Answers to Referees (Comments from the Referees are in black, answers in blue)

Referee Marc Aubinet

1. General
First I must say that I'm never been convinced by the interest of publishing software inter-comparisons. Indeed, when differences between software appear, they are either due to errors in the software or to a difference in the use or the implementation of the computation procedures. In the first case, the solution is to correct the error in the concerned software and this does not deserve publication; in the second case, it is clearer and of much more interest for the readers to focus the discussion on the impact of these procedures themselves. When we performed the first software inter-comparison, in the frame of Euroflux, in 1996, we didn't see any relevance to publish these results but the exercise led us conclude to the necessity of a procedure clarification, which led finally to the publication of the Euroflux methodology paper.
This said, the present paper brings interesting material in that it points out some sources of uncertainty in the computation procedure. In this way, Figure 2 potentially presents a nice synthesis of these impacts. It would however be more useful to scientists if it was structured differently and rather focus on the impact of each computation step on the flux, in relation to the site and to the analyzer. It is also important to specify that the procedures that were identified as critical are those that were treated differently by the software. This list is thus strongly related to the software used and could be not exhaustive. It is also specific to the sites investigated.

We thank the Reviewer for this comment. Indeed, the magnitude of the corrections are site and instrument-specific (as already pointed out also in our paper), and in fact in Fig 2 we have listed the ranges of reported impacts derived from this and other studies. Of course, this may be not exhaustive, but we also think that the recent advances in the instrumentation technology help to reduce the impacts of some of these processing steps on the final fluxes, and then reducing the flux uncertainty.

I found of particular interest the discussion relative to the spectral correction. This point is important as I think that standardized procedures have not been proposed yet and their impact on the fluxes is large and do not concern LE only (see below). I would like to see this part expanded. Concerning the reference cospectrum, in a recent research, we pointed out that, at sites where local cospectra deviated from theoretical ones, the choice of the Kaimal cospectrum as reference could lead to an important overcorrection (in our case, it made the site switch from a carbon sink to a source!). A paper by Mamadou et al. is presently submitted to AFM on this point. I can provide it if you are interested.

We have added some discussion in the revised paper. I am looking forward to see the study by Mamadou et al.

Concerning the theoretical approach, I'm not in favor of its use, at least for closed path.
Indeed this approach does not take all high cut filtering processes into account. As an example, it was found recently that the presence of a rain cap at the tube inlet could affect greatly the system cut off frequency. An enclosed system with a 25 mL rain cap would never provide cut off frequencies larger than 1 Hz at 15 L/min! This was established recently by two teams who published in AMTD (Aubinet et al, Atmos. Meas.Tech. Discuss., 8, 10735–10754, 2015; Metzger et al, Atmos. Meas. Tech. Discuss., 8, 10983–11028, 2015). At present, this effect was not taken into account by theoretical functions which therefore should underestimate the spectral correction factor. An experimental approach appears always preferable to me as it provide validated transfer functions.

We definitely agree on this point. We will add a sentence on the Discussion and make it more clear in the revised manuscript, also referring to these two AMTD papers (which look vey interesting).

2. Clarifications needed I think that some concepts presented in the paper need to be simplified or clarified. I suppose that "spectroscopic correction" refers to the additional cross-sensitivity of some analyzers to water vapor. This is really an instrumental effect, due to collisional broadening of absorption lines, and it affects all analyzers. The importance of this effect depends on the measurement technique and on the absorption line of concern. It potentially affects both closed and open analyzers and is physically independent of dilution (density?) corrections. I thus think that it could be misleading to associate systematically density and spectroscopic corrections (even if the correction procedure often mixes the two).

We agree with the Reviewer, and in the revised paper we have separated the density and spectroscopic corrections.

In addition, the figure 2 suggests that spectroscopic correction would not be necessary for the closed path systems, which is wrong to my opinion (I suppose that it is included in the conversion to molar fraction). The necessity to apply or not this correction depends mainly of the analyzer: some manufacturers include the correction in the analyzer software (Aerodyne, for example); some do not and suggest a way to implement it like did McDermitt or Rella.

Yes, we agree and we clarify this in the text and we will modify the Fig.2 accordingly.

I think that the authors present WPL (or density or dilution) corrections in a too complicated way. First they use different terms (WPL, dilution or density correction) to characterize what appears to be the same correction to me. Maybe do some expressions refer to only one of the WPL term – density to the H term and dilution to the LE term? If it is the case, this should be clearly stated. Secondly, the effect of WPL correction is exactly the same as the conversion to dry mole fraction. The fact that one or the other procedure is used does not depend on the fact that the analyzer is an open or closed path but mainly on the availability of high frequency measurements of water vapor and temperature (or, in closed paths, the fact that high frequency temperature fluctuations are supposed negligible).

Ok, thanks for the comment. We are now using only the term "WPL correction" (and for closed-path sensor also "dilution correction), following also the suggestion by the 2nd reviewer.

