**Answer to Reviewer 1,2**

**This is a simple, clear and useful manuscript.**
**Implementation of the proposed quality flagging scheme when WPL correction is large when compared to the actual gas flux makes good sense, and will improve flux data quality.**
**I have several minor, a few medium, and one major comment. Minor editorial suggestions are comments # 1-3, 6-10, 13, 15 in the attached file.**

We thank George Burba for his helpful comments and have modified or supplemented the paper accordingly.
We have adopted the minor comments given in the supplement, except for comment 9, so that we do not explicitly address them here.

The possible range of $CO_2$ concentrations is 400 – 1500 ppm, but the range for which factor 3 applies is only 1000 – 1500 ppm. Thus, we have not made any changes. But, we have added a "for".

**Medium comments:**

**# 4 – This is quite good but to brief. It may be helpful to explain in more detail.**

See our comment to Rev. 3, line 52-55

**#5 - The surface heating correction seems constant in Fig 1, but it is supposed to change with temperature. Is this a maximum range for coldest temperature on that day?**

In the resolution of the ordinate, the surface heating correction appears constant, since the air temperature changed only insignificantly during the period shown during the polar night (268-270 K). We added the temperature range in the figure caption.

**#11 - The sentence is perhaps a bit too broad. This would not hold over water surfaces.**

We have added "land surfaces" as this is also the main application of flux measurements. The extension to water surfaces would require too much information on temperatures and stratification. Thanks for pointing this out, otherwise the sentence could be misunderstood.

**#12 and #16 - It may be important to also add that relatively imperfections in both zero and slope of the sensor calibrations will affect open-path fluxes during periods with very small natural flux. Especially if calibrations are performed in the field, these calibrations related errors in small fluxes are quite frequent.**

We agree and have added this.

**#14 - Surface heating correction should not really be applied in warm periods. It is still there, but very small, and modeling it brings more harm than good.**

We agree with the reviewer, but unfortunately the correction is sometimes still applied in warm periods, e.g. by Kittler et al. (2017). We have added a note.

**#17 and #18 - I get nearly identical numbers for Table 1 in some cases but very-slightly different for others. It may be because I am using 50% humidity and constant Cp and lambda. It may be helpful to list what humidity, Cp, and lambda are used in this table.**

The quantities cp and lambda are valid for 15 °C (standard atmospheric conditions), see our comment to Rev. 3, line 52-55. The humidity was assumed to be 100 %.
At lower humidity values, the correction factor is a few percent lower in the term of the sensible heat flow. We have added this to the table caption.

**#19 - This sentence seems too categorical. Surely people gave special attention to WPL related uncertainties in the past. Tables 3,4,5 from Burba et al, 2018 (https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.14614) is just one example, but many others did similar calculations before that.**

Thank you for pointing out your paper. We have reworded the sentence and cited your paper. Nevertheless, we generally have the feeling that although the original paper is very clearly presented (see line 142), the fluxes are calculated schematically without paying the necessary attention to the WPL correction.

**What did not seem to have happened in the past is the QC scheme involving such uncertainties in relation to the flux itself. This is a great idea and seems novel.**
**By the way, something similar can also be achieved by simply assuming a WPL uncertainty error bar of 15%. When this error bar crosses zero, the flux should be flagged. But your proposed QC scheme looks much less arbitrary and more sophisticated**.

The suggested control of the error bar in comparison to the zero line is certainly also a possibility. We are not sure whether this could really be applied by all users, as these values are not always implemented in the software. Our method - which is obviously also the opinion of the reviewer - seems to us to be generalisable here and easy to implement in common software packages. We have not made any additions in this regard.

