protocols have been developed for CO_2~~ EC data processing ~~, which follows the steps detailed in Sabbatini et al. (2018):~~ chain includes despiking and filtering raw data~~to eliminate spikes and erroneous data, coordinate~~, rotation of the ~~wind components according to the prevailing wind direction~~coordinate system to align with the prevailing streamlines, determining the time lag between the sonic anemometer and the gas analyzer signals, trend removal ~~(~~to separate the turbulent fluctuations from the mean trend~~)~~, calculating covariances and correcting for flux losses at low and high frequencies. ~~Flux post-processing then includes quality filtering and quality flagging~~ After processing, fluxes are quality filtered and flagged according to atmospheric turbulence characteristics and stationarity.

~~During nighttime, decoupling layers may form between the measurement height and the surface in a stable atmosphere. In conditions of low atmospheric turbulence, the EC flux is usually underestimated and thus needs to be either removed or corrected. The nighttime problem is well studied in the EC community (Aubinet et al., 2003, 2005, 2010) but not solved yet. In addition to decoupling layers, advective transport may become important in stable conditions. Vertical advection has been used to explain the nighttime flux observations (Lee, 1998; Mammarella et al., 2007; Rannik et al., 2009), but with the lack of horizontal advection measurements is not recommended as a correction anymore (Aubinet et al., 2003). Storage change flux correction and friction velocity (u_*) filtering are the most commonly used methods for correcting and filtering out low-turbulence occasions. Stable stratification may also enlarge the flux footprint so that it extends across other surface cover types. The final step of the post-processing is then filling the gaps in the data using gap-filling functions based on environmental data.~~

Studies on ecosystem COS flux measurements with the ~~eddy covariance (EC)~~ EC technique are still limited ~~(Asaf et al., 2013; Billesbach~~ (Asaf et al., 2013; Billesbach et al., 2014; Maseyk et al., 2014; Commane et al., 2015; Gerdel et al., 2017; Wehr et al., 2017; Yang et al., 2 and there is no standardized flux processing protocol for COS EC fluxes. ~~Currently there are two commercial models of closed-path spectroscopic analyzers for measuring COS at 10 Hz from Aerodyne Research (Billerica, MA, USA) and Los Gatos Research (San Jose, CA, USA), respectively~~Table 1 summarizes the processing steps used in earlier studies. Most studies do not report all necessary steps, and in particular, often ignore the storage change correction. COS EC flux measurements and data processing have similarities with other trace gases (e.g., CH_4 and N_2O) that often have low signal-to-noise ratios, especially regarding ~~lag time~~ time lag determination and frequency response corrections. ~~Accurate lag time~~ Time lag determination is essential for ~~correctly synchronizing the~~ aligning wind and gas concentration measurements ~~, a basis of EC measurements, and an incorrect lag time would bias the~~ to minimize biases in flux estimates. Frequency response corrections, on the other hand, are needed for correcting flux underestimation due to signal losses both at high and low frequencies ~~and are always increasing the fluxes (Aubinet et al., 2012). However, unlike~~ (Aubinet et al., 2012). Unlike for CH_4 or N_2O, there are no sudden bursts or sinks expected for COS, and in that sense some ~~of the processing steps~~ processing steps for COS are more like those for CO_2 (e.g., despiking, storage change correction and friction velocity (u_*) filtering). Gerdel et al. (2017) describes the issues of different detrending methods~~and~~, high-frequency spectral correction, time lag determination and u_* filtering. However, there has not been any study ~~on the lag time determination and the nighttime measurement issues~~comparing different methods for time lag determination or high-frequency spectral correction in terms of their effects on COS fluxes. This weakens our ability to assess uncertainties in COS flux measurements.

In this study, we ~~provide suggestions for a robust EC flux measurement setup for COS studies, give recommendations for processing COS EC flux data~~ compare different methods for detrending, time lag determination and high frequency spectral correction. In addition, we compare two methods for storage change flux calculation, discuss the nighttime low turbulence problem in the context of COS EC measurements, introduce a method for gap-filling COS fluxes for the first time and discuss the most important sources of random and systematic errors. ~~In addition, we introduce a method for gap-filling COS fluxes for the first time~~Through the evaluation of these processing steps, we aim to settle on a set of recommended protocols for COS flux calculation.

Table 1. Processing steps used in previous COS eddy covariance studies. Detrending methods include linear detrending (LD), block averaging (BA), and recursive filtering (RF). Spectral corrections include an experimental method and a theoretical method by Moore (1986).

Reference	Sampling frequency	Detrending	Time lag	Spectral corrections	Storage change correction	u_* filtering
Asaf et al. (2013)	1 Hz	LD	max $\lvert\overline{w'\chi'_{COS}}\rvert$	Exp. method	-	-
Billesbach et al. (2014)	10 Hz	-	max $\lvert\overline{w'\chi'_{COS}}\rvert$	Moore (1986)	-	$u_* > 0.15$ ms^{-1}
Maseyk et al. (2014)	10 Hz	-	max $\lvert\overline{w'\chi'_{COS}}\rvert$	Moore (1986)	-	$u_* > 0.15$ ms^{-1}
Commane et al. (2015)	4 Hz	BA	max $\lvert\overline{w'\chi'_{H2O}}\rvert$	-	Neglected	$u_* > 0.17$ ms^{-1}
Gerdel et al. (2017)	10 Hz	BA/ LD/ RF	max $\lvert\overline{w'\chi'_{COS}}\rvert$	Exp. method	EC meas.	$u_* > 0.12$ ms^{-1}
Kooijmans et al. (2017)	10 Hz	LD	max $\lvert\overline{w'\chi'_{CO2}}\rvert$	Exp. method	Profile meas.	$u_* > 0.3$ ms^{-1}
Wehr et al. (2017)	4 Hz	BA	-	Exp. method, CO_2 spectrum	Profile meas.	$u_* > 0.17$ ms^{-1}
Yang et al. (2018)	1 Hz	-	-	-	Neglected	-
Kooijmans et al. (2019)	10 Hz	LD	max $\lvert\overline{w'\chi'_{CO2}}\rvert$	Exp. method	Profile meas.	$u_* > 0.3$ ms^{-1}
Spielmann et al. (2019)	5 or 10 Hz	LD	max $\lvert\overline{w'\chi'_{COS}}\rvert$	Exp. method	-	-

2 Materials and Methods

In this study we utilized used COS and CO_2 EC flux datasets collected in at the Hyytiälä ICOS station in Finland from 26 June to 2 November 2015. The site has a long history of flux and concentration observations (Hari and Kulmala, 2005) and a COS analyzer was introduced to the site in March 2013. In this section, we go through all processing steps in EC flux processing. In the next section, the different processing schemes are compared to a "standard scheme", which consists of linear detrending, 2D wind rotation, using CO_2 lag time for COS and experimental spectral correction. The dataset used in this study was partly published by Kooijmans et al. (2017), where only the nighttime data is utilized with the standard processing schemedescribe methods used in the reference and alternative data processing schemes.

2.1 Site description

Measurements were made in a boreal Scots pine (*Pinus sylvestris*) stand in at the Station for Measuring Forest Ecosystem-Atmosphere Relations (SMEAR II) in Hyytiälä, Finland (61°51′ N, 24°17′ E, 181 m above sea level). The Scots pine stand was established in 1962 and reaches at least 200 m to all directions and about 1 km to the north (Hari and Kulmala, 2005). The site is characterized by modest height variation and an oblong lake is situated about 700 m to the southwest of the forest station (Rannik, 1998; Vesala et al., 2005). Canopy height was 17 m and the all-sided leaf area index (LAI) was approximately 8 m^{2}m^{-2} in 2015. EC measurements were done at 23 m height. The site became a certified ICOS class 1 atmospheric station in 2017 and class 1 ecosystem station in 2018. Sunrise time varied from 2:37 am in June to 7:55 am in November, while sunset was at 10:14 pm in the beginning and 4:17 pm in the end of the measurement period(all times presented in . All results are presented in Finnish winter time (UTC +2, local winter time)) and nighttime is defined as periods when photosynthetic photon flux density PPFD < 3 μmol m^{-2}s^{-1}.

2.2 EC measurement setup

The EC setup consisted of an ultrasonic anemometer (Solent Research HS1199, Gill Instruments Ltd., England, UK) for measuring wind speed in three dimensions and sonic temperature, an Aerodyne quantum cascade laser spectrometer (QCLS) (Aerodyne Research Inc., Billerica, MA, USA) for measuring COS, CO_2 ~~, carbon monoxide (CO)~~ and water vapor (H_2O) mole fractions and an LI-6262 infrared gas analyzer (Licor, ~~Nebraska~~Lincoln, NE, USA) for measuring H_2O and CO_2 mole fractions. All measurements were ~~done~~ recorded at 10 Hz frequency and were made with a flow rate of approximately 10 ~~liter~~ liters per minute (LPM) for the QCLS and 14 LPM for LI-6262, respectively. The PTFE sampling ~~tube was~~ tubes were 32 m ~~long~~ and 12 m long for QCLS and LI-6262, respectively, and both had an inner diameter of ~~8~~ 4 mm. Two PTFE filters were used upstream of the ~~instrument~~ QCLS inlet to prevent any contaminants from entering the analyzer sample cell: one coarse filter (0.45 μm, Whatman), followed by a finer filter (0.2 μm, Pall corporation), at approximately 50 cm distance to the analyzer inlet. The Aerodyne QCLS ~~uses~~ used an electronic pressure control system to control the pressure fluctuations in the sampling cell. The QCLS was ~~running~~ run at 20 Torr sampling cell pressure. An Edwards XDS35i scroll pump (Edwards, England, UK) was used to pump air through the sampling cell, while LI-6262 had flow control by a LI-670 flow control unit.

Background measurements of high-purity nitrogen (N_2) were done every 30 min for 26 s to remove background spectral structures in the QCLS (Kooijmans et al., 2016). In addition, a calibration cylinder was sampled each night at 00:00:45 for 15 s. The calibration cylinder consisted of COS at 429.6 $\pm$ 5.6 ppt, CO_2 at 408.37 $\pm$ 0.05 ppm and CO at 144.6 $\pm$ 0.2 ppb. The cylinder was calibrated against the NOAA-2004 COS scale, WMO-X2007 CO_2 scale and WMO-X2004 CO scale using cylinders that were calibrated at the Center for Isotope Research of the University of Groningen in the Netherlands (Kooijmans et al., 2016). The standard deviation calculated from the cylinder measurements was 19 ppt for COS mixing ratios and 1.3 ppm for CO_2 at 10 Hz measurement frequency~~, as calculated from the cylinder measurements~~.

It has previously been shown that water vapor in the sample air can affect the measurements of COS through spectral interference of the COS and H_2O absorption lines (Kooijmans et al., 2016). This spectral interference was corrected for by fitting the COS spectral line separately from the H_2O spectral line.

The computer embedded in the Aerodyne QCLS and the computer that controlled the sonic anemometer and logged the LI-6262 data were synchronized once a day with a separate server computer. ~~Data were logged~~ Analogue data signals from LI-6262 were gathered to the Gill anemometer sensor input unit, which digitised the analogue data and appended it to the digital output data string. Digital Aerodyne data were collected to the same computer with a serial connection and recorded in separate files with ~~a~~ home-made software (COSlog).

2.3 Profile measurements

Atmospheric concentration profiles were measured with another Aerodyne QCLS at a sampling frequency of 1 Hz. Air was sampled at 5 heights: 125 m, 23 m, 14 m, 4 m and 0.5 m. A multi-position Valco valve (VICI, Valco Instruments Co. Inc.) was used to switch between the different profile heights and calibration cylinders. Each measurement height was sampled for 3 min each hour. One calibration cylinder was measured twice for 3 min each hour to correct for instrument drift, and two other calibration cylinders were measured once for 3 min each hour to assess the long-term stability of the measurements. A background spectrum was measured once every six hours using high-purity nitrogen (N 7.0) (for more details, see Kooijmans et al. (2016)). The overall uncertainty of this analyzer was determined to be 7.5 ppt for COS and 0.23 ppm for CO_2 at 1 Hz frequency (Kooijmans et al., 2016). The measurements are described in more detail in Kooijmans et al. (2017).

2.4 Eddy covariance fluxes

In this section, we go through describe the processing steps of EC flux calculation from raw data handling to final flux gap-filling and uncertainties. Fig. 1 provides a graphical outline of all processing steps. The different processing options presented here are applied and discussed in SeeSect. 3. In the next section, the different processing schemes are compared to a "reference scheme", which consists of linear detrending, planar fit coordinate rotation, using CO_2 time lag for COS and experimental spectral correction. A subset of the data – nighttime fluxes processed with the reference scheme – was published in Kooijmans et al. (2017).

2.4.1 Pre-processing

For flux calculation, the sonic anemometer and gas analyzers analyzer signals need to be synchronized. This is particularly relevant for fully digital systems where digital data streams are gathered from different instruments that can be completely asynchronous to each other (Fratini et al., 2018). The following procedure was done used to combine two data files of 30 min length (of which one includes sonic anemometer and LI-6262 data and the other includes Aerodyne QCLS data): 1) QCLS data were first forced to have the same time lag as LI-6262 CO_2, 2) then the cross-covariance of the two CO_2 signals (QCLS and LI-6262) was calculated 3) finally 2) the QCLS data were shifted so that the cross-covariance of the CO_2 signals was maximized. Note that this will result in having the same time lag for QCL and LI-6262. The time shift was maximum between the two computers was a maximum of 10 seconds, most often with most varying between 0 s to 2 s during one day. It is also possible to shift the time series by maximizing the covariance of CO_2 and w, which will then already account for the time lag (Fig. S1) or combine files according to their time stamps and allow a longer window in which the time lag is searched. However, in this case it is important that the time lag (and computer time shift) is determined from CO_2 measurements only, as using COS data might result in several covariance peaks in longer time frames due to low signal-to-noise ratios and small fluxes.

