"Atmospheric CO2, CH4, and CO with CRDS technique at the Izaña Global GAW station: instrumental tests, developments and first measurement results" by Angel J. Gomez-Pelaez et al. (https://www.atmos-meas-tech-discuss.net/amt-2017-375/)

Author's Replies to the Comments of Referee #1

This work describes the GHG measurement system at Izana Observatory (IZO). Several novel measurement methods or improvements to standard methods are described, including a new correction to CO2 in Picarro analyzers that is caused by inhomogeneous mixing in the measurement cell of the analyzer that is affected by flow rate. **It should be published after corrections outlined below.**

We acknowledge the constructive and extensive comments of the referee.

R1.M1) One major comment that should be addressed is the determination of the co2 correction for inlet pressure variation should be tested more. Details and recommendations are below.

We have performed additional tests to determine this correction. They confirm our previous findings. In the revised version of the manuscript we are going to detail these new tests and expand the information about the previous tests.

R1.M2) A second is that the authors use a few more complex fits and equations for corrections than reported previously in the literature for these same instruments, and there is little justification or evaluation of the improvement that these complex corrections make on the data, or reduction in error/uncertainty (e.g. quadratic fit to the CO2, and water vapour corrections).

In the new version of the manuscript we justify better and more these points mentioned. We agree with your comment that for CO2 we are using a more complex fitting in mole fraction than it is usually done in the literature (quadratic instead of linear), but this is not the case for the water vapour corrections. Concerning this H2O correction, for CO2 and CH4 we use equations of the same type than in the literature, whereas for CO we use almost the same equation than in the literature, indeed our equation is a bit simpler since one of the polynomials is cubic instead of quartic (see Eq. [5] of Chen et al., 2013, AMT amt-6-1031-2013; and section 2.2.5 of Karion et al. , 2013, AMT amt-6-511-2013). The novelties we present in the H2O correction for CO are in the way the processing is done. We are going to explain this more in the new version of the manuscript.

Major & Minor comments:

R1.m1) Acronyms: Some are introduced without definition in the abstract (CO2, WMO); Others esp. in the introduction are defined backwards from the usual convention, i.e. the acronym comes first with the expansion in parentheses. But not consistently. And some are introduced (WMO) prior to definition. The authors should carefully go through the text and fix these.

Thanks. These is fixed in the new version of the manuscript.

R1.m2) L29: awkward (requirements are higher as higher the lifetime...). Requirements are more stringent for gases with a longer lifetime perhaps?

Done as suggested.

R1.m3) P2, L4: American billion?

"American" removed. Now just "billion" (one thousand million). In the past, in British English one billion was one million million (this continues being the case in some languages).

R1.m4) P2, L4, ppm should be defined here as micromoles per mole of dry air, or dry mole fraction. (and similar for ppb).

We have added these suggestions.

R1.m5) P2 L8 awkward: "therefore, the required frequency... is much lower".

Done.

R1.m6) P2 L30 Tertiary shoudl not be capitalized I don't think.

Agreed.

R1.m7) P2 L32 until now

Done.

R1.m8) P3 L12 "being" should be ", rejecting the first hour as stabilization time."

Done.

R1.m9) Table 1 caption should say "for two averaging times". Some note should be made explaining the negative water vapour values here.

Done. Note that the precision for H2O measurements indicated by the manufacturer of the CRDS is: <200 ppm for 5-second averages, and < 50 ppm for 5-minute averages. Therefore, it is reasonable obtaining negatives values (-2.8 ppm) for H2O when measuring dry air, since, taking into account the precision of the instrument, these values are completely compatible with 0.0 ppm. A note clarifying this point is added.

R1.m10) Table 2 - wouldn't this be a more appropriate place to report the average H2O? It would be good to know if the water vapour had completely left the stream by the time these SD's were taken. Some explanation of the higher SD's over these 10 minutes compared to the Table 1 results should be made - perhaps the cylinder is still "stabilizing", but then did the authors see a trend in the measurement over the 10 minutes? Was any such trend or noise caused by water vapour still drying out?

We agree with your first comment. We are going to include in this table the average H2O. What any of the SD of tables 1 and 2 means is the SD of the sample of 1-minute, 10-minute or 60-minute means obtained along the full duration of the experiment (we clarify it better in the new version of the manuscript). We re-write your second comment as: the SD´s (associated to 10-minute means) in table 2 are larger than the SD´s in table 1 associated with 1-minute means. Apparently, this might seem contradictory since we expect a decrease in the random noise when increasing the averaging period. However, table 1 considers a 24-hour-long test, whereas table 2 considers a 72-hour-long test, and because of this, the SD of the set of 10-minute means is larger, since it includes the drift in the response of the instrument. Indeed, the purpose of this second test is monitoring the instrument drift along 3 days. We also include a note in the new version to point out this fact.