**Major concerns:**

**My only major concern is with how to use MAD in this case and avoid excluding perfectly good data when fluxes are highly variable (sun-shadow for CO2, ebullition for CH4, N2O episodic emissions, etc.). There certainly should be a way to use MAD but successful use would greatly depend on how exactly MAD is implemented.**

The MAD test has been widely used in recent years in the analysis of flux data after its first application by Papale et al. (2006). The relevant equation for the MAD-test is

$$\langle d \rangle - \frac{q \cdot MAD}{0.6745} \leq d_i \leq \langle d \rangle + \frac{q \cdot MAD}{0.6745},$$

with the scaling parameter q
One can distinguish three time ranges in which the test has been applied:
- Detection of spikes in the time range < 1 s (Mauder et al., 2013), q=7
- Detection of individual periods with exceptional fluxes in the time range 5-60 min: short-term methane emissions (Schaller et al., 2019), q=6; Birch effect after forest fires (Oliveira et al., 2021), q=5, $CO_2$ fluxes during ice/snow (Jentzsch et al., 2021), q=4
- Time series of $CO_2$ fluxes, time range > 30 min (Papale et al., 2006), q=4.

The scaling of the MAD test (factor q) is adapted by all authors to the process under investigation. We take up the reviewer's hint that it may be necessary to adapt the MAD test when investigating special events in order to enable reliable detection.

## References

Jentzsch, K., Schulz, A., Pirk, N., Foken, T., Crewell, S., and Boike, J.: High levels of CO2 exchange during synoptic-scale events introduce large uncertainty into the Arctic carbon budget, Geophys. Res. Letters, 48, e2020GL092256, doi: https://doi.org/10.1029/2020GL092256, 2021.

Kittler, F., Eugster, W., Foken, T., Heimann, M., Kolle, O., and Göckede, M.: High-quality eddy-covariance $CO_2$ budgets under cold climate conditions, J. Geophys. Res.: Biogeosci., 122, 2064-2084, doi: 10.1002/2017JG003830, 2017.

Mauder, M., Cuntz, M., Drüe, C., Graf, A., Rebmann, C., Schmid, H. P., Schmidt, M., and Steinbrecher, R.: A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements, Agric. For. Meteorol., 169, 122-135, doi: 10.1016/j.agrformet.2012.09.006, 2013.

Oliveira, B. R. F., Schaller, C., Keizer, J. J., and Foken, T.: Estimating immediate post-fire carbon fluxes using the eddy-covariance technique, Biogeosci., 18, 285-302, doi: 10.5194/bg-18-285-2021, 2021.

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, Biogeoscience, 3, 571-583, doi: 10.5194/bg-3-571-2006, 2006.

Schaller, C., Kittler, F., Foken, T., and Göckede, M.: Characterisation of short-term extreme methane fluxes related to non-turbulent mixing above an Arctic permafrost ecosystem, Atmos. Chem. Phys., 19, 4041-4059, doi: 10.5194/acp-19-4041-2019, 2019.

**Answer to Reviewer 3**

**General comment:**
**This manuscript presents a new criterion for quality selection of small carbon dioxide fluxes. Those fluxes are mostly subject to a relatively large Webb Pearman, and Leuning (WPL) correction affecting flux direction and magnitude significantly. Therefore, the proposed WPL quality criterion could improve flux data quality. Overall, the concept is introduced well and could be implemented in flux processing software with manageable effort.**
**However, I have some minor points and one major, which should be addressed by the authors before the manuscript is suitable for publication in AMT.**

**Minor comments:**

**Line 11: The word "multiple" is a bit imprecise. Maybe you can provide numbers instead?**

See answer to Line 97.

**Line 19: Could you give examples of eddy-covariance packages using WPL? Flux processing scripts using mixing ratios do not need the WPL correction for flux calculation. I would suggest rewording it to "in software packages using molar densities, molar fractions, etc. as concentration unit for …". Consider replacing "uniform" with similar.**

Thank you for pointing out that it is certainly important that the WPL correction be applied primarily to open-path instruments when the input variables are measured as densities or concentrations and not as mixing ratios. We have added a note to this effect.