Raw data were then despiked so that the difference between subsequent data points was maximum 5 ppm for CO_2, 1 mmol mol^{-1} for H_2O, 12 ppb for CO and 200 ppt for COS and 5 ms^{-1} for w. After despiking, the missing values were gap-filled by linear interpolation.

We used the 2D coordinate rotation planar fit method to rotate the coordinate frame so that the turbulent flux divergence is as close as possible to the total flux divergence. First, the average u-component was forced to be along the prevailing wind direction. The second rotation was performed to force the mean vertical wind speed ($\bar{w}$) to be zero (Kaimal and Finnigan, 1994) . In this way, the x-axis is parallel and z-axis perpendicular to the mean flow. In addition, for determining the vertical advection and flux uncertainties (described in detail later in the text), we used the planar fit method for coordinate rotation. In this

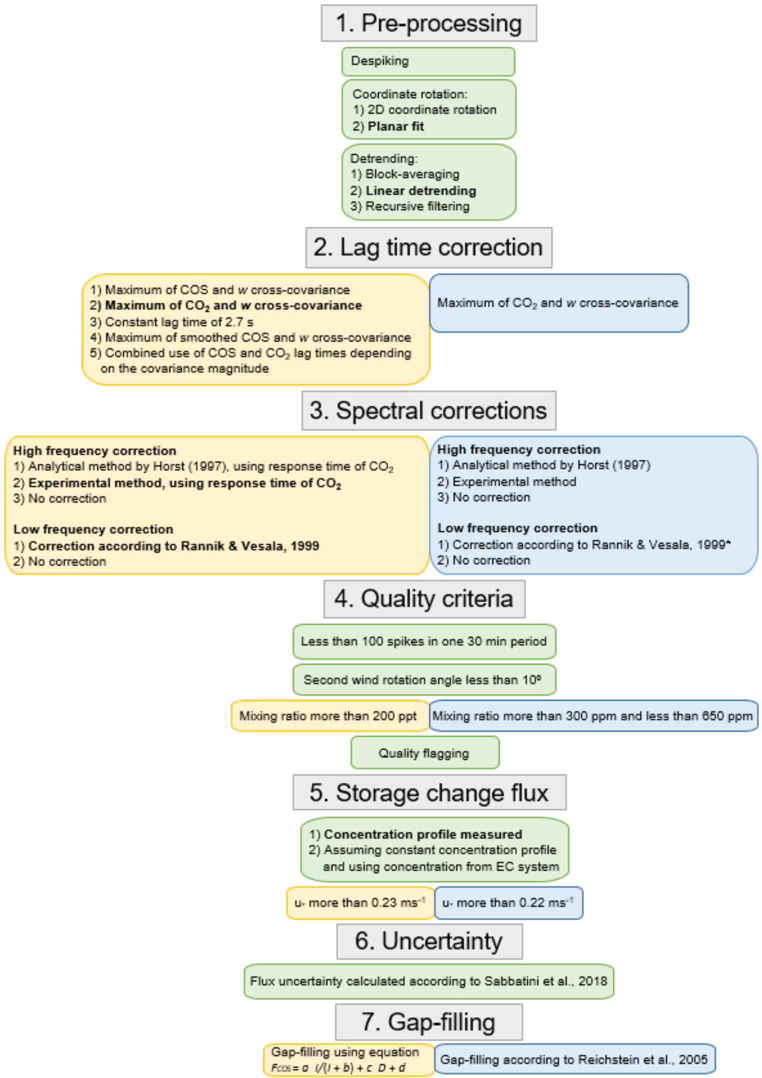


Figure 1. Different EC processing steps summarized. Yellow boxes refer to steps only used for COS data processing, blue boxes to steps used only to CO_2 data and green boxes to steps that are relevant for both gases. Recommended options are ~~marked with a star~~written in bold. ~~Standard options~~ Options that are used ~~as a~~ in the reference processing scheme for COS in this study are: ~~2D~~ planar fit coordinate rotation, linear detrending, ~~combined use of COS and~~ CO_2 time lag~~times~~, experimental ~~high frequency~~ high-frequency correction, ~~low frequency~~ low-frequency correction according to Rannik and Vesala (1999) and storage change flux from measured concentration profile.~~Advection correction (that was ignored in this study) would follow storage change flux calculation.~~

method, $\bar{w}$ is assumed to be zero only on longer time scales (weeks or even longer). A mean streamline plane is fitted to a long set of wind measurements. Then the z-axis is fixed as perpendicular to the plane and $\bar{v}$ wind component to be zero (Wilczak et al., 2001). In addition, we used the 2D coordinate rotation for coordinate rotation in two processing schemes to determine the flux uncertainty that is related to processing (Sect. 2.4.6). First, the average u-component was forced to be along the prevailing wind direction. The second rotation was performed to force the mean vertical wind speed ($\bar{w}$) to be zero (Kaimal and Finnigan, 1994). In this way, the x-axis is parallel and z-axis perpendicular to the mean flow. While 2D coordinate rotation is the most commonly used rotation method, the planar fit method (even though more demanding by computation) brings benefits especially in complex terrain (Lee and Finnigan, 2004) and is nowadays recommended as the preferred coordinate rotation method (Sabbatini et al., 2018).

To separate the mixing ratio time series into mean and fluctuating parts, we tested three different detrending options: 1) 30 min block averaging (BA), 2) linear detrending (LD) and 3) recursive filtering (RF) with a time constant of 30 s. BA is the most commonly used method for averaging the data with the benefit of dampening the turbulent signal the least. On the other hand, it BA may lead to an overestimation of the fluctuating part (and thus overestimating the flux), especially for example due to instrumental drift or large scale changes in atmospheric conditions, that are not related to turbulent transfer (Aubinet et al., 2012). The LD method fits a linear regression to the averaging period and thus gets rid of instrumental drift and to some extent of weather changes, but may lead to underestimation of the flux if the linear trend was related to actual turbulent fluctuations in the atmosphere. The third method, RF, uses a time window (here 30 s) for a moving average over the whole averaging period. RF brings the biggest correction and thus lowest flux estimate compared to other methods, but effectively removes biased low-frequency contributions to the flux. In the ICOS protocol for small fluxes, it is recommended to use BA for estimating the background signal (Nemitz et al., 2018)An Allan variance was determined for a time period when the instrument was sampling from a gas cylinder (Werle, 2010). The time constant of 30 s for RF was determined from the Allan plot (Fig. S4), as the system starts to drift in non-linear fashion at 30 s, following the approach suggested by Mammarella et al. (2010). The effect of different detrending methods will be shown is shown and discussed in Sect. 3.

2.4.2 Lag time Time lag determination

The lag time time lag between w and COS signals was determined using the following five different methods:

1) From the maximum difference of the cross-covariance of the COS mixing ratio and w ($\overline{w'\chi'_{\mathrm{COS}}}$) within lag window 1.5–3.8 s to a line between covariance values at the lag window limits (referred hereafter as COSlag). $\mathrm{COS}_{\mathrm{lag}}$). This applies also to other covariance methods explained below, and prevents the time lag to be exactly at the lag window limits. Lag window limits (form 1.5 s to 3.8 s) were determined based on the nominal lag time time lag of 2.6 s calculated from the flow rate and tube dimensions. More flexibility was given to the longer upper end of the lag window as lag times time lags have been found to be longer than the nominal lag time time lag (Massman, 2000; Gerdel et al., 2017).

2) From the maximum difference of the cross-covariance of the CO_2 mixing ratio and w ($\overline{w'\chi'_{\mathrm{CO_2}}}$) within to a line between covariance values at the lag window limits within the lag window 1.5–3.8 s (referred hereafter as CO_2 lag$\mathrm{CO2}_{\mathrm{lag}}$).

3) Using a constant lag time of 2.7 time lag of 2.6 s, which is was the nominal time lag and the most common lag for CO_2 with

our setup (referred hereafter as Const~~lag~~$_{\text{lag}}$).

4) From the maximum ~~of~~ difference of the smoothed $\overline{w'\chi'_{\text{COS}}}$ ~~when the~~ to a line between covariance values at the lag window limits. The cross-covariance was smoothed with a 1.1 s running mean (referred hereafter as RM~~lag~~$_{\text{lag}}$). The averaging window was chosen so that it provided a more distinguishable covariance maximum while still preventing a shift in the timing of the maximum.

5) A combination of COS~~and CO_2 lag times~~$_{\text{lag}}$ and CO2$_{\text{lag}}$. First, the random flux error due to instrument noise was calculated according to Mauder et al. (2013):

$$RE = \sqrt{\frac{(\sigma_c^{noise})^2 \sigma_w^2}{N}} \tag{1}$$

where instrumental noise σ_c^{noise} was estimated from the method proposed by Lenschow et al. (2000), σ_w is the standard deviation of the vertical wind speed and N the number of data points in the averaging period. The random error was then compared to the raw maximum covariance. If the maximum covariance was higher than three times the random flux error, then the COS~~lag~~$_{\text{lag}}$ method was used for ~~lag time~~ time lag determination. If the covariance was dominated by noise (the ~~random error~~ covariance being smaller than three times the random error) ~~then the CO_2~~ or COS$_{\text{lag}}$ was at the lag window limit, then the CO2$_{\text{lag}}$ lag method was selected, as proposed in (Nemitz et al., 2018) (referred hereafter as the DetLim~~lag~~$_{\text{lag}}$).

2.4.3 Frequency response correction

Some of the turbulence signal is lost both at high and low frequencies due to losses in sampling lines, inadequate frequency response of the instrument and inadequate averaging times, among ~~others~~ other reasons (Aubinet et al., 2012). In this section, we describe both high and low frequency loss corrections in detail. We tested two high-frequency correction methods, described below, simultaneously correcting for low frequency losses. One run was performed without both low-frequency and high-frequency response corrections.

High-frequency correction

Especially in closed-path systems ~~cause the fluctuations to~~ the high frequency turbulent fluctuations of the target gas dampen at high frequencies due to long sampling lines. Other reasons for high-frequency losses include sensor separation and inadequate frequency response of the sensor. In turn, high-frequency losses cause the normalized co-spectrum of the gas with w to be lower than expected at high frequencies, resulting in lower flux. The flux attenuation factor (FA) for a scalar s is defined as

$$FA = \frac{\overline{w's'}_{meas}}{\overline{w's'}} = \frac{\int_0^\infty T_{ws}(f)Co_{ws}(f)df}{\int_0^\infty Co_{ws}(f)df} \tag{2}$$

where $\overline{w's'_{meas}}$ and $\overline{w's'}$ are the measured and unattenuated covariances, respectively, $T_{ws}(f)$ is the net transfer function, specific to the EC system and scalar s, and $Co_{ws}(f)$ is the cospectral density of the scalar flux $\overline{w's'}$. For solving FA, a cospectral model and transfer function are needed. In this study, we tested the effect of high-frequency spectral correction by applying either an analytical correction for high-frequency losses ~~(Moore, 1986; Horst, 1997)~~ (Horst, 1997) or an experimental correction (Aubinet et al., 2000). The analytical correction was based on scalar co-spectra defined in Horst (1997) and the experimental approach was based on the assumption that temperature co-spectrum is measured without significant error and the normalized scalar co-spectra were compared to the normalized temperature co-spectrum ~~(Aubinet et al., 2000; Mammarella et al., 2009)~~ (Aubinet et al., 2000; Wohlfahrt et al., 2005; Mammarella et al., 2009).

~~Analytical high frequency attenuation for the scalar fluxes were determined by~~ In the analytical approach by Horst (1997) the integral in Eq. 2 is solved analytically by assuming a model cospectrum of the form $\frac{fCo_{ws}(f)}{\overline{w's'}} = \frac{2}{\pi}\frac{f/f_m}{1+(f/f_m)^2}$ and a transfer function ~~by Horst (1997):~~ $T_{ws}(f) = \frac{1}{1+(2\pi f \tau_s)^2}$. The flux attenuation then results to

$$\sout{H_{ws}(n)} FA_H = [1+(2\pi\tau_s f_m)^{\alpha}]^{-1} \tag{3}$$

where α = 7/8 for neutral and unstable stratification and α = 1 for stable stratification in the surface layer~~.~~, τ_s is the ~~sensor specific time constant (0.68 s in our case for the Aerodyne QCLS and 0.032 s for LI-6262)~~ overall EC system time constant and f_m is the frequency of the logarithmic co-spectrum peak estimated from ~~$f = \frac{n_m \bar{u}}{z-d}$~~ $f_m = \frac{n_m \bar{u}}{z_m - d}$, where n_m is the normalized frequency of the co-spectral peak, $\bar{u}$ the mean wind speed, ~~z~~ z_m the measurement height and d the displacement height. The normalized frequency of the co-spectral peak n_m is dependent on ~~stability $\zeta = \frac{z-d}{L}$ (Horst, 1997):~~ atmospheric stability $\zeta = \frac{z_m - d}{L}$ (Horst, 1997):

$$n_m = 0.085, \text{ for } \zeta \leq 0 \tag{4}$$

$$n_m = 2.0 - 1.915/(1+0.5\sout{z/L}\zeta), \text{ for } \zeta > 0 \tag{5}$$

where ~~z is the measurement height and~~ L is the ~~Monin-Obukhov length.~~ Obukhov length. The model cospectrum for the analytical high-frequency spectral correction was given as

$$\frac{fCo_{w\theta}(n)}{\overline{w'\theta'}} = \begin{cases} \frac{1.05n/n_m}{(1+1.33n/n_m)^{7/4}} & \text{for } \zeta \leq 0,\ n < 1 \\ \frac{0.387n/n_m}{(1+0.38n/n_m)^{7/3}} & \text{for } \zeta \leq 0\ n \geq 1 \\ \frac{0.637n/nm}{(1+0.91n/n_m)^{2.1}} & \text{for } \zeta > 0 \end{cases} \tag{6}$$

~~Empirical estimation of co-spectral transfer functions and flux attenuation was done according to Mammarella et al. (2009) using a site-specific co-spectral model , as recommended in De Ligne et al. (2010). The transfer function, which describes the attenuation, was defined as~~ In the experimental approach, we solved the Eq. 2 numerically as

$$\frac{fCo_{w\theta}(n)}{\overline{w'\theta'}} = \begin{cases} \frac{10.36n}{(1+4.82n)^{3.05}} & \text{for } \zeta \leq 0,\ n < 1 \\ \frac{1.85n}{(1+3.80n)^{7/3}} & \text{for } \zeta \leq 0,\ n \geq 1 \\ \frac{0.094n/n_m}{(1+9.67n/n_m)^{1.74}} & \text{for } \zeta > 0 \end{cases} \tag{7}$$

where the stability-dependency of the cospectral peak frequency n_m followed the equation

$$n_m = 0.0956, \text{ for } \zeta \leq 0 \quad (8)$$

$$n_m = 0.0956(1 + 2.4163\zeta^{0.7033}), \text{ for } \zeta > 0 \quad (9)$$

In both approaches (analytical and experimental), the time constant τ_s was empirically estimated by fitting the transfer function $T_{ws}(f)$ to the normalized ratio of ~~the cospectral densities~~ cospectral densities

$$T_{ws} = \frac{N_\theta C_{ws}(f)}{N_s C_{w\theta}(f)} \frac{N_\theta Co_{ws}(f)}{N_s Co_{w\theta}(f)} \quad (10)$$

where N_θ and N_s are normalization factors and ~~C_{ws} and $C_{w\theta}$~~ Co_{ws} and $Co_{w\theta}$ the scalar and temperature cospectra, respectively. ~~A sigmoidal function $\frac{1}{1+(2\pi f \tau_s)^2}$ was fitted to the transfer function and τ_s determined as the parameter of the least squares regression~~The estimated time constant was 0.68 s for the Aerodyne QCLS and 0.35 s for LI-6262.