R1.m11) P4 L4: How was the ambient pressure test conducted? This paragraph should have more detail (pressure chamber?). I agree the effect of ambient pressure is small -you could note that this analyzer is not meant for use on aircraft where large pressure changes might occur. But I still wonder if this effect on the measurements is transient or dependent only on the absolute ambient pressure. Yver Kwok et al state that they find no ambient pressure dependence on CO2 and CH4.

In the new version of the manuscript we are including more details about this test. We will also add the note you mention. Thanks. We have not used a pressure chamber for the test, but simply taking into account the atmospheric pressure changes during the 72-hour test, in a similar fashion than ICOS-ATC (ICOS Atmospheric Thematic Centre for Atmosphere) did. Therefore, the test provides an upper-limit for the ambient pressure sensitivity, since there might be instrumental drift not attributable to atmospheric pressure changes. According to our knowledge, Yver Kwok et al. (2015) proceed in the same way than us (not using a pressure chamber). Our main purpose with this test was to discard we had one of the CRDS units whose CO measurements are affected by natural ambient pressure changes. We do think the sensitivities we have obtained for CO2 and CH4 are within what Yver Kwok calls not significant (we obtain 0.0038 ppm/hPa for CO2 and 0.047 ppb/hPa for CH4), as it happens for CO (0.04 ppb/hPa)

R1.m12) P4 L10: these pressures hould be given in millibar or SI units - presumably this is gauge pressure, psig. Absolute pressure would be better, perhaps stating the regulator gauge pressure in parentheses. Was the effect only noted for these two pressures, or was there more testing done at intermediate pressures? Was it repeatable, or are these numbers given only for the single test between two pressures? (this is a question I had later, on how the fit was made to the outlet valve - was the fit based only on two points, and was it evaluated for a repeated test?).

We agree, hPa are going to be used as units. Yes, this was gauge pressure. We are going to use absolute pressure following the referee´s advice. Since the effect is quite small, these two "extreme" pressures were used to maximize the signal to noise ratio. The test was repeated two different days with the same results. When accounting for this effect the fitting to the CO2 response function improved significantly in the first calibrations, when we were not so skilled adjusting precisely the CRDS inlet pressure corresponding to each gas cylinder. Therefore, intermediate pressures had been tested indirectly. The same considerations apply for the outlet valve.

We have performed additional tests to determine this correction, including intermediate pressure points and even a larger pressure range. They confirm our previous findings. In the revised version of the manuscript we are going to detail these new tests and expand the information about the previous tests.

R1.m13) p4 L23: "during that first calibration" - what calibration? perhaps "During the calibration described above,"

Done.

R1.m14) P4 L27, awkward phrasing "to the computation". Perhaps just "Pre-processing refers to the computation of raw data 3-second means... as well as the computation of some derived variables". (what variables?). Also "no" should read "not".

Done. Thanks.

R1.m15) P5 L3-10. It is not clear why certain variables are associated with certain species codes. I thought that the species code had to do with which laser was making the measurement (or the frequency range it was covering). One sentence explaining this perhaps would be useful to the reader. It might not be useful to mention the CO2_dry and CH4_dry variables, as they confuse things and they are not used in this work at all - they get updated after the h2o_reported scan. I find the explanation of each variable useful however. I think CO2_dry is erroneously mentioned twice: once with sepcies=2 and once when species=3.

Yes, you are right, the species code indicates in which spectral range the measurements are performed. There are 4 spectral ranges and 3 lasers. Two nearby spectral ranges are scanned with the same laser. We are going to add a sentence explaining it in the revised version of the manuscript. CO2_dry and CH4_dry are mentioned by completeness and for helping readers to "disentangling" the CRDS raw files (you are right, they are updated once the main spectral line of h2o is measured). The time sequence in which the species are scanned is going also to be indicated: 1, 2, 4 and 3. CO2_dry appears in two species: 2 and 3. This is not a mistake. The manufacturer might have decided two obtain an updated CO2_dry value using the H2O of the previous measurement cycle without waiting to the end of the current measurement cycle.

R1.m16) Why are both the MPV and SV fields used - should mention why these are monitored - presumably your gas handling uses both solenoid and MPV valves controlled by the Picarro software?

