

Remaining minor points:

Page 1, line 16: change "of sample" to "between sample"

P2, L2: "This compatibility between laboratories is ensured .." -> "This compatibility is ensured .."

P2, L19: "In many previous" -> "In previous"

P2, end of Introduction section: I suggest to formulate the hypothesis that the biases are due to a sensitivity of the internal cavity pressure to water vapor already at this point. Actually, you should explain that, depending on water vapor, the internal cavity pressure sensor produces an erroneous reading, which translates into a bias in dry CO₂ and CH₄. Then explain that experiments were designed to show this issue and to characterize the biases, which ultimately allowed you to formulate a correction model. Without this, the paper is hard to read, since many experiments and results start to make sense only later in the text.

P3, caption of Table 1: "in and experiments" -> "in experiments"

P3, L5: The experiments listed in Table 1 should be better motivated, rather than just stating that experiments were conducted with five Picarros. Please explain the purpose of these experiments. Table 1 is very hard to understand without a brief motivation of the individual experiments. Please also explain the meaning of "usable trace gas measurements" (4th column in Table 1). Why would one list experiments that were not "usable" at all?

P3, Table 1: The table suggests that no H₂O experiment was conducted with the O₂ Picarro #6, in contradiction to the results presented in Sect. 3.2.3.

P4, L10: Shouldn't it be "rather stable"?

P4, L11: Mentioning the fact that CO₂ and CH₄ readings from this experiment were not used seems irrelevant here.

P4, L20: The setup with external pressure sensor doesn't look very complex to me. Wouldn't it be better to write "Due to issues with this setup explained later, .."

Section 2.3.1: Please explain why the external pressure sensor was placed before the inlet valve (or after the outlet valve) of the cavity (as in the response to the reviewer), since this placement is clearly not optimal. Then explain that this allowed monitoring cavity pressure "indirectly" and that the relation between internal cavity pressure and external pressure sensor was established/calibrated in separate experiments with dry air.

P5, L4-6: I couldn't find any indications on Picarro datasheets that the G2207-i instrument returns information on O₂ line width and the optical path length. Does this require operating the instrument in a special mode, or is this part of the housekeeping data?

P5, L13: replace "scale" by "magnitude"

P5, L16: Why do you say "We therefore expect a linear dependence"? Did these studies suggest a linear dependence? If so, please reformulate to make this point clearer.

P5, L26: "their range" -> "the range". This sentence is unclear to me: How can there be a "range" between dry and humid air experiments, if the internal cavity pressure is always regulated to the same value?

Section 2.4, last sentence: Change into a regular sentence (brackets are not needed).

P6, L4: Why should the reader be interested in the median value of 40 min?

P6, L5: Probably one should say "drifted relative to the internal cavity pressure on a timescale ..". What do you mean by "they were calibrated"? Calibrated against what?

P6, L6: It should better be explained why "average readings" were used. Because they were temporally centered on the humidity experiments?

P7, Section 2.6: What was the motivation for using two different H₂O ranges for the cycles?

P8, L8: The title of Section 3.1 is very unspecific.

P8, L10: Probably one should add "as expected" at the end of the sentence.

P8, L13: "were very similar to" -> "differed by only a few percent from"

P8, L14: "this analyzer". Which one?

P9, L6: As mentioned earlier, the paper is confusing if the hypothesis that the internal cavity pressure sensor produces erroneous readings depending on water vapor levels is not formulated earlier on in the manuscript. The sentence starting with "Cavity pressure was estimated .." is a good example for this. You should state again, that since the internal pressure reading was suspected to be wrong, the cavity pressure was additionally estimated based on external pressure readings.

P11, L7: "throughout the experiment". At this stage it is not clear (anymore) what type of experiment this was.

P12, Caption of Fig. 6: Change "The slopes" to "The slopes of the linear parts of the two methods ..".

P13, L15: "yields" -> ", which yields"

P14, caption of Fig. 7: The last line of this caption should be moved to the main text.