Low-frequency correction

Detrending the turbulent time series, especially with LD or RF methods, may also remove part of the real low frequency variations in the data (Lenschow et al., 1994; Kristensen, 1998), which should be corrected for in order to avoid flux underestimation. Low frequency correction in this study for different detrending methods was done according to Rannik and Vesala (1999). ~~One run was performed without both low frequency and high frequency response corrections.~~

2.4.4 Flux quality criteria

The calculated fluxes ~~pass the quality criteria when : the~~ were accepted when following criteria were met: the second wind rotation angle (θ) ~~is~~ was below 10 °, the number of spikes in one half hour ~~is~~ was less than 100, the COS mixing ratio ~~is~~ was higher than 200 ppt, the CO_2 mixing ratio ranges between 300 ppm and 650 ppm and the H_2O mixing ratio ~~is~~ was higher than 1 ppb.

~~For quality flagging , the standard criteria by Sabbatini et al. (2018) were used both~~ We used a similar flagging system for micrometeorological quality criteria as Mauder and Foken (2006) for COS and CO_2: flag 0 was given if flux stationarity was less than 0.3 (meaning that covariances calculated over 5 min intervals deviate less than 30 % from the 30 min covariance), kurtosis was between 1 and 8 and skewness was within a range from -2 to 2. Flag 1 was given if flux stationarity was from 0.3 to 1 and kurtosis and skewness were within the ranges given earlier. Flag 2 was given if these criteria were not met.

In addition to these filtering and flagging criteria, we added friction velocity (u_*) filtering to screen out ~~low turbulence. However, it has to be kept in mind that the assumptions under which the u_* filtering are applied (that fluxes do not go down under low turbulenceconditions)~~ data collected under low turbulence conditions. A drop in measured EC flux is usually observed under low turbulence conditions, although it is assumed that gas exchange should not decrease due to low turbulence. While this assumption holds for CO_2, it may not be justified for COS (Kooijmans et al., 2017), as will be further discussed in Sect. 3.5. The appropriate u_* threshold was derived from a 99 % threshold criterion (Papale et al., 2006; Reichstein et al., 2005). The lowest acceptable u_* value was determined from both COS and CO_2 nighttime fluxes.

2.4.5 Storage change flux calculation

Storage change fluxes were calculated from concentration gas mixing ratio profile measurements and from EC system concentration measurements by assuming a constant concentration profile throughout the canopy using EC system mixing ratio measurements. Storage change fluxes from concentration mixing ratio profile measurements were calculated using the formula

$$F_{stor} = \frac{p}{RT_a} \int_0^{h\, z_m} \frac{\partial \chi_c(z,t)}{\partial t} dz, \tag{11}$$

where p is the atmospheric pressure, T_a air temperature, R the universal gas constant , h the EC measurement height (23 m) and $\chi_c(z)$ the gas mixing ratio at each measurement height. The integral was determined from hourly measured profile concentrations χ_c profile at 0.5 m, 4 m, 14 m, and 23 m by integrating an exponential fit through the data (Kooijmans et al., 2017). When the profile measurement was not available, storage was calculated from COS (or CO_2) concentration mixing ratio measured by the EC setup.

Another storage change flux calculation was done assuming a constant profile from the EC measurement height (23 m) to the ground level. A running average over a 5 hour window was applied to the COS concentration mixing ratio data to reduce the random noise of COS concentration measurements.

The storage change fluxes were used to correct the EC fluxes for storage change of COS and CO_2 below the flux measurement height.

2.4.6 Flux uncertainty

The flux uncertainty was calculated according to ICOS recommendations presented by Sabbatini et al. (2018). First, flux random error was estimated as the variance of covariance, according to Finkelstein and Sims (2001):

$$\epsilon_{rand} = \frac{1}{N}\left(\sum_{j=-m}^{m} \hat{\gamma}_{w,w}(j)\hat{\gamma}_{c,c}(j) + \sum_{j=-m}^{m} \hat{\gamma}_{w,c}(j)\hat{\gamma}_{c,w}(j)\right) \tag{12}$$

where N is the number of datapoints (18 000 for 30 min of EC measurements at 10 Hz), m number of samples sufficiently large to capture the integral timescale (18 000 was used in this study), $\hat{\gamma}_{w,w}$ is the variance and $\hat{\gamma}_{w,c}$ the covariance of the measured variables w and c (in this case, the vertical wind velocity and gas mixing ratio).

As the chosen processing schemes affect the resulting flux, the uncertainty related to the used processing options have to be accounted for. This uncertainty was estimated as

$$\epsilon_{proc} = \frac{max(F_{c,j}) - min(F_{c,j})}{\sqrt{12}} \tag{13}$$

where $F_{c,j}$ is the flux calculated according to j=1,...,4 different processing schemes: BA with 2D wind rotation, BA with planar fitting, LD with 2D wind rotation and LD with planar fitting. However, as all the different processing schemes will lead to slightly different random errors as well, the flux random error was estimated to be the mean of Eq. (12) for different processing schemes:

$$\overline{\epsilon_{rand}} = \frac{\sqrt{\sum_{j=1}^{4} \epsilon_{rand,j}^2}}{4} \tag{14}$$

The combined flux uncertainty is then the summation of ϵ_{rand} and ϵ_{proc} in quadrature:

$$\epsilon_{comb} = \sqrt{\overline{\epsilon_{rand}}^2 + \epsilon_{proc}^2} \tag{15}$$

To finally get the total uncertainty as the 95th percentile confidence interval, the total uncertainty becomes

$$\epsilon_{total} = 1.96\epsilon_{comb} \tag{16}$$

2.4.7 Advection

Horizontal and vertical advection are generally assumed negligible compared to EC and storage change fluxes, and advection is usually ignored as it is difficult to measure (Aubinet et al., 2010, 2012). However, advective transport may become important in low turbulence conditions, when turbulent exchange is restricted. Horizontal advection can only be determined from several EC towers placed in proximity, but this method comes with the uncertainty of natural spatial differences (Aubinet et al., 2005). Vertical advection, however, can be determined from the concentration profile as in Mammarella et al. (2007) and Lee (1998):

$$F_{VA} = \int_{0}^{z_r} \overline{w}(z) \frac{\partial \overline{c}(z)}{\partial z} dz = \overline{w}_{zr}(\overline{c}_{zr} - \langle \overline{c} \rangle)$$

where

$$\langle \overline{c} \rangle = z_r^{-1} \int_{0}^{z_r} \overline{c}(z) dz$$

where $\overline{w}$ is the mean vertical wind velocity, determined with the planar fitting method (Wilczak et al., 2001).

2.4.7 Flux gap-filling

Missing values of CO_2 fluxes were gap-filled according to Reichstein et al. (2005) while missing COS fluxes were replaced by simple model estimates or by mean hourly fluxes hourly mean fluxes if model estimates were not available, in a way comparable to gap-filling of CO_2 fluxes (Reichstein et al., 2005).

The COS gap-filling function was parameterised in a moving time window of 15 days to capture the seasonality of the fluxes. To calculate gap-filled fluxes, the parameters were interpolated daily. Gaps where any driving variable of the regression model was missing were filled with the mean hourly flux during the 15-day period.

We tested different combinations of linear or saturating (rectangular hyperbola) functions of the COS flux on photosynthetic photon flux density (PPFD) PPFD and linear functions of the COS flux against VPD or RHvapor pressure deficit (VPD) or relative humidity (RH). The saturating light response function with fixed offset (=the mean nighttime flux) as a fixed offset term explained the short-term variability of COS flux relatively well but the residuals as a function of temperature, RH and VPD were clearly systematic. Therefore, for the final gap-filling, we used a combination of saturating function on PPFD and linear function on VPD that showed good agreement with the measured fluxes while having a relatively small number of parameters:

$$F_{COS} = a * I/(I+b) + c * D + d \tag{17}$$

where I is PPFD (μmol m^{-2}s^{-1}), D is VPD and a, b, c(kPa) and a (pmol m^{-2}s^{-1}), b (μmol m^{-2}s^{-1}), c (pmol m^{-2} s^{-1}kPa^{-1}), and d (pmol m^{-2}s^{-1}) are fitting parameters.

3 Results and Discussion

Here, the various processing schemes (see Fig. 1)are compared to a "standard scheme", consisting of linear detrending,

3.1 Detrending methods

In order to check the contribution of different detrending methods on the resulting flux, we made flux calculations with different methods: block averaging (BA), linear detrending (LD) and recursive filtering (RF) using the same time lag (CO_2 time lag) for all runs (Fig. S5).

The largest median COS flux was obtained from RF (-12.0 pmol m^{-2}s^{-1}) while the smallest median flux resulted from BA (-10.7 pmol m^{-2}s^{-1}) and LD (-11.3 pmol m^{-2}s^{-1}) differed by 5.3 % from BA (Table 2). The range of the 30 min COS flux was the largest (from -183.2 to 82.56 pmol m^{-2}s^{-1}) when using the BA detrending method, consistent with a similar comparison in Gerdel et al. (2017). In comparsion, it was from -107.3 to 73.1 pmol m^{-2}s^{-1} for LD and from -164.9 to 36.8 pmol m^{-2}s^{-1} for RF. While it was surprising that RF resulted in a more negative median flux than BA, it is likely explained by the large variation in BA with compensating high negative and positive flux values, as the positive flux values with BA and LD are higher than with RF. In addition, it has to be kept in mind that the fluxes were corrected for low frequncy losses, which in part also smooths the effects of detrending.

For CO_2lag time, experimental high frequency correction and 2D wind rotation in the flux calculation. All results are presented in Finnish winter time (UTC +2) and nighttime is defined as periods when PAR < 3 , the largest median flux resulted from BA (-0.62 μmol m^{-2}s^{-1}). The smallest median flux resulted from LD (-0.56 μmol m^{-2}s^{-1}) and RF (-0.59 μmol m^{-2}s^{-1}). The difference between median CO_2 flux with BA and LD was 10.7 %.

3.2 Effect of lag time correction

The most commonly recommended averaging methods are BA (Sabbatini et al., 2018; Nemitz et al., 2018; Moncrieff et al., 2004) and LD (Rannik and Vesala, 1999) because they have less impact on spectra (Rannik and Vesala, 1999) and require less low-frequency corrections. These are also the most used detrending methods in previous COS EC flux studies (Table 1). RF may underestimate the flux (Aubinet et al., 2012) but as spectroscopic analyzers are prone to e.g. fringe effects under field conditions, the use of RF might still be justified (Mammarella et al., 2010). Regular checks of raw data provides information on instrumental drift and helps to determine the optimal detrending method. It is also recommended to check the contribution of each detrending method on the final flux to better understand what is the low frequency contribution in each measurement site and setup.

Different lag time methods resulted in slightly different lag times and COS fluxes. As COS measurements are often characterized by random noise, the maximization of the absolute value of cross-covariance may get stuck at local maxima, as demonstrated in Fig.??

Table 2. Median COS and CO_2 fluxes (in pmol $m^{-2}s^{-1}$ and μmol $m^{-2}s^{-1}$, respectively) during 26 June to 2 November 2015 and their difference to the reference fluxes when using different processing options. Median reference fluxes are -11.3 pmol $m^{-2}s^{-1}$ and -0.56 μmol $m^{-2}s^{-1}$ for COS and CO_2, respectively, and median daytime (night-time) fluxes -16.8 (-4.1) pmol $m^{-2}s^{-1}$ and -4.58 (1.23) μmol $m^{-2}s^{-1}$ when using linear detrending, CO_2 time lag and experimental high frequency spectral correction.