Finally, the case of the LI-7200 is specific as it is not really a closed path, so that "H" term of WPL (or conversion to molar fraction considering air density fluctuations) is necessary but it must be computed on the basis of chamber temperature and not on air temperature. Uncertainties linked to this specific case deserve a discussion. Figure 2 should also be adapted to take this into account.

The conversion to wet (and dry) molar fraction is done by the internal software of Li-7200 using the chamber temperature and pressure (and H_2O), and it is always preferable to use the dry mole fraction for flux calculation, avoiding in this way extra steps in the post-processing.

Some miscellaneous clarification required: P6L9: Spectroscopic correction is sometimes called line broadening correction. Please harmonize.

Done.

P18L23 ISO 8000- 9 recommends using the term "Density" only to characterize single components. In the present case, the right term is "molar concentration".

Done.

The sign convention could lead to confusion: equations 1-4 suggest that micrometeorological sign convention is followed, i.e. downward fluxes are considered as negative and upward fluxes as positive. However, this is contradicted in Figure 3 where FCO2 are considered as positive. As a result, an expression like "a 7% higher FCO2 flux" (P11L27) appears ambiguous: does it finally mean that the sink is over or underestimated?

Indeed we follow the micrometeorological sign convention. The positive CO_2 fluxes (emission) seen in Figure 3h,i are real, and due to the fact that Erottaja is an urban site.

3. Some additional comments on the chapters 3.1 Material and methods The material and methods part is imbalanced as it only presents EddyUH and not Eddy Pro. As said before, the main differences between the results will come from differences in computation procedure and it is therefore important to understand what these differences are.
In addition this part rather looks as an instruction leaflet to the UH program and is of few utility for the present analysis. I thus suggest that this part presents the computation procedures implemented by the two software and focuses on their differences. If the authors find necessary to give an organigram, as in Figure 1, it should be given for both software.
I agree with the experimental protocol that consists in performing different runs, introducing progressively the computation steps. This is indeed a good way to estimate the impact of each individual step.

We guess the reviewer means the chapter 2.1 Software description in Materials and Methods. The original idea to give mainly a description of EddyUH came from the fact that EddyPro is a software widely documented, e.g. via Li-COR Biosciences Inc. webpage, software manual and literature (e.g. Fratini et al., 2015), while similar extensive documentation for EddyUH is not yet available. However, we agree that this part is only partly useful for the present analysis, and we decide to move it in Appendix A (together with the Table 1 and Figure 1). The remaining text introducing the two software was then combined with chapter 2.4.

3.2 Results The result presentation is a little bit intricate and difficult to follow in that it mixes two sites, four systems, two software and seven to eight correction/computation procedures. I suggest to better organize the presentation and to focus on the impact on flux of each correction procedure according to site characteristics and analyzer type.

We thank the reviewer for the suggestion. However, even though we admit that the large number of combinations mentioned by the reviewer might make part(s) of the text difficult to follow, we did not figure out better sequence or way of presentation than we have. At moment the results are not mixed with the discussion, and nicely divided in four subchapters, focusing on 1) software inter-comparison; 2) instrument inter-comparison; 3) Impact of correction steps; 4) inter-comparison of cumulative sums. It is true that we have a lot of information, but this is also a novelty of this study (respect to the previous ones), which includes a wide range of EC systems/gas analyzers in two contrasting ecosystems. Finally, we want to definitely keep the results on software inter-comparison, since it brings several new aspects important for processing the EC data (e.g. difference in spectral correction approaches, including site specific limitation of correction algorithms for the sensor separation; effect of the use of different inputs for certain corrections, like in WPL T-term for LE LI-7500, etc).

3.3 Discussion and conclusions Here again, this part would benefit from a reorganization focusing on the impact of computation procedures on the fluxes rather than on a software comparison. In particular, in the conclusion (P14 L23), if I agree that further methodical researches are still necessary, to my opinion, they don't have to focus on software inter-comparison but rather on these impacts and contemplate a larger spectrum of sites (different

ecosystem types, different climates).

We think that the focus in the Discussion is already quite much on the impact of computation procedures (as can be also recognized from the subchapter titles). However, in the revised manuscript we will add some more text related to spectral correction approaches (theoretical vs experimental), and WPL correction. We will also add a sentence in the conclusion and make it more clear that further research has to focus more on the impacts of calculation steps in different ecosystem types and climates.

4. Miscellaneous P6 L27-28: Units are probably not correct (not the same for each flux). Please clarify.

Yes, we have now specified the units for each flux.

P7 L8-9 : I suggest to rephrase : “was obtained for LE and FCO2, measured by LI-7000, LE and FCH4 by G1301-f, at Siikaneva (Figs. 3d,e and 3c,a) and FCO2 by LI-7200 at Erottaja (Fig. 3i).”

Done.

P12L15 : not really the ecosystem type : more specifically the LE importance.

We have rephrased it.

P22 : reference to Rella is incomplete.

Corrected.

Fig 8 : CO2 and H2O curves are difficult to differentiate.

We have improved the Fig 8.