**Line 21: add commas: , for instance of CO2,**

Was done

**Lines 52-55: Are these values adapted from literature or measurements? Generally, the findings are interesting. Could you give more details to these numbers or support them by a figure which shows the temperature dependence?**

The numerical values are taken from a book chapter with extensive tables based on the current ITS-90 (Preston-Thomas, 1990) and TEOS-10 (Wright et al., 2010;Feistel et al., 2010) standards. The book should have been published long ago but is now scheduled for December 2021. The article can be requested from the author (TF). Most of the tables will be freely available online. As both reviewers have obviously asked for more information, we have created a graphic which will be attached to the article as a supplement**.**

**Lines 57-61: I would suggest to move these lines to the introduction, for example to line 30ff. since they seem repetitive to me.**

We have gladly realised this suggestion.

**Line 64 and 65: Consider adding town and state to the description of the instruments.**

Because of the international character of the research station on Svalbard, we have refrained from using a country name, as is common practice. Settlement is probably the best name for Ny Ålesund (https://en.wikipedia.org/wiki/Ny-%C3%85lesund).

**Line 65: Please add the version number of EddyPro and a citation.**

Was done.

**Lines 66-67: "measurements of this site were published several times" but only two publications are mentioned. Consider rewording of this sentence or provide more references**

We have replaced several times by e.g., as other works are not or hardly relevant in the sense of this publication.

**Line 68: I would add that negative fluxes indicate uptake, positive fluxes emission.**

Was done.

**Line 69: "Also, ..."**

Was done.

**Line 69: "... is well below 50%" of the corrected flux?**

Was done**.**

**Line 70: "During the short period, ..."**

Was done**.**

**Figure 1 and 2: I would suggest to use colors instead of different shades of gray, denote the 2 in CO2 as subscript, and maybe put a box around the legend.**

We have chosen not to use colour for two reasons: (i) When not printed in colour, the graphics must still be legible. (ii) The colours must be chosen in such a way that they are clearly recognisable even for the colour-blind.

The subscript was corrected**.**

**Caption of Fig. 2: Denote the 2 in CO2 as subscript and add a left bracket to Eq. 1)**

Was done.

**Line 76: "In Fig. 2, ..."**

Was done**.**

**Lines 83-84: What do you mean with "The typical measurement error"?**

The addition of "flux error" by Rev. 1 should make it clearer. Typical error of 10 % is a simplification, because there is always an absolute value (lowest limit) and above a certain value the error is about 10 % of the flux. Since different fluxes have different dimensions for the absolute value, we have given the handbook for eddy-covariance measurements as a reference, where this is given in detail. However, we have added "for most of the fluxes".

**Line 97: Can you provide a range of factors instead of using multiple?**

Since reference has already been made to Table 1, in which the magnitude of the WPL correction is given and in the same line it is stated that the problem is particularly relevant for fluxes < 5 µmol $m^{-2}$ $s^{-1}$, there is a very large range of factors between measured flux and WPL correction up to several orders of magnitude, so that we do not want to change the previous formulation.

**Major comment:**

**You showed the applicability of a new quality criterion for flux measurements on a certain event. I think it would strengthen the manuscript significantly if you show the impact of the new criterion on the entire flux time series of 2015. Testing the criterion for a completely different ecosystem may also be possible. How does the WPL criterion affect annual carbon budget? How much of the**

**measured fluxes are filtered out by the new criterion? Are physically plausible events affected by the WPL quality check?**

The annual fluxes in 2015 were investigated comprehensively in Jentzsch et al. (2021). The events of strong $CO_2$ exchange considered here occurred during strong downward sensible heat fluxes, resulting in a large negative WPL correction. The purpose of this publication was to generalise the detection method for these events and to define a quality criterion. In principle, however, there is no difference, so that the results of Jentzsch et al. (2021) can be interpreted as if the proposed test had been applied.

In general, it can be said that measurements with very small fluxes are excluded by gap-filling routines because of the then frequently low wind speeds. They would also have little influence on annual balances unless they have a bias. The test plays a role especially in very specific investigations of individual processes and, as stated in conclusion iii, when data of small respiration fluxes are used to determine gap-filling parameters. In this sense, we see the paper as a suggestion for the FLUXNET community and programmes like NEON and ICOS to review their gap-filling procedures. But this is beyond the scope of this paper.