	Detrending		Time lag				Spectral corrections	
	BA	RF	COS lag	Const lag	RM lag	DetLim lag	Horst (1997)	None
Median F_{COS}	-10.7	-12.0	-11.5	-11.1	-11.1	-13.1	-11.0	-9.7
Difference to reference	5.3 %	6.2 %	1.8 %	1.8 %	1.8 %	15.9 %	2.7 %	14.2 %
Daytime median F_{COS}	-16.0	-17.6	-17.4	-16.6	-16.6	-19.2	-16.9	-15.0
Night-time median F_{COS}	-4.2	-4.1	-4.4	-4.1	-4.3	-4.9	-4.0	-3.4
Median F_{CO2}	-0.62	-0.59	NA	NA	NA	NA	-0.54	-0.48
Difference to reference	10.7 %	5.4 %	NA	NA	NA	NA	3.6 %	14.3 %
Daytime median F_{CO2}	-4.49	-5.00	NA	NA	NA	NA	-4.77	-4.18
Night-time median F_{CO2}	1.17	1.26	NA	NA	NA	NA	1.18	1.02

3.2 Effect of time lag correction

Different time lag methods resulted in slightly different time lags and COS fluxes. The most common lag times were : 3.6 time lags were 2.6 s from the COSlag method, 2.7 s from $_{lag}$, CO2lag, 3.8 $_{lag}$ and DetLim$_{lag}$ methods, and 1.5 s from RMlag ($_{lag}$, which was the lag window limit), and 2.7 s from the the DetLim lag (Fig. 2). Lag times Time lags determined from the COSand RM lag methods were mostly at the upper limit of $_{lag}$ and RM$_{lag}$ methods were distributed evenly throughout the lag window(3.8 s), whereas the lag from the CO_2 and DetLimmethods reached the global maxima around CO2$_{lag}$ and DetLim$_{lag}$ methods were normally distributed with most lags detected at the window center.

The lag time determined directly from $\overline{w'\chi'_{COS}}$ We tested determining time lags also from the most commonly used method of maximizing the absolute covariance. If the time lag was determined from the absolute covariance maximum instead of the maximum difference to a line between covariance values at the lag window limits, the COS$_{lag}$ and RM$_{lag}$ had most values at the lag window limits (Fig. S2). This resulted in a “mirroring effect” (Langford et al., 2015), i.e., fluxes close to zero were not detected as often as with other methods, since the covariance is always maximized and the derived flux switches between positive and negative values of similar magnitude (Fig. 3). The effect was not as strong with smoothed $\overline{w'\chi'_{COS}}$ but was still visible. Other S3). This issue should be taken into account in COS EC flux processing, as absolute COS covariance maximum is by far the most commonly used method in determining the time lag in COS studies (Table 1). To make the different methods more comparable, the time lags were, in the end, determined from the maximum difference to a line between covariance values at the lag window limits in this study. In this way, time lags were not determined at the window borders and most of the methods had the final flux PDF peak approximately at the same flux values but and had otherwise small differences in the distributions (Fig. 3). A constant lag time The only method that was clearly different from the others was DetLim$_{lag}$, that produced higher fluxes than other methods.

A constant time lag has been found to bias the flux calculation as the lag time likely varies time lag likely can vary over time due to, for example, fluctuations in pumping speed (Langford et al., 2015; Taipale et al., 2010; Massman, 2000). However, as the CO2$_{lag}$ was often the same as the chosen constant lag 2.6 s, we did not observe major differences between these two methods. A reduced bias in the flux calculation with smoothed cross-covariance was introduced by Taipale et al. (2010), who recommended using this method for any EC system with a low signal-to-noise ratio. However, we do not recommend this as a first choice since we still the time lags do not have a clear distribution and if maximum covariance method is used, we find a mirroring effect with the RMlag $_{lag}$ in the final flux distributions. If other supporting measurements, like CO_2 in the case of COS measurements, are not available to help the lag time determination, then RM lag would be a more suitable method than the maximum cross-covariance method when determining lag times for fluxes with low signal-to-noise ratio.

About 40 % of the lag times were COSlag times with the DetLim lag methodBy using the DetLim$_{lag}$ method, the COS time lag was estimated for 54 % of cases from COS$_{lag}$, while the CO2$_{lag}$ was used as proxy for the COS time lag in about 46 % of cases. Fig. 3 shows that the raw covariance of COS only exceeds the noise level at higher COS flux values and thus the COSlag time $_{lag}$ is chosen by this method only at higher fluxes, as expected. At lower flux rates, and especially close to zero, the COS fluxes are not high enough to surpass the noise level and thus the CO_2 lag time CO2$_{lag}$ is chosen.

The cumulative COS flux (summarized in Table 2) was median COS fluxes were highest when the lag time time lag was determined from the detection limit method (-74.3 nmol DetLim$_{lag}$ (-13.1 pmol $m^{-2}s^{-1}$) and maximum $\overline{w'\chi'_{COS}}$ (-72.1 nmol COS$_{lag}$ (-11.5 pmol $m^{-2}s^{-1}$) methods, respectively (Table 2). Using the COS$_{lag}$ results in both higher positive and negative fluxes and might thus have some compensating effect on the median fluxes. Using just the COSlag time biases against small fluxes, as it is always maximizing the covariance, but it also produces more positive fluxes compared to the DetLim method. This results in slightly smaller cumulative fluxes(smaller negative values). Maximizing the smoothed $\overline{w'\chi'_{COS}}$ produced the third largest cumulative uptake (-69.7 nmol Const$_{lag}$ and RM$_{lag}$ produced the smallest median uptake (-11.1 pmol $m^{-2}s^{-1}$) while the CO_2 lag time produced a slightly higher cumulative flux (-65.9 nmol $m^{-2}s^{-1}$) than using a constant lag time of 2.7 s (-65.1 nmol CO2$_{lag}$ had a median flux of -11.3 pmol $m^{-2}s^{-1}$). The difference between each of the cumulative sums was 12 % or less. This difference is large in the annual scale, and we recommend using the detection limit method in lag time determination of small fluxes, as in Nemitz et al. (2018).

Lag time determined from different methods on 15 July 2015 at 12:00–12:30 for COS (a) and CO_2 (b).

Distribution of lag times derived from different methods: COS lag (a), CO_2 lag (b), COS lag from a running mean cross-covariance (c) and combination of COS and CO_2 lag times (d), see Sec. 2.4.2.

Normalized COS flux distributions using different lag time methods: COS lag (a), CO_2 lag (b), constant lag time of 2.7 s (c), COS lag from a running mean cross-covariance (d) and combination of COS and CO_2 lag times (e) (see Sec. 2.4.2), and a summary of all PDFs (f).

Cumulative sums of COS and CO_2 fluxes (in nmol $m^{-2}s^{-1}$ and mmol $m^{-2}s^{-1}$, respectively) and their difference to the reference fluxes when using different processing options. Reference fluxes are indicated in bold.

3.3 Detrending methods

In order to check the contribution of different detrending methods on the resulting flux, we made flux calculations with different methods: block averaging (BA), linear detrending (LD) and recursive filtering (RF) using the same lag time (CO_2 lag) for all runs (Fig. ??). An Allan variance was determined for a time period when the instrument was sampling from a gas cylinder (Werle, 2010). The time constant of 30 s for recursive filtering was determined from the Allan plot (Fig. ??), as the system starts to drift in non-linear fashion at 30 s.

The largest cumulative COS flux was obtained from BA (-66.0 nmol $m^{-2}s^{-1}$), since BA retains the fluctuations the most. The smallest cumulative sum resulted from RF (-65.2 nmol $m^{-2}s^{-1}$) while LD (-65.9 nmol $m^{-2}s^{-1}$) did not differ much from BA (Table 2) . The difference between BA and RF in the total cumulative COS flux was only 1.2 %. The variation of the COS flux was highest (from -234.2 to 154.9 the reference flux and the flux from $DetLim_{lag}$ method was clearly higher (15.9 %) than with other methods (1.8 %) and had largest variation (from -113.8 pmol $m^{-2}s^{-1}$) when using the BA detrending method, consistent with a similar comparison in Gerdel et al. (2017). Variation for other detrending methods was from -119.1 to 79.7 to 81.6 pmol $m^{-2}s^{-1}$ for LD and from -75.3 to 33.3) and standard deviation (16.1 pmol $m^{-2}s^{-1}$ for RF. While BA results in higher negative fluxes in general, it also includes higher positive fluxes, thus resulting in less variation in the cumulative sum.

For CO_2, the largest cumulative flux resulted from LD (-12.74 mmol $m^{-2}s^{-1}$) although the difference to BA (-12.71 mmol $m^{-2}s^{-1}$) was small. The smallest cumulative flux resulted from RF (-11.31 mmol $m^{-2}s^{-1}$), similar to COS. The difference between cumulative CO). This difference might become important in the annual scale, and as the most commonly used covariance maximization method does not produce a clear time lag distribution for $DetLim_{lag}$ or COS_{lag}, we recommend using the $CO2_{lag}$ for COS fluxes, as in most ecosystems the CO_2 flux with BA and LD was 11 %. The relative difference between RF and BA methods is thus smaller for COS than for CO_2, possibly because instrumental noise is aliasing part of the COS fluctuations.

The most commonly recommended averaging methods are BA (Sabbatini et al., 2018; Nemitz et al., 2018; Moncrieff et al., 2004) and LD (Rannik and Vesala, 1999) because they have less impact on spectra (Rannik and Vesala, 1999) and require less corrections in the low frequency. RF may underestimate the flux (Aubinet et al., 2012). As spectroscopic analyzers are prone to e.g. fringe effects under field conditions, the use of RF might still be justified (Mammarella et al., 2010). Regular checks of raw data provides information on instrumental drift and helps to determine the optimal detrending method. It is also recommended to

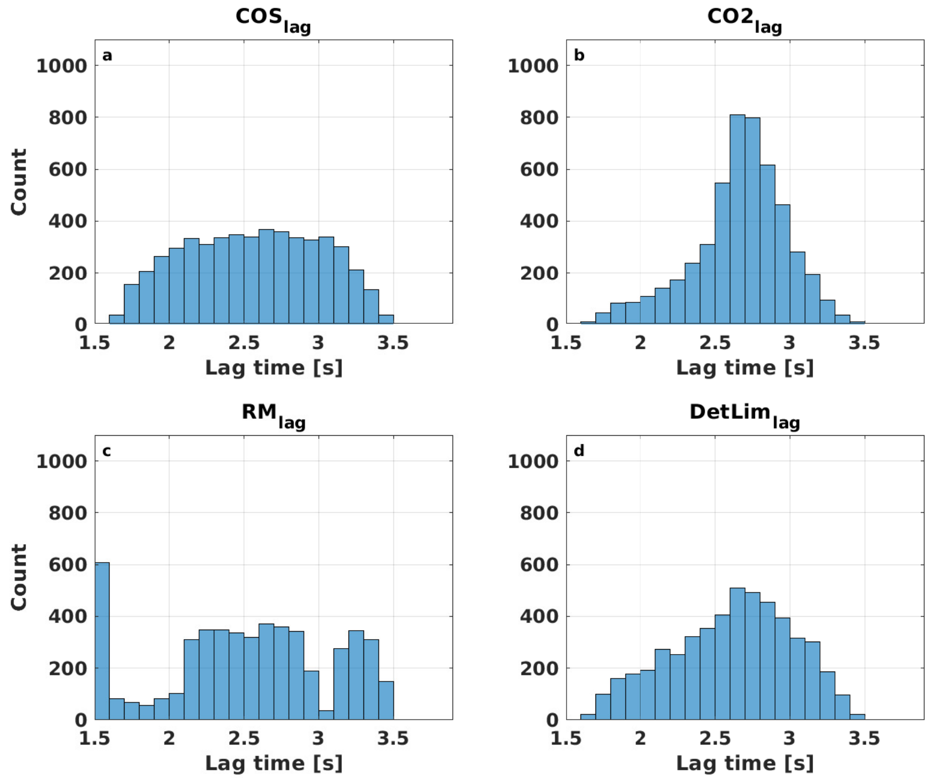


Figure 2. Comparison Distribution of time lags derived from different detrending methods: COS$_{lag}$ (recursive filtering, linear detrending and block averaginga), see Sec. 3.1CO2$_{lag}$ (b)applied to raw , COS (time lag from a running mean cross-covariance (RM$_{lag}$, c) and CO_2 combination of COS$_{lag}$ and CO2$_{lag}$ (bDetLim$_{lag}$, d)mixing ratio data on 15 July 2015 12:00–12:30..

check the contribution of each detrending method on the final flux to better understand what is the low frequency contribution in each measurement site and setup. cross-covariance with w is more clear than the cross-covariance of COS and w signals.