You are right, our gas handling system uses both solenoid and MPV valves controlled by the Picarro software. We are going to add a sentence mentioning it.

R1.m17) P5 L19, I believe the flow in the orifice itself is sonic, and is supersonic immediately downstream, depending on the ratio of upstream to cavity pressure. Also pressure units should be in mb or other SI, with torr in parentheses.

You are right, the flow has Mach number one just at the orifice, and larger than one just downstream de orifice. We are going to indicate it in the new version of the manuscript. As it is mentioned in Appendix A, within the inlet pressure range recommended by the manufacturer, the flow is always supersonic just after the orifice. We are going to use hPa units in the new version of the manuscript, as recommended by the referee.

R1.m18) P5 L20 was this linear relationship between OV and inlet pressure confirmed in your tests? This should be investigated and shown, not just assumed, as the inlet pressure/flow correction is using the OV variable as a direct proxy for flow rate essentially.

This linear relationship between OV and inlet pressure has been confirmed in numerous additional tests we have performed (using several pressures in each test). We are going to show this in the new version of the manuscript.

R1.m19) P5 Eq 1 - it would be good to see some more information on this fit. Was it really based only on the two different pressures investigated in the test? Could a fit to more pressures be performed? how good or bad is the fit - what are the residuals from the test itself? This is one of the main findings of this work, and has not been reported before, as the authors point out, so more work is needed to show the reader how the correction was determined, how it performs, and its consistency over time and after a reboot. The fit should

be tested by performing a test again and checking for consistency and for errors after the fit was applied, esp. its reliance on the OV variable.

Please, refer to our replies to comments R1.m12) and R1.m18).

I do believe that because the inlet orifice is not a physical orifice but rather a proportional valve fixed at a certain opening, after the analyzer has a power cycle (reboot), it may not come back to the same spot exactly (the valve is not so precise, so given the same voltage might open to a different diameter slightly), and following that, the OV may not be in the same spot, i.e. for the same inlet pressure as prior to the power cycle, the OV value may differ. I think that this would render equation 1 incorrect and would need to be re-calculated. This should be checked and determined. It may be necessary to install a pressure sensor upstream of the CRDS instrument and perform a linear fit on that rather than relying on the OV. This may not be true if there is a physical critical orifice upstream whose aperture never changes. The aircraft analyzers do have a physical critical orifice so that they can maintain a constant given flow rate, even after the instrument is rebooted.

According to the information provided to me by the manufacturer (Rella, private communication), "there is both a proportional valve and a physical orifice in the inlet system of the G2401. The proportional valve is opened slowly at startup to ensure that the flow smoothly changes, but after this startup procedure, the valve is set to full open, and the flow is set by the orifice".

Note also that if Eq. (1) is expanded, only the slope term in OV is kept fixed, since the independent term is, in practice, combined with the independent term of Eq. (6), and the latter equation is updated each time a mole fraction calibration is performed. Therefore, it automatically takes into account hypothetical drifts in the independent term of Eq. (1). The impact in the slope (in OV) by hypothetical drifts in the outlet valve controller is quite limited as it is going to be shown in the revised version of the manuscript.

The main advantage of the aircraft analysers is that they have the critical orifice at the outlet of the cavity instead of at the inlet. Since at critical orifices the (supersonic) flow rate is determined by the pressure and temperature at its inlet, for aircraft analysers the flow rate is constant no matter the huge changes in the pressure at the CRDS inlet that takes place during a flight.

R1.m20) p6 L22: perhaps rephrase to : "provided results that satisfied our accuracy requirements".

Done. Thanks.

R1.m21) P6 L24: This helps avoid...

Done.

R1.m22) P6 L25: that might propagate to the interior of the cylinders...

Done.

R1.m23) P7 L6: What reason is given for going to a quadratic fit for CO2 here? In other literature (including Yver Kwok et al) a linear fit is generally used. Perhaps the authors can justify the use of a more complex curve for this instrument. It is useful to note that the residuals of a calibration should be compared to the assigned 1-sigma uncertainty of the gas

standard relative to the scale. For NOAA tanks, this is something close to 0.02-0.03 ppm (as cited in Andrews et al., but could be checked for specific cylinders perhaps?). You wouldn't expect the residuals to be better than that.

No reason was provided in this part of the manuscript, but in the last paragraph of Sect. 2. In the new version of the manuscript we are going to justify also before Eq. (6) the reasons for using a quadratic fit.

R1.m24) P7 L6, L9: I would suggest replacing "real dry mole fraction" by "the dry mole fraction assigned by