## References


Feistel, R., Wright, D. G., Kretzschmar, H. J., Hagen, E., Herrmann, S., and Span, R.: Thermodynamic properties of sea air, Ocean Sci., 6, 91-141, doi: 10.5194/os-6-91-2010, 2010.

Jentzsch, K., Schulz, A., Pirk, N., Foken, T., Crewell, S., and Boike, J.: High levels of CO2 exchange during synoptic-scale events introduce large uncertainty into the Arctic carbon budget, Geophys. Res. Letters, 48, e2020GL092256, doi: https://doi.org/10.1029/2020GL092256, 2021.

Preston-Thomas, H.: The international temperature scale of 1990 (ITS-90), Metrologia, 27, 3-10, doi, 1990.

Wright, D. G., Feistel, R., Reissmann, J. H., Miyagawa, K., Jackett, D. R., Wagner, W., Overhoff, U., Guder, C., Feistel, A., and Marion, G. M.: Numerical implementation and oceanographic application of the thermodynamic potentials of liquid water, water vapour, ice, seawater and humid air – Part 2: The library routines, Ocean Sci., 6, 695-718, doi: 10.5194/os-6-695-2010, 2010.

# Supplement
# Importance of the WPL correction for the measurement of small $CO_2$ fluxes

Katharina Jentzsch[1], Julia Boike[1,2], Thomas Foken[3]

[1]Alfred Wegener Institute, Telegrafenberg A6, D-14473 Potsdam, Germany

[2] Geography Department, Humboldt-Universität zu Berlin, D-12489 Berlin, Germany

[3] University of Bayreuth, Bayreuth Center of Ecology and Environmental Research (BayCEER), D-95440 Bayreuth, Germany

*Correspondence to*: Thomas Foken (thomas.foken@uni-bayreuth.de)

**Temperature dependence of quantities in Eq. (1)**

The WPL correction is given by

$$F_c = \overline{w'\rho_c'} + \mu \frac{\overline{\rho_c}}{\overline{\rho_d}} \overline{w'\rho_w'} + (1+\mu\,\sigma)\overline{\rho_c}\,\frac{\overline{w'T'}}{\overline{T}}, \qquad \text{(S1)}$$

Most quantities in Eq. (S1) are temperature dependent. The relevant temperature-dependent quantities of the 2nd term of Eq. (1) are $\overline{\rho_c}/\overline{\rho_d}$ and in the 3rd term $(1+\mu\,\sigma)\overline{\rho_c}/\overline{T}$ (essentially $\sigma$ and $\overline{T}$). The numerical values are taken from a book chapter (Foken et al., 2021) with extensive tables based on the current ITS-90 (Preston-Thomas, 1990) and TEOS-10 (Feistel et al., 2010;Wright et al., 2010) standards. The book should have been published long ago but is now scheduled for December 2021.

Especially the third term is of interest since it has the more significant influence on the WPL correction according to the present study. The results for a constant assumed value of $\rho_c$= 0.00078 kg m$^{-3}$ are given in Fig. S1. Note: The density of $CO_2$ is about a factor of 1.55 greater than that of dry air, and the factor is almost independent of temperature.