3.3 Frequency High-frequency response correction

3.3.1 High frequency correction

The mean COS co-spectrum cospectrum was close to the normal mean CO_2 co-spectrum cospectrum (compare Fig. 4a and 4bc). The power spectrum of COS was dominated by white noise as can be seen from the increasing power spectrum with increasing frequency for normalized frequencies greater than 0.2, which is similar to what was measured for COS by Gerdel et al. (2017) and for N_2O by Eugster et al. (2007). The fact that COS measurements are dominated by white noise at high frequencies means that those measurements are limited by the precision and that they do likely likely do not capture the

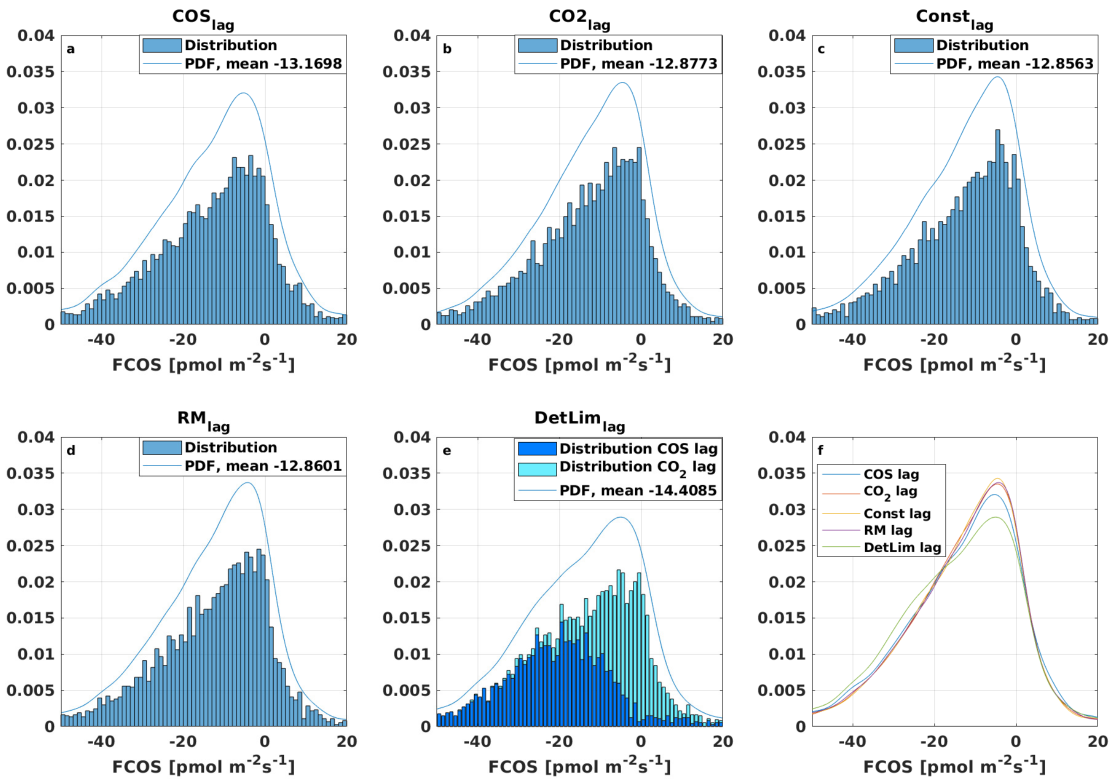


Figure 3. ~~Allan plot for~~ Normalized COS flux distributions using different time lag methods: COS$_{lag}$ (a), CO2$_{lag}$ (b), constant time lag of 2.6 s (Const$_{lag}$, c), time lag from a running mean COS cross-covariance (RM$_{lag}$, d) and combination of COS and CO_2 time lags (~~b~~DetLim$_{lag}$, e)~~mixing ratios versus averaging time τ. The dashed lines represent slopes for white noise~~, ~~linear drifting~~ and ~~non-linear drifting~~a summary of all probability distribution functions (PDFs) (f).

true variability in COS turbulence signals. This is less of a problem for CO_2, where white noise only starts to dominate at higher frequencies (normalized frequency higher than 3~~s). Co-spectral~~). Cospectral attenuation was found for both COS and CO_2 at high frequency. ~~The response time of the analyzer (0.68 s) was calculated from CO_2 measurements and the same response time applied to COS high frequency spectral correction, assuming that both fluxes are affected by the same attenuation (Wehr et al., 2017).~~

~~Both experimental and analytical approaches increased the COS and CO_2 fluxes, as expected.~~ High frequency losses due to e.g. attenuation in sampling tubes and limited sensor response times are expected to decrease fluxes if not corrected for (Aubinet et al., 2012). The ~~cumulative sum of the~~ median COS flux, when using the CO_2 ~~lag time~~ time lag and keeping low-frequency correction and quality filtering the same, was the lowest without any high-frequency correction (~~-59.4 nmol~~ -9.7

pmol $m^{-2}s^{-1}$), highest with the experimental correction (-65.9 nmol -11.3 pmol $m^{-2}s^{-1}$) and in between with the analytical correction (-61.2 nmol -11.0 pmol $m^{-2}s^{-1}$). Correcting for the high frequency attenuation thus made a maximum of 9.9 14.2 % difference in the cumulative median COS flux. In addition, daytime median fluxes increased from -15.0 to -16.9 pmol $m^{-2}s^{-1}$ with the analytical correction, and to -16.8 pmol $m^{-2}s^{-1}$ with the experimental high frequency correction. However, the relative difference was larger during night-time when fluxes increased from -3.4 to -4.0 and -4.1 pmol $m^{-2}s^{-1}$ with analytical and experimental methods, respectively.

Similar results were found for the CO_2 flux but the differences were smaller: without any high frequency correction the cumulative sum median flux was lowest at -12.18 mmol -0.48 μmol $m^{-2}s^{-1}$, highest with experimental correction (-12.74 mmol -0.56 μmol $m^{-2}s^{-1}$) and in between with the analytical correction (-12.54 mmol -0.54 μmol $m^{-2}s^{-1}$), thus making a maximum of 4.4 14.3 % difference in the cumulative median CO_2 flux, similar to COS. Very similar results were found for CH_4 and CO_2 fluxes in Mammarella et al. (2016), where the high frequency correction made the largest difference in the final flux processing after dilution and Webb corrections.

From the empirical method we get a stability-dependent function for the cospectral peak frequency n_m similar to the analytical method by Horst (1997) (Eq. 4 and 5):

$$n_m = 0.0956, \text{ for } \zeta \leq 0$$

$$n_m = 0.0956(1+2.4163\zeta^{0.7033}), \text{ for } \zeta > 0$$

This result is compared to the analytical method (Eq. 4 and 5) in Fig. ??. In contrast to Rannik et al. (2004), we find a difference between the empirical method (Eq. 18) and Horst (1997) (Eq. 5) in stable conditions (when $\zeta < 0$, Fig. ??for closed-path analysers. Similarly to COS, also CO_2 fluxes were increased more during night-time due to spectral correction than during daytime. Daytime median flux increased from -4.18 to -4.77 and -4.58 μmol $m^{-2}s^{-1}$ and night-time fluxes increased from 1.02 to 1.18 and 1.23 μmol $m^{-2}s^{-1}$ when using analytical and experimental high frequency spectral corrections, respectively. Flux attenuation was dependent on stability and wind speed for both correction methods, as also found in Mammarella et al. (2009) (Fig. S7).

Co-spectrum and power spectrum for COS (a and b, respectively) and CO_2 (c and d) in July 2015.The site specific model captures the cospectrum better than the model cospectrum by Horst (1997), as shown in Fig. 4. For high-frequency spectral corrections, it is recommended to use the site specific cospectral model, as has been done in most previous COS studies (Table 1).

3.3.1 Low frequency correction

As shown in Sect. 3.1, the difference between different detrending methods on the cumulative COS flux was negligible. From the example raw COS data (Fig. ?? a) it is also seen that the detrending methods are not very different from each other. Thus it is expected that the low frequency correction for COS does not make a large effect on the final flux, as noise is covering part of the low frequency variability in COS measurements. However, as we find larger differences in CO_2 detrending and the

~~Co-spectral peak frequency as a function of stability according to experimental method and Horst (1997).~~

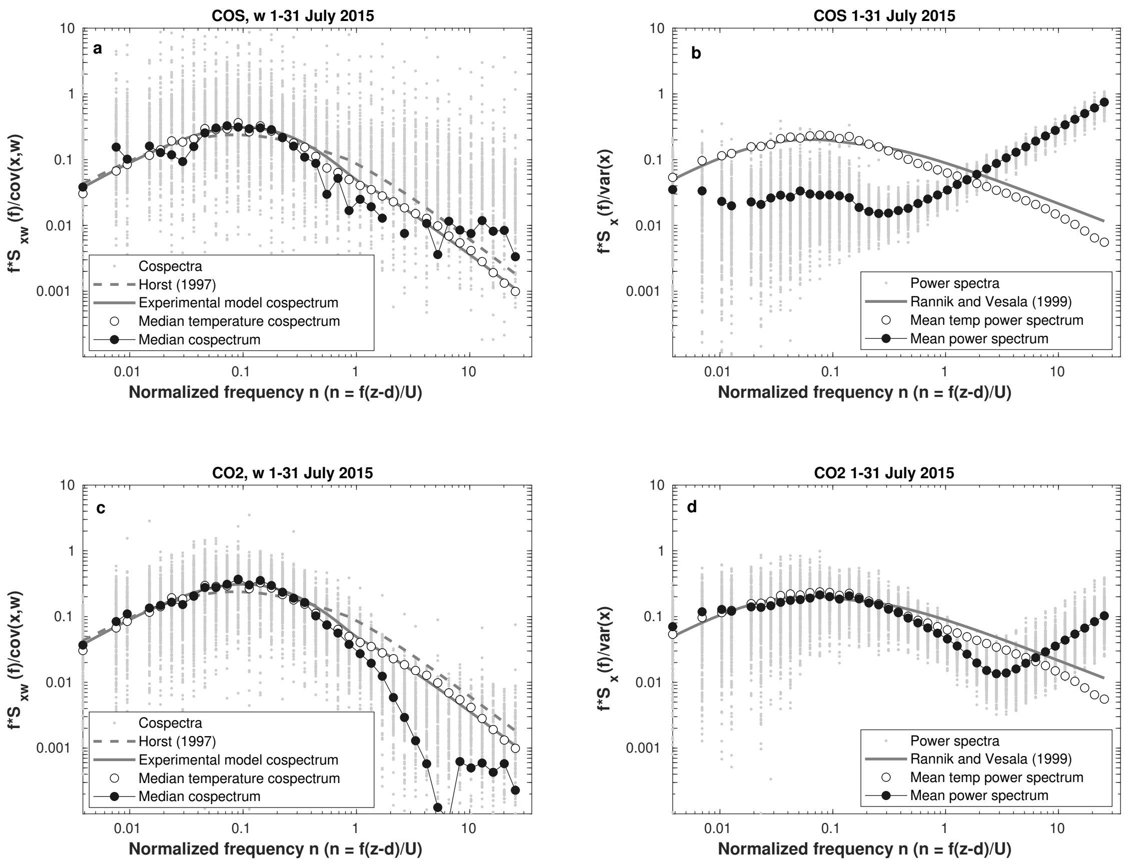


Figure 4. Cospectrum and power spectrum for COS (a and b, respectively) and CO_2 (c and d) in July 2015. All data were filtered by the stability condition $-2 < \zeta < -0.0625$ and COS data were only accepted when the covariance was higher than three times the random error due to instrument noise (Eq. 1). The cospectrum models by experimental method and Horst (1997), that were used in the high-frequency spectral correction, are shown in grey continuous and dashed lines, respectively.

~~detrending methods differ from each other more than for COS (Fig. ?? b), the low frequency correction is expected to be more relevant for CO_2 fluxes. Rannik and Vesala (1999) found a 15 % underestimation of the uncorrected fluxes for CO_2, which is also comparable to the difference between the CO_2 fluxes determined using different detrending methods in our study.~~

3.4 Storage change fluxes

In the following, storage change fluxes based on profile measurements are listed as default, with fluxes based on the constant profile assumption listed in brackets.

The COS storage change flux was negative from 15:00 in the afternoon until 06:00 in the morning with a minimum of -1.0 pmol $m^{-2}s^{-1}$ (-0.6 pmol $m^{-2}s^{-1}$) reached at 20:00 in the evening. A negative storage change flux of COS indicates that there is a COS sink in the ecosystem when the boundary layer and effective mixing layer is shallow. Neglecting this effect would lead to overestimated uptake at the ecosystem level later when the air at the EC sampling height is better mixed. The COS storage change flux was positive from around 6:00 in the morning until 15:00 in the afternoon and peaked at 9:00 with a magnitude of 1.9 pmol $m^{-2}s^{-1}$ (0.8 pmol $m^{-2}s^{-1}$). The storage change flux made the highest relative contribution to the sum of measured EC and storage change fluxes at midnight with 18 % (13 %) (Fig. 5 c). The difference between the two methods was minimum 13 % at 11:00 and maximum 56 % at 09:00. The two methods made a maximum of 7 % difference on the resulting cumulative ecosystem flux, as already reported in Kooijmans et al. (2017).

The CO_2 storage change flux was positive from 15:00 in the afternoon until around 4:00 in the morning with a maximum value of 0.62 μmol $m^{-2}s^{-1}$ (0.38 μmol $m^{-2}s^{-1}$) reached at 21:00 in the evening. A positive storage change flux indicates that the ecosystem contains a source of carbon when the boundary layer is less turbulent and accumulates the respired CO_2 within the canopy. As turbulence would increase later in the morning, the accumulated CO_2 would result in an additional flux that could mask the gas exchange processes occurring at that time step. The CO_2 storage change flux minimum was reached with both methods at 08:00 with magnitude -1.01 μmol $m^{-2}s^{-1}$ (-0.52 μmol $m^{-2}s^{-1}$) when the boundary layer has already started expanding, and leaves are assimilating CO_2. The maximum contribution of the storage change flux was as high as 89 % (36 %) compared to the EC flux at 18:00, when the CO_2 exchange is turning to respiration and storage change flux increases its relative importance (Fig. 5 d). The difference between the two storage change flux methods for CO_2 was maximum 53 % at 21:00 and minimum 13 % at midnight. The maximum difference of 5 % was found in the cumulative ecosystem CO_2 flux when the different methods were used.

In conclusion, the storage change fluxes are not relevant for budget calculations ~~- as expected - but they are important to account for the delayed capture of fluxes by the EC system under low turbulence conditions.~~

~~Diurnal variation of the storage change flux, determined from COS (a) and CO_2 (b) profile measurements (blue) and by assuming a constant profile up to 23 m height (purple). Contribution of storage change flux to the total ecosystem EC flux with the profile measurements and assuming a constant profile for COS (c) and CO_2 (d).~~

3.5 ~~Vertical advection~~

Vertical advection of COS ($F_{VA,COS}$) was negligible during daytime and negative during nighttime (Fig. ??). During the night, $F_{VA,COS}$ was more important than the – as expected – and have not been widely applied in previous COS studies (Table 1), even though storage change flux and was half the size of the total ecosystem flux in magnitude.

Vertical advection of CO_2 ($F_{VA,CO2}$) was also negligible during the day, and in contrast to COS, it was slightly positive during night. $F_{VA,CO2}$ was still less than 50 % of the total ecosystem CO_2 flux and was comparable to the storage flux , similar to Mammarella et al. (2007) and Aubinet et al. (2003). Similar findings were reported also in Aubinet et al. (2005) at slightly sloping sites, for which the importance of measurements are mandatory in places where the EC system is placed at a height of 4 m or above according to the ICOS protocol for EC flux measurements (Montagnani et al., 2018). In addition, storage change fluxes and advection varied with weather conditions. As we were not able to measure horizontal advection, we did not include vertical advection in the nighttime flux correction, to avoid an overestimated correction and an underestimated total flux (Aubinet et al., 2003). Horizontal and vertical advection were found to be large but of opposite signs in Aubinet et al. (2003) and they suggested to exclude the correction based solely on vertical advection, proposed by Lee (1998)are important in the diurnal scale to account for the delayed capture of fluxes by the EC system under low turbulence conditions.