P15, Table 5: I don't really understand how the standard deviation for the expanded pressure correction model can be larger (in the case of CH₄) than that of the pressure-correction model. Isn't the expanded pressure correction model directly fitted to the CH₄ measurements, so that it should minimize the differences from the individual data points? (the same question applies to Table 8).

P15, L24: I can't make any sense out of the statement within brackets.

P17, L2: "Dashes lines: as" -> "Dashed lines as"

P17, Table 7: The coefficient h_p of the "joint correction with data from both experiments" is indicated to be $(0.16 \pm 0.04) \% \text{H}_2\text{O}$. The uncertainty range of this coefficient seems too small, since the coefficients of the two experiments separately (0.079 and 0.26) are outside of the range.

P18, L7: "and expanded" -> "and the expanded"

P23, L7-8: This sentence tells the reader at the same time that there were no differences in response between dry and humid air experiments and that, nevertheless, there might be a water-dependent bias. This is very confusing and needs to be briefly explained here (with more details given in the supplement).

P23, L11: What do you mean by "and CH4 data"?? (should probably be deleted)

P23, L18-21: I don't understand how an experiment with dry air can provide useful information on the question, whether cavity pressure may adjust to a new water vapor level on a time scale longer than that of the humid air experiments. This whole paragraph sounds highly speculative to me. Is this really needed?

P23, L29: delete "instead"

P24, L8-10: This sentence rather belongs to the next section 4.3.

P24, L12: " The standard water correction model caused biases" -> "Applying the standard water correction model resulted in biases"

P24, L14: " directly links cavity pressure" -> " directly links cavity pressure estimated from an external pressure sensor"

P24, L22-23: " was based on the parabolic water correction model from the literature and our" -> " combined the parabolic water correction model from the literature with our"

P25, L5: "may help spotting inconsistencies" -> "provides useful information on potential inconsistencies".

P25, L7: You may add that the experiments with stable water vapor levels need to resolve the range of low water vapor levels between 0 to 0.2%.

P25, L12-13: Change sentence " Simultaneously, cavity pressure estimated based on the external pressure sensor was too low and inconsistent in this domain, with the slowest-evaporating droplet closest to the data from experiment with stable water vapor levels" to "Cavity pressure estimated based on the external pressure sensor was lower around the pressure bend position in experiments with fast evaporating droplets than with the slowest-evaporating droplet."

and continue with

"This suggests that the fast water vapor variations .."

P25, L16: "captured exaggerated and inconsistent" -> "tended to exaggerate the"

P25, L18: "slower" -> "more slowly"

P25, L19: delete "than the faster-evaporating droplets"

P25, L24: I suggest to slightly change the structure of Section 4 as follows: Delete title 4.5, change title of 4.5.1 to "4.5 Temporal stability of expanded water correction model" and change title of 4.5.2 to "4.6 Differences of expanded water correction model between analyzers"

P25, L26-29: The sentences referring to the non-useful droplet experiments should be deleted.

P26, L1-2: "far from" -> "well above"

P26, L2-3: delete "between the two experiments" (this should be clear by now)

P26, L9: " with the exception that the effect on CO2 of Picarro #3 appeared diminished" -> "except that the effect on CO2 was reduced for Picarro #3"

P27, L3: "largest at water vapor mole fractions" -> "largest at low water vapor mole fractions"

P27, start of conclusions: I agree with one of the reviewers that it should be stated more clearly that the overall effect is small, especially in the conclusions. Currently, the conclusions section refers to the

WMO compatibility goals (which people assume to be +/- 0.1 ppm (+/-0.05 ppm in S.Hem.) for CO₂ and +/- 2 ppb for CH₄), but actually it seems that you are referring to the "internal reproducibility goals", which is only half the compatibility goals.

P27, L12: "reported on here" -> "reported here"

P27, L21: "we used" -> "used"

P27, L22-27: These new sentences are very knotty (therefore, however, therefore) and could probably be reduced to half the length.