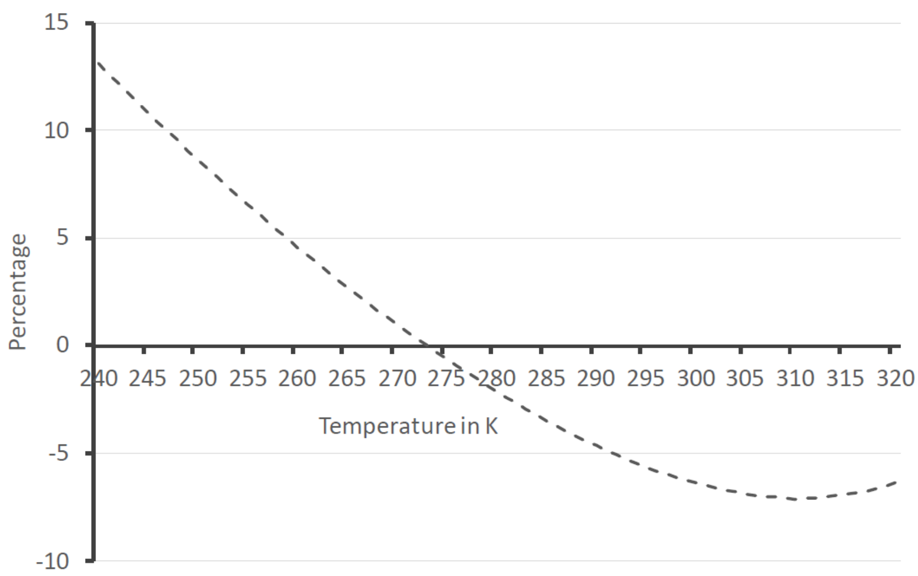


**Figure S1: Percentage of the temperature dependent parts of the 3rd term in relation to the value at 273 K (constant density of carbon dioxide) for kinematic sensible heat flux**

If the sensible heat fluxes are used in energetic units (correction with air density and specific heat at constant pressure), there is nearly no temperature dependency (Fig. S2).

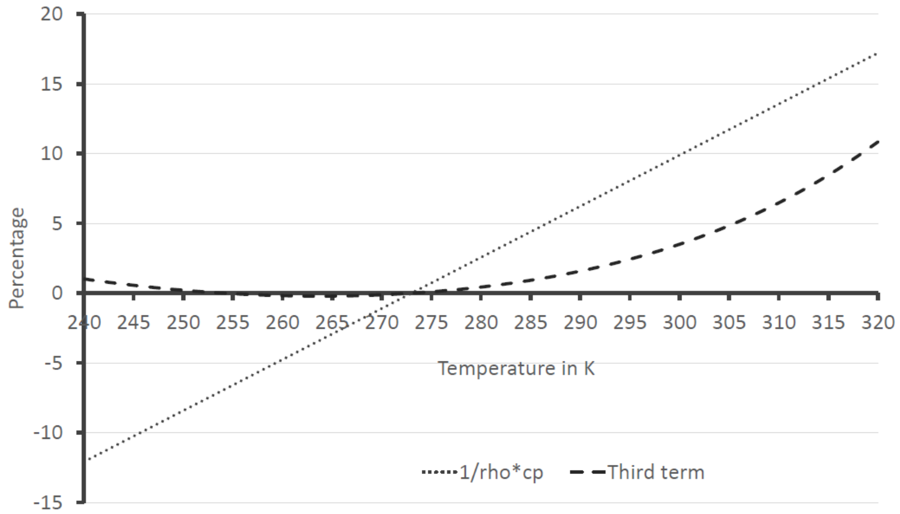


**Figure S2: Percentage of the temperature dependent parts of the 3rd term in relation to the value at 273 K (constant density of carbon dioxide) for energetic sensible heat flux. Additionally, the percentage of the term $1/(\rho_d\, c_p)$ in relation to the value at 273 K is shown**

## References

Feistel, R., Wright, D. G., Kretzschmar, H. J., Hagen, E., Herrmann, S., and Span, R.: Thermodynamic properties of sea air, Ocean Sci., 6, 91-141, doi: 10.5194/os-6-91-2010, 2010.

Foken, T., Hellmuth, O., Huwe, B., and Sonntag, D.: Physical quantities, in: Handbook of Atmospheric Measurements, edited by: Foken, T., Springer Nature, Switzerland, 107-151, doi: 10.1007/978-3-030-51171-4_5, 2021.

Preston-Thomas, H.: The international temperature scale of 1990 (ITS-90), Metrologia, 27, 3-10, 1990.

Wright, D. G., Feistel, R., Reissmann, J. H., Miyagawa, K., Jackett, D. R., Wagner, W., Overhoff, U., Guder, C., Feistel, A., and Marion, G. M.: Numerical implementation and oceanographic application of the thermodynamic potentials of liquid water, water vapour, ice, seawater and humid air – Part 2: The library routines, Ocean Sci., 6, 695-718, doi: 10.5194/os-6-695-2010, 2010.