3.5 u_* filtering

Especially during nighttime it is common that there is not enough turbulence to mix the surface layerCalm and low turbulence conditions are especially common during nights with stable atmospheric stratification. In this case, storage and advection start to play a bigger change and advective fluxes have an important role and the measured EC flux of a gas does not only reflect the atmosphere-biosphere exchange, typically underestimating the exchange. This often leads to a systematic bias in the annual flux budgets (Moncrieff et al., 1996; Aubinet et al., 2000; Aubinet, 2008). Even after studies of horizontal and vertical advection, the u_* filtering still keeps its place as the most efficient and reliable tool to filter out erroneous data data that is not representative of the surface-atmosphere exchange under low turbulence (Aubinet et al., 2010).

For COS, the nighttime filtering is a more complex issue than it is for CO_2. In contrast to CO_2, COS is taken up by the ecosystem during nighttime (Kooijmans et al., 2017, 2019), depending on stomatal conductance and the concentration gradient between the leaf and the atmosphere. When the atmospheric COS concentration goes down mixing ratio decreases under low turbulence conditions (due to nighttime COS uptake in the ecosystem), the concentration gradient between the leaf and the atmosphere goes down, such that a decrease in COS uptake can be expected (Kooijmans et al., 2017). The u_* filtering is applied to conform Thus, the assumption that fluxes do not go down under low turbulence conditions, as is the case for respiration of CO_2, but which does not does not necessarily apply to COS uptake. The u_* filtering may therefore bias COS fluxes due to false assumptions. StillHowever, as we did not see u_* dependency disappearing even with a concentration gradient-normalized flux, the u_* filtering is applied here normally to overcome the EC measurement limitations under low turbulence conditions.

We determined u* u_* limits of 0.23 ms^{-1} for COS and 0.22 ms^{-1} for CO_2 (Fig. 6). Filtering according to these u_* values would remove 12 % and 11 % of data, respectively. If the storage change flux was excluded when determining the u_* threshold, the limits were 0.39 ms^{-1} and 0.24 ms^{-1} from CO_2 and COS fluxes, respectively. The increase in the u_* threshold with CO_2 is because the fractional storage change flux is larger for CO_2 than for COS (Fig. 5, panels c and df). On the other hand, the u_*

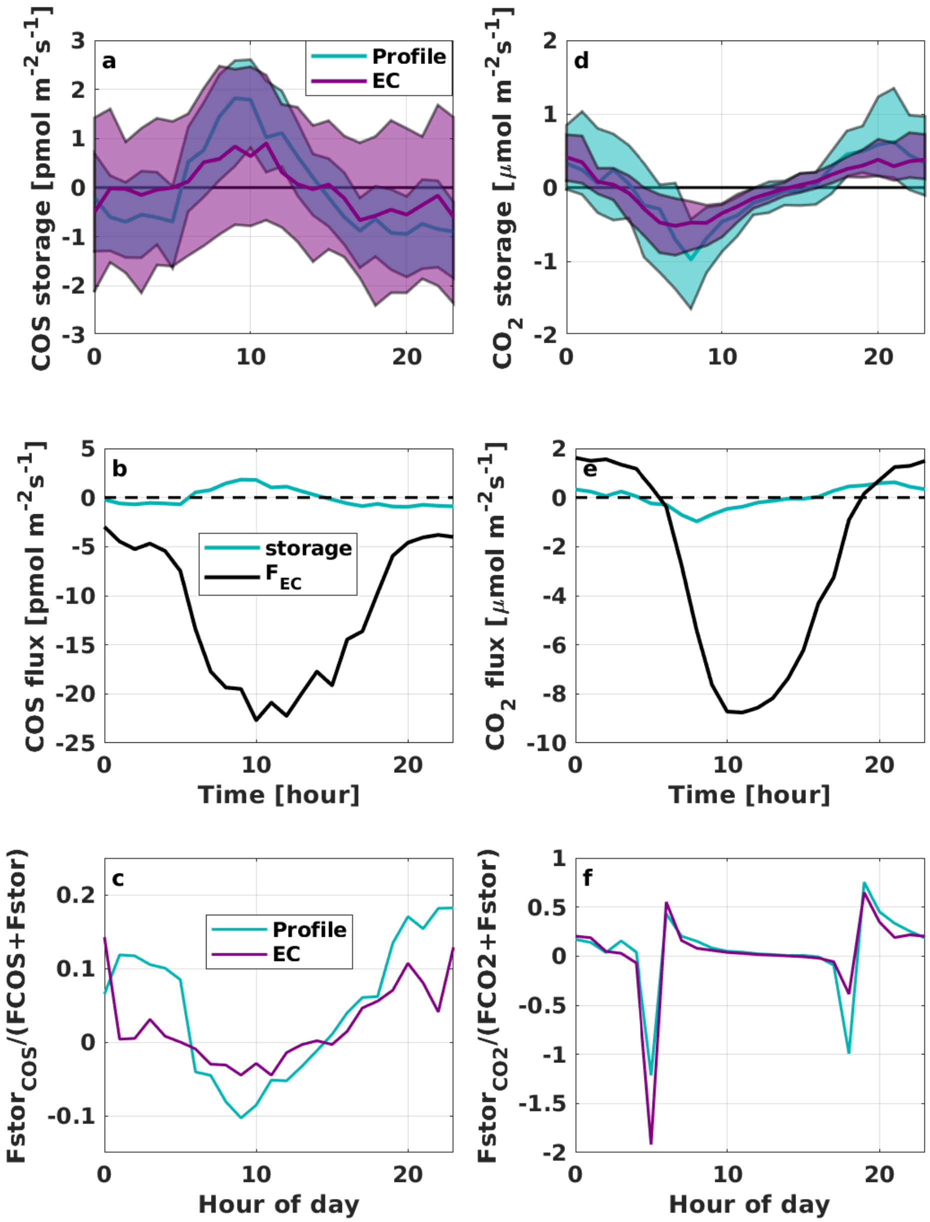


Figure 5. Diurnal variation of the storage change flux, determined from COS (a) and CO_2 (d) profile measurements (blue) and by assuming a constant profile up to 23 m height (purple), ~~vertical advection~~ and diurnal variation of the EC flux (~~yellow~~black) and ~~total ecosystem EC~~ storage change flux with the profile method (blue) (b and e for COS ~~(~~and CO_2 fluxes, respectively) during the measurement period 26 June to 2 November 2015. Contribution of storage change flux to the total ecosystem EC flux with the profile measurements and assuming a constant profile for COS (c) and CO_2 (~~b~~f).

limit for COS stayed similar to the previous one. With these u_* thresholds the filtering would exclude 30 % and 13 % of the data for COS and CO_2 respectively.

If fluxes are not corrected for storage before deriving the u_* threshold, there is a risk of flux overestimation due to double accounting. ~~Without storage correction done first, the fluxes during low turbulence in nighttime would be removed and storage flux ignored. The filtered flux would be gapfilled~~The flux data filtered for low turbulence would be gap-filled, thereby accounting for storage by the canopy, but then accounted for again when the storage is released and measured by the EC system during the flushing hours in the morning (Papale et al., 2006). Thus, it is necessary to make the storage change flux correction before deriving u_* thresholds and applying the filtering.

3.6 Gap-filling

Three combinations of environmental variables (PAR, PAR and relative humidity, PAR and VPD) were tested using the gap-filling function Eq. 17. These environmental parameters were chosen because COS exchange has been found to depend on stomatal conductance, which in turn depends especially on radiation and humidity (Kooijmans et al., 2019). Development of the gap-filling parameters a, b, c and d over the measurement period is presented in the Supplementary material Fig. S10. While saturating function of PAR only captured the diurnal variation already relatively well, adding a linear dependency on VPD or RH made the diurnal pattern even closer to the measured one (Fig. 7). Therefore, the combination of saturating light response curve and linear VPD dependency was chosen. Furthermore, we chose a linear VPD dependency instead of a linear RH dependency due to smaller residuals in the former (Fig. S9).

For COS fluxes, 44 % of daytime flux measurements were discarded due to the above-mentioned quality criteria (Sect. 2.4.4) and low-turbulence filtering. As expected, more data (66 %) were discarded during nighttime. Altogether 52 % of all COS flux data were discarded and gap-filled with the gap-filling function presented in Eq. (17). The cumulative sum of the final, corrected and gap-filled, COS fluxes during the whole measurement period totalled up to ~~-139~~ -136 μmol COS m^{-2}, while without gap-filling the cumulative sum would be ~~47~~ 43 % smaller at ~~-73.7~~ -77.2 μmol COS m^{-2}.

For CO_2, 41 % daytime CO_2 fluxes were discarded, while 67 % of fluxes were discarded during nighttime. Altogether comprising 53 % of all CO_2 flux data. CO_2 fluxes were gap-filled according to standard gap-filling procedures presented in (Reichstein et al., 2005). The cumulative NEE after all corrections and gap-filling was ~~-22.1~~ -22.73 mol CO_2 m^{-2}, while without gap-filling the cumulative sum would be only ~~-19.5~~ -19.65 mol CO_2 m^{-2}.

Although the COS community has not been interested in the cumulative COS fluxes or yearly COS budget so far, it is important to fill short-term gaps in COS flux data to properly capture e.g. the diurnal variation. The gap-filling method presented here is one option to be tested also at other measurements sites.

3.7 Errors and uncertainties

The uncertainty due to different processing schemes in the flux processing contributed 33 % to the total uncertainty of the COS flux, the rest was composed of the random flux uncertainty (Fig. 8). For the CO_2 flux uncertainty, the processing was more

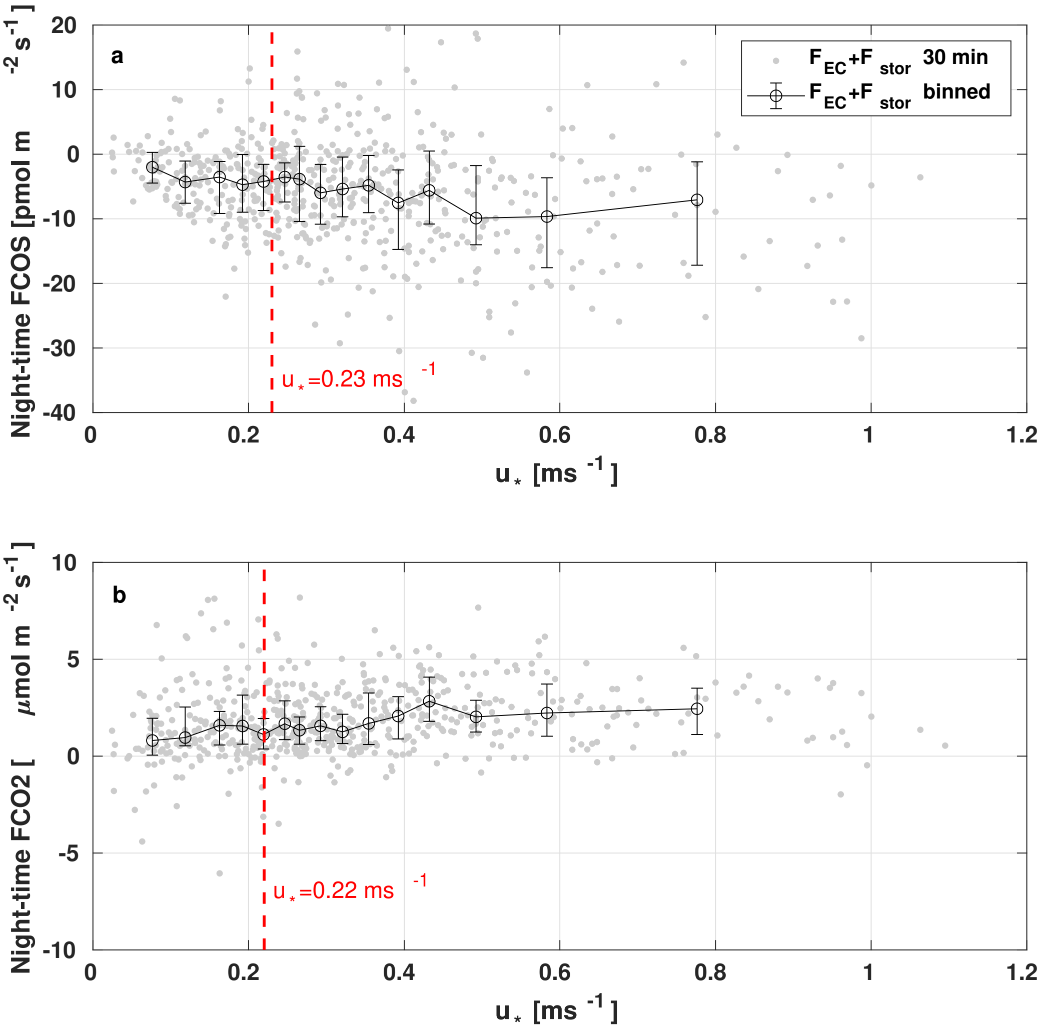


Figure 6. Nighttime median ecosystem fluxes (black) of COS (a) and CO_2 (b) ~~fluxes~~ binned ~~to~~ into 15 equal-sized friction velocity bins. Error bars indicate ranges between 25th and 75th percentiles. Friction velocity thresholds are shown with ~~the~~ red dashed lines and single data points of 30-min fluxes with light gray dots.

