Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-313-RC1, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "Towards standardized processing of eddy covariance flux measurements of carbonyl sulfide" *by* Kukka-Maaria Kohonen et al.

Georg Wohlfahrt (Referee)

georg.wohlfahrt@uibk.ac.at

Received and published: 31 October 2019

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.

(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.

(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.

(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.

(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).

(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 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.

(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.

(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.

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 ...“ l. 17: replace "due to the use“ with something more suitable, e.g. "motivated by ...“ l. 20: ", but in contrast to CO2, COS is destroyed ...“ 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 l. 29: "... the assumptions underlying the EC method ...“ l. 29-34 and l. 35-41: these two paragraphs are in my view too general to meaningfully add to the introduction 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 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 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 l. 41: the first ones to report on this issue were Ibrom et al. (2007) 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) 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 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 l. 56: if environmental data are lacking too, mean diurnal variation may be used as a last resort l. 64: if the cross-covariance is flat, then a wrong lag time will not have a large effect l. 66: I do not get the "However, . . ." which links to the previous sentence, which does not appear to make sense here l. 69: Gerdel et al. did study lag determination (their section 3.1) and u*-filtering (their Fig. 6) l. 71: actually you do not discuss the "EC flux measurement setup" at all l. 71-73: the introduction should finish with a statement of objectives l. 79: coordinate rotation l. 78-88: does this have any relevance for this study? l. 92: and sonic temperature l. 95: flow rate through Li-6262, same pump as QCL, tube diameter, length, is the same tube as for the QCL? l. 106: is this the mean? what is the standard deviation? 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?! 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. l. 131: might be worth showing this as a histogram in the supplement? 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? l. 142-143: not sure that computation time is a relevant issue nowadays with regard to coordinate rotation 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°? l. 200: in my memory, the first to propose this approach were Aubinet et al. (2000, 2001) l. 201: a site-specific cospectral model was already used by Wohlfahrt et al. (2005) l. 213: which rotation angle? l. 232: but fluxes are available half-hourly – how do you come up with a half-hourly storage change estimate? 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? l. 264: clearly here only Lee (1998) is to credit with this approach l. 270: "missing CO2 fluxes"? l. 282: add units l. 284-286: isn't this a repetition from above? l. 288-289: all measurements are characterised by noise to a certain degree . . . 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 l. 298-299: so what? How does that sentence relate to your results? 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 %) 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? 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 need to calculate a daily or longerterm 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. Table 1: instead of repeating the values for the reference three times, list it only once?! l. 315: how do you know this difference is large on an annual scale? What is the basis for this recommendation? l. 335-336: "low-frequency corrections" l. 336-337:

combine both sentences l. 337-338: this is kind of trivial l. 344: what is the "normal" CO_2 cospectrum? Fig. 7a shows the COS cospectrum, 7b the COS power spectrum – this sentence does not make sense 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 l. 349: in Fig. 7 you are using normalized frequency, thus no units l. 350: while the attenuation is clearly visible for CO_2, it does not show well for COS, whose cospectrum mostly overlaps with the sensible heat cospectrum l. 350-351: how was this calculated? l. 353: indeed this is as expected . . . remove? 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 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 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) 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?! 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 Figure 9: to what period do the data shown refer to? l. 400-401: again, this is well established - l. 418: I would not call data erroneous – they simply do not fulfil the assumptions

underlying our simplified model of surface-atmosphere exchange 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? 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? 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 Figure 11: also shows the vertical advection without being mentioned in the figure legend?! 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 CO_2 flux literature to that end) and then compare the gap-filled against the measured (during the artificially created gaps) fluxes! 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?

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-313, 2019.