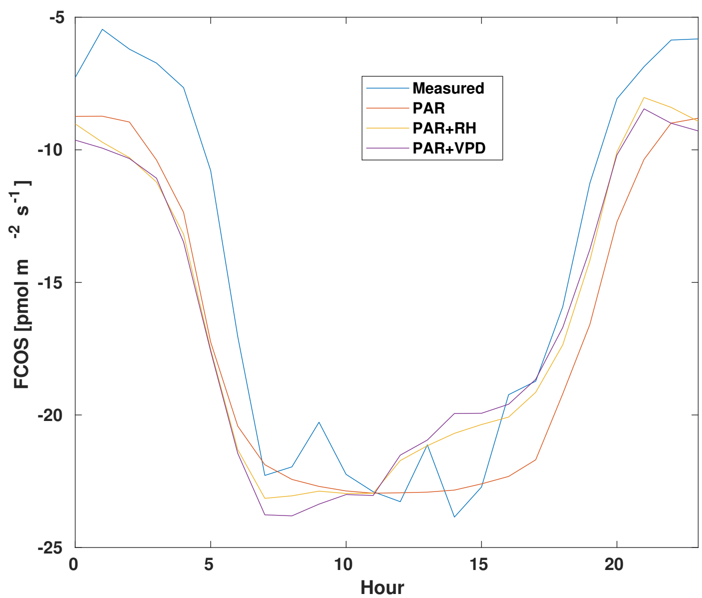


Figure 7. Diurnal variation of the measured COS flux (blue) and the flux from different gap-filling methods: gap-filling with only saturating PAR function (orange), saturating PAR and linear dependency on RH (yellow) and saturating PAR and linear dependency on VPD (purple). Diurnal variation is calculated from 1 July to 31 August 2015 for periods when measured COS flux existed.

important than for COS (40 %), but the random error was still dominating the combined flux uncertainty. The random error of the CO_2 flux was found to be lower than in Rannik et al. (2016) for the same site, probably related to differences in the gas analyzers and overall setup. The mean noise estimated from Lenschow et al. (2000) was 0.06 μmol $m^{-2}s^{-1}$ for our QCLS CO_2 fluxes while in Rannik et al. (2016) it was approximately 0.08 μmol $m^{-2}s^{-1}$ for LI-6262 CO_2 fluxes at the same site. Gerdel et al. (2017) found the total random uncertainty of COS fluxes to be mostly around 3–8 pmol $m^{-2}s^{-1}$, comparable to our results.

The relative flux uncertainty for COS is very high at low flux (-3 pmol $m^{-2}s^{-1}$ < ~~FCOS~~ F_{COS} < 3 pmol $m^{-2}s^{-1}$) values (8 times the actual flux value) but levels ~~out~~ off to 45 % at fluxes higher (meaning more negative fluxes) than -27 pmol $m^{-2}s^{-1}$ (Fig. 8c). The total uncertainty of the CO_2 flux was also high at low fluxes (-1.5 μmol $m^{-2}s^{-1}$ < ~~FCO2~~ F_{CO2} < 1 μmol $m^{-2}s^{-1}$, uncertainty reaching 120 % of the flux at 0.17 μmol $m^{-2}s^{-1}$) and decreased to 15 % at fluxes higher than -11 μmol $m^{-2}s^{-1}$ (Fig. 8d). Higher relative uncertainty at low flux levels is probably due to detection limit of the measurement system.

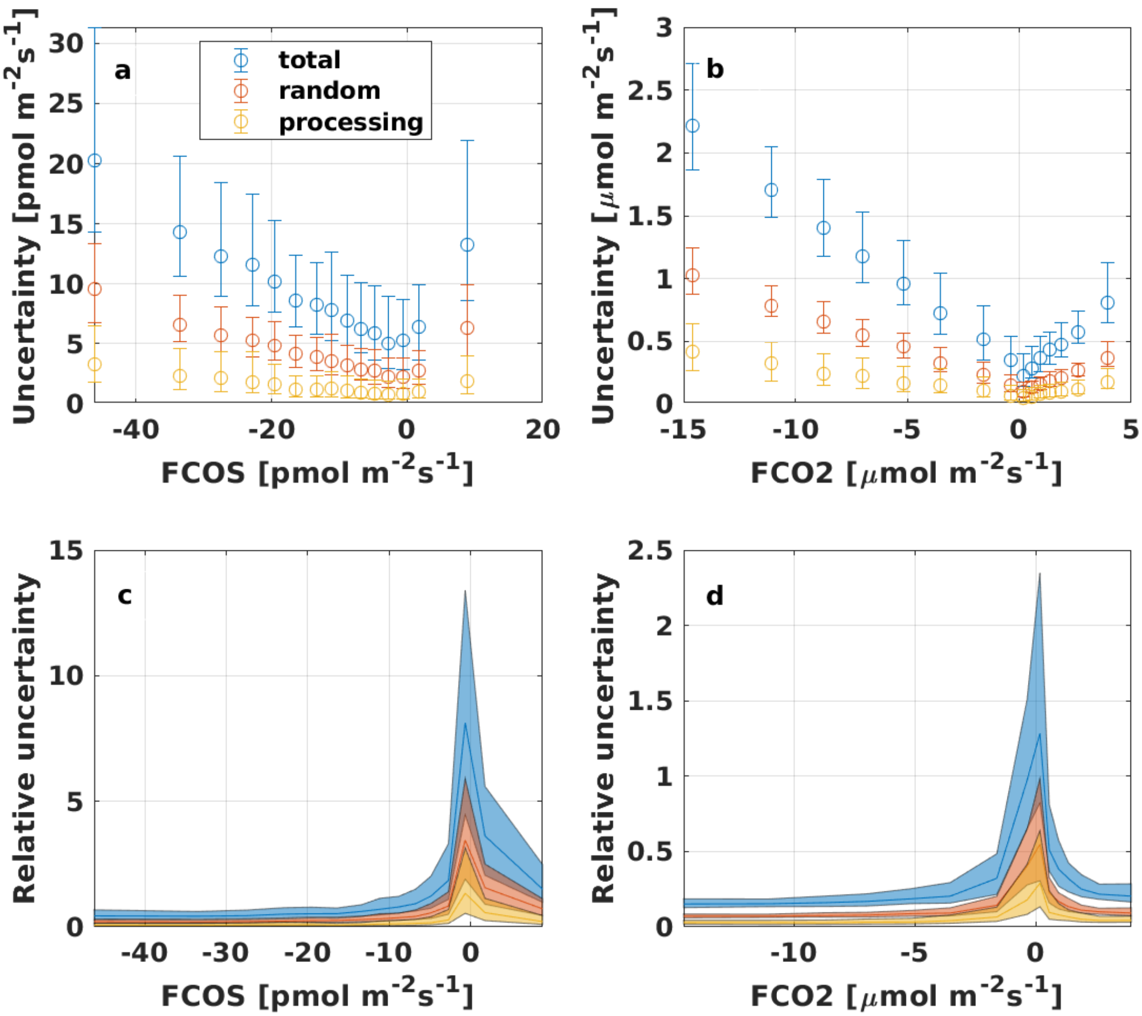


Figure 8. Uncertainty of COS and CO_2 fluxes, binned to 15 equal-sized bins that represent median values (a,b). Error bars show the 25th and 75th percentiles. Total uncertainty is represented as the 95 % confidence interval (1.96 ϵ_{comb}) Panels c and d represent the relative uncertainty, i.e. the uncertainty divided by the flux, for COS and CO_2, respectively.

4 Conclusions

In this study, we examined the effects of different processing steps on COS EC fluxes and compared them to CO_2 flux processing. COS fluxes were calculated with five lag time time lag determination methods, three detrending methods, two high frequency high-frequency spectral correction methods and with no spectral corrections. We calculated the storage change fluxes of COS and CO_2 from two different concentration profiles and investigated the diurnal variation in the storage change fluxesand vertical advection. We also applied u_* filtering and introduced a gap-filling method for COS fluxes. We also quantified the uncertainties of COS and CO_2 fluxes.

The largest differences in the final fluxes came from lag time time lag determination and spectral corrections. Different lag time time lag methods made a difference of maximum 12.7 15.9 % in the cumulative median COS flux while spectral corrections influenced the cumulative flux by 9.9 median flux by 14.2 %. We suggest to use a combination of COS and CO_2 lag times time lag for COS flux calculation , depending on the COS flux uncertainty, so that potential biases in the determined lag times for small fluxes due to low signal-to-noise ratio of COS mixing ratio measurements can be eliminated. CO_2 mixing ratio is measured simulatenously with COS mixing ratio with the Aerodyne QCLS and in most cases has a higher signal-to-noise ratio and more clear cross-covariance with w than COS. Experimental high frequency correction is recommended for accurately correcting for site specific spectral losses. Different detrending methods made a maximum of 1.1 6.2 % difference in the cumulative median COS flux while it was more important for CO_2 (11.22 10.7 % difference between linear detrending and recursive filteringblock averaging). We recommend comparing the effect of different detrending methods on the final flux for each site separately, to determine the site and instrument specific trends in the raw data.

Flux uncertainties of COS and CO_2 followed a similar trend of higher relative uncertainty at low flux values and random flux uncertainty dominating over uncertainty related to processing in the total flux uncertainty. The relative uncertainty was more than 5 times higher for COS than for CO_2 at low flux values (absolute COS flux less than 3 pmol m^{-2} s^{-1}), while at higher fluxes they were more similar.

We emphasize the importance of lag time time lag method selection for small fluxes, whose uncertainty may exceed the flux itself, to avoid systematic biases. COS EC flux processing follows similar steps as other fluxes with low signal-to-noise ratio, such as CH_4 and N_2O, but as there are no sudden bursts of COS expected and its diurnal behaviour is close to CO_2, some processing steps are more similar to CO_2 flux processing. In particular, lag time time lag determination and high-frequency spectral corrections should follow the protocol of low signal-to-noise ratio fluxes (Nemitz et al., 2018), while QA/QC, despiking, u_* filtering and storage change correction should follow the protocol produced for CO_2 flux measurements (Sabbatini et al., 2018). Our recommendation for time lag determination (CO_2 cross-covariance) differs from the most commonly used method so far (COS cross-covariance), while experimental high frequency spectral correction has been widely applied already before. Many earlier studies have neglected the storage change flux, but we emphasize its importance in the diurnal variation of COS exchange. In addition, we encourage implementing gap-filling to future COS flux calculations for eliminating short-term gaps in data.

Data availability. Data sets will be open and available before the final submission.

Author contributions. KMK and IM designed the study and KMK processed and analyzed the data. PK developed the gap-filling function for COS and participated in data processing. IM, LK, US and HC participated in field measurements. KMK and IM wrote the manuscript with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank the Hyytiälä Forestry Field Station staff for all their technical support, especially Helmi-Marja Keskinen and Janne Levula. ~~KME~~ KMK thanks The Vilho, Yrjö and Kalle Väisälä foundation for their kind support. The authors thank ICOS-FINLAND (319871), the Academy of Finland Center of Excellence (307331) and Academy Professor projects 284701 and 282842, and the ERC-advanced funding scheme (AdG 2016 Project number: 742798, Project Acronym: COS-OCS).

References

Asaf, D., Rotenberg, E., Tatarinov, F., Dicken, U., Montzka, S. A., and Yakir, D.: Ecosystem photosynthesis inferred from measurements of carbonyl sulphide flux, Nature Geoscience, 6, 186–190, https://doi.org/10.1038/NGEO1730, 2013.

Aubinet, M.: Eddy covariance CO2 flux measurements in nocturnal conditions: an analysis of the problem, Ecological Applications, 18, 1368–1378, 2008.

Aubinet, M., Grelle, A., Ibrom, A., Rannik, Ü., Moncrieff, J., Foken, T., Kowalski, A. S., Martin, P. H., Berbigier, P., Bernhofer, C., Clement, R., Elbers, J., Granier, A., Grünwald, T., Morgenstern, K., Pilegaard, K., Rebmann, C., Snijders, W., Valentini, R., and Vesala, T.: Estimates of the annual net carbon and water exchange of forests: the EUROFLUX methodology, in: Advances in ecological research, vol. 30, pp. 113–175, 2000.

Aubinet, M., Heinesch, B., and Yernaux, M.: Horizontal and vertical CO2 advection in a sloping forest, Boundary-Layer Meteorology, 108, 397–417, 2003.

Aubinet, M., Berbigier, P., Bernhofer, C., Cescatti, A., Feigenwinter, C., Granier, A., Grünwald, T., Havrankova, K., Heinesch, B., Longdoz, B., Marcolla, B., Montagnani, L., and Sedlak, P.: Comparing CO2 storage and advection conditions at night at different carboeuroflux sites, Boundary-Layer Meteorology, 116, 63–93, 2005.

Aubinet, M., Feigenwinter, C., Heinesch, B., Bernhofer, C., Canepa, E., Lindroth, A., Montagnani, L., Rebmann, C., Sedlak, P., and Van Gorsel, E.: Direct advection measurements do not help to solve the night-time CO_2 closure problem: Evidence from three different forests, Agricultural and forest meteorology, 150, 655–664, 2010.

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

Billesbach, D., Berry, J., Seibt, U., Maseyk, K., Torn, M., Fischer, M., Abu-Naser, M., and Campbell, J.: Growing season eddy covariance measurements of carbonyl sulfide and CO_2 fluxes: COS and CO_2 relationships in Southern Great Plains winter wheat, Agricultural and Forest Meteorology, 184, 48–55, https://doi.org/10.1016/j.agrformet.2013.06.007, 2014.

Blonquist, J. M., Montzka, S. A., Munger, J. W., Yakir, D., Desai, A. R., Dragoni, D., Griffis, T. J., Monson, R. K., Scott, R. L., and Bowling, D. R.: The potential of carbonyl sulfide as a proxy for gross primary production at flux tower sites, Journal of Geophysical Research, 116, https://doi.org/10.1029/2011JG001723, 2011.

Commane, R., Meredith, L., Baker, I. T., Berry, J. A., Munger, J. W., and Montzka, S. A.: Seasonal fluxes of carbonyl sulfide in a midlatitude forest, Proceedings of the National Academy of Sciences, 112, 14 162—-14 167, https://doi.org/10.1073/pnas.1504131112, 2015.

De Ligne, A., Heinesch, B., and Aubinet, M.: New transfer functions for correcting turbulent water vapour fluxes, Boundary-layer meteorology, 137, 205–221, 2010.

Eugster, W., Zeyer, K., Zeeman, M., Michna, P., Zingg, A., Buchmann, N., and Emmenegger, L.: Methodical study of nitrous oxide eddy covariance measurements using quantum cascade laser spectrometery over a Swiss forest, Biogeosciences, 4, 927–939, 2007.

Finkelstein, P. and Sims, P.: Sampling error in eddy correlation flux measurements, Journal of Geophysical Research: Atmospheres, 106, 3503–3509, 2001.

Franz, D., Acosta, M., Altimir, N., Arriga, N., Arrouays, D., Aubinet, M., Aurela, M., Ayres, E., López-Ballesteros, A., Barbaste, M., et al.: Towards long-term standardized carbon and greenhouse gas observations for monitoring Europe's terrestrial ecosystems: a review, Int. Agrophys., 32, 471–494, https://doi.org/10.1515/intag-2017-0039, 2018.

Fratini, G., Sabbatini, S., Ediger, K., Riensche, B., Burba, G., Nicolini, G., Vitale, D., and Papale, D.: Eddy covariance flux errors due to random and systematic timing errors during data acquisition, 2018.

Gerdel, K., Spielmann, F., Hammerle, A., and Wohlfahrt, G.: Eddy covariance carbonyl sulfide flux measurements with a quantum cascade laser absorption spectrometer, Atmos. Meas. Tech., 10, 3525–3537, 2017.

Hari, P. and Kulmala, M.: Station for Measuring Ecosystem–Atmosphere Relations (SMEAR II), Boreal Environment Research, 10, 315–322, 2005.

Horst, T. W.: A Simple Formula for Attenuation of Eddy Fluxes, Boundary-Layer Meteorology, 82, 219–233, 1997.

Kaimal, J. C. and Finnigan, J. J.: Atmospheric boundary layer flows: their structure and measurement, Oxford university press, 1994.

Kooijmans, L. M., Uitslag, N. A., Zahniser, M. S., Nelson, D. D., Montzka, S. A., and Chen, H.: Continuous and high-precision atmospheric concentration measurements of COS, CO2, CO and H2O using a quantum cascade laser spectrometer (QCLS), Atmospheric Measurement Techniques, 9, 5293–5314, https://doi.org/10.5194/amt-9-5293-2016, 2016.

Kooijmans, L. M., Sun, W., Aalto, J., Erkkilä, K.-M., Maseyk, K., Seibt, U., Vesala, T., Mammarella, I., and Chen, H.: Influences of light and humidity on carbonyl sulfide-based estimates of photosynthesis, Proceedings of the National Academy of Sciences, 116, 2470–2475, 2019.

Kooijmans, L. M. J., Maseyk, K., Seibt, U., Sun, W., Vesala, T., Mammarella, I., Kolari, P., Aalto, J., Franchin, A., Vecchi, R., Valli, G., and Chen, H.: Canopy uptake dominates nighttime carbonyl sulfide fluxes in a boreal forest, Atmospheric Chemistry and Physics, 17, 11 453–11 465, https://doi.org/10.5194/acp-17-11453-2017, 2017.

Kristensen, L.: Time series analysis, Dealing with imperfect data, Riso National Laboratory, Roskilde, Denmark, 1998.

Langford, B., Acton, W., Ammann, C., Valach, A., and Nemitz, E.: Eddy-covariance data with low signal-to-noise ratio: time-lag determination , uncertainties and limit of detection, Atmospheric Measurement Techniques, 8, 4197–4213, https://doi.org/10.5194/amt-8-4197-2015, 2015.

Lee, X.: On micrometeorological observations of surface-air exchange over tall vegetation, Agricultural and Forest Meteorology, 91, 39–49, 1998.

Lee, X. and Finnigan, J.: Coordinate systems and flux bias error, in: Handbook of Micrometeorology, pp. 33–66, Springer, 2004.

Lenschow, D., Mann, J., and Kristensen, L.: How long is long enough when measuring fluxes and other turbulence statistics?, Journal of Atmospheric and Oceanic Technology, 11, 661–673, 1994.

Lenschow, D., Wulfmeyer, V., and Senff, C.: Measuring Second- through Fourth-Order Moments in Noisy Data, Journal of Atmospheric and Oceanic Technology, 17, 1330–1347, 2000.

Mammarella, I., Kolari, P., Rinne, J., Keronen, P., Pumpanen, J., and Vesala, T.: Determining the contribution of vertical advection to the net ecosystem exchange at Hyytiälä forest, Finland, Tellus, Series B: Chemical and Physical Meteorology, 59, 900–909, https://doi.org/10.1111/j.1600-0889.2007.00306.x, 2007.

Mammarella, I., Launiainen, S., Grönholm, T., Keronen, P., Pumpanen, J., Rannik, Ü., and Vesala, T.: 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, 1856–1866, https://doi.org/10.1175/2009JTECHA1179.1, 2009.

Mammarella, I., Werle, P., Pihlatie, M., Eugster, W., Haapanala, S., Kiese, R., Markkanen, T., Rannik, U., and Vesala, T.: A case study of eddy covariance flux of N_2O measured within forest ecosystems: quality control and flux error analysis, Biogeosciences, 7, 427–440, https://doi.org/https://doi.org/10.5194/bg-7-427-2010, 2010.

Mammarella, I., Peltola, O., Nordbo, A., Järvi, L., and Rannik, Ü.: Quantifying the uncertainty of eddy covariance fluxes due to the use of different software packages and combinations of processing steps in two contrasting ecosystems, Atmospheric Measurement Techniques, 9, 4915–4933, https://doi.org/10.5194/amt-9-4915-2016, 2016.

Maseyk, K., Berry, J. A., Billesbach, D., Campbell, J. E., Torn, M. S., Zahniser, M., and Seibt, U.: Sources and sinks of carbonyl sulfide in an agricultural field in the Southern Great Plains, Proceedings of the National Academy of Sciences, 111, 9064–9069, https://doi.org/10.1073/pnas.1319132111, 2014.

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

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

Mauder, M., Cuntz, M., Drüe, C., Graf, A., Rebmann, C., Peter, H., Schmidt, M., and Steinbrecher, R.: A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements, Agricultural and Forest Meteorology, 169, 122–135, https://doi.org/10.1016/j.agrformet.2012.09.006, 2013.

Moncrieff, J., Clement, R., Finnigan, J., and Meyers, T.: Averaging, detrending, and filtering of eddy covariance time series, in: Handbook of micrometeorology, edited by Lee, X., Massman, W., and Law, B., pp. 7–31, Springer, 2004.

Moncrieff, J. B., Malhi, Y., and Leuning, R.: The propagation of errors in long-term measurements of land-atmosphere fluxes of carbon and water, Global change biology, 2, 231–240, 1996.

Montagnani, L., Grünwald, T., Kowalski, A., Mammarella, I., Merbold, L., Metzger, S., Sedlàk, P., and Siebicke, L.: Estimating the storage term in eddy covariance measurements : the ICOS methodology, International Agrophysics, 32, 551–567, https://doi.org/10.1515/intag-2017-0037, 2018.

Montzka, S., Calvert, P., Hall, B., Elkins, J., Conway, T., Tans, P., and Sweeney, C.: On the global distribution, seasonality, and budget of atmospheric carbonyl sulfide (COS) and some similarities to CO_2, Journal of Geophysical Research: Atmospheres, 112, 2007.

Moore, C.: Frequency response corrections for eddy correlation systems, Boundary-Layer Meteorology, 37, 17–35, 1986.

Nemitz, E., Mammarella, I., Ibrom, A., Aurela, M., Burba, G. G., Dengel, S., Gielen, B., Grelle, A., Heinesch, B., Herbst, M., Hörtnagl, L., Klemedtsson, L., Lindroth, A., Lohila, A., McDermitt, D., Meier, P., Merbold, L., Nelson, D., Nicolini, G., Nilsson, M., Peltola, O., Rinne, J., and Zahniser, M.: Standardisation of eddy-covariance flux measurements of methane and nitrous oxide, International agrophysics, 32, 517–549, https://doi.org/10.1515/intag-2017-0042, 2018.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., and Yakir, D.: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation, Biogeosciences, 3, 571–583, 2006.

Rannik, Ü.: On the surface layer similarity at a complex forest site, Journal of Geophysical Research: Atmospheres, 103, 8685–8697, 1998.

Rannik, Ü. and Vesala, T.: Autoregressive filtering versus linear detrending in estimation of fluxes by the eddy covariance method, Boundary-Layer Meteorology, 91, 259–280, 1999.

Rannik, Ü., Keronen, P., Hari, P., and Vesala, T.: Estimation of forest-atmosphere CO2 exchange by eddy covariance and profile techniques, Agricultural and Forest Meteorology, 126, 141–155, https://doi.org/10.1016/j.agrformet.2004.06.010, 2004.

Rannik, Ü., Mammarella, I., Keronen, P., and Vesala, T.: Vertical advection and nocturnal deposition of ozone over a boreal pine forest, Atmospheric Chemistry and Physics, 9, 2089–2095, 2009.

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

Rebmann, C., Aubinet, M., Schmid, H., Arriga, N., Aurela, M., Burba, G., Clement, R., De Ligne, A., Fratini, G., Gielen, B., et al.: ICOS eddy covariance flux-station site setup: a review, International Agrophysics, 32, 471–494, 2018.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, 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.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm, Global Change Biology, 11, 1424–1439, https://doi.org/10.1111/j.1365-2486.2005.001002.x, 2005.

Sabbatini, S., Mammarella, I., Arriga, N., Fratini, G., Graf, A., Hörtnagl, L., Ibrom, A., Longdoz, B., Mauder, M., Merbold, L., Metzger, S., Montagnani, L., Pitacco, A., Rebmann, C., Sedlak, P., Sigut, L., Vitale, D., and Papale, D.: Eddy covariance raw data processing for CO2 and energy fluxes calculation at ICOS ecosystem stations, International Agrophysics, 32, 495–515, https://doi.org/10.1515/intag-2017-0043, 2018.

Sandoval-Soto, L., Stanimirov, M., Von Hobe, M., Schmitt, V., Valdes, J., Wild, A., and Kesselmeier, J.: Global uptake of carbonyl sulfide (COS) by terrestrial vegetation: Estimates corrected by deposition velocities normalized to the uptake of carbon dioxide (CO_2), Biogeosciences, 2, 125–132, 2005.

Spielmann, F., Wohlfahrt, G., Hammerle, A., Kitz, F., Migliavacca, M., Alberti, G., Ibrom, A., El-Madany, T. S., Gerdel, K., Moreno, G., Kolle, O., Karl, T., Peressotti, A., and Delle Vedove, G.: Gross primary productivity of four European ecosystems constrained by joint CO2 and COS flux measurements, Geophysical Research Letters, 46, 5284–5293, 2019.

Taipale, R., Ruuskanen, T. M., and Rinne, J.: Lag time determination in DEC measurements with PTR-MS, Atmospheric Measurement Techniques, 3, 853–862, 2010.

Vesala, T., Suni, T., Rannik, Ü., Keronen, P., Markkanen, T., Sevanto, S., Grönholm, T., Smolander, S., Kulmala, M., Ilvesniemi, H., et al.: Effect of thinning on surface fluxes in a boreal forest, Global Biogeochemical Cycles, 19, 2005.

Wehr, R., Commane, R., Munger, J. W., McManus, J. B., Nelson, D. D., Zahniser, M. S., Saleska, S. R., and Wofsy, S. C.: Dynamics of canopy stomatal conductance, transpiration, and evaporation in a temperate deciduous forest, validated by carbonyl sulfide uptake, Biogeosciences, 14, 389–401, https://doi.org/10.5194/bg-2016-365, 2017.

Werle, P.: Time domain characterization of micrometeorological data based on a two sample variance, Agricultural and forest meteorology, 150, 832–840, 2010.

Whelan, M. E., Lennartz, S., Gimeno, T. E., Wehr, R., Wohlfahrt, G., Wang, Y., Kooijmans, L., Hilton, T. W., Belviso, S., Peylin, P., Commane, R., Sun, W., Chen, H., Kuai, L., Mammarella, I., Maseyk, K., Berkelhammer, M., Li, K.-F., Yakir, D., Zumkehr, A., Katayama, Y., Ogée, J., Spielmann, F., Kitz, F., Rastogi, B., Kesselmeier, J., Marshall, J., Erkkilä, K.-M., Wingate, L., Meredith, L., He, W., Bunk, R., Launois, T., Vesala, T., Schmidt, J., Fichot, C. G., Seibt, U., Saleska, S., Saltzman, E., Montzka, S., Berry, J. A., and Campbell, J.: Reviews and syntheses: Carbonyl sulfide as a multi-scale tracer for carbon and water cycles, Biogeosciences, 15, 3625–3657, 2018.

Wilczak, J. M., Oncley, S. P., and Stage, S. A.: Sonic anemometer tilt correction algorithms, Boundary-Layer Meteorology, 99, 127–150, 2001.

Wohlfahrt, G., Anfang, C., Bahn, M., Haslwanter, A., Newesely, C., Schmitt, M., Drösler, M., Pfadenhauer, J., and Cernusca, A.: Quantifying nighttime ecosystem respiration of a meadow using eddy covariance, chambers and modelling, Agricultural and Forest Meteorology, 128, 141–162, 2005.

Yang, F., Qubaja, R., Tatarinov, F., Rotenberg, E., and Yakir, D.: Assessing canopy performance using carbonyl sulfide measurements, Global Change Biology, 24, 3486–3498, https://doi.org/10.1111/gcb.14145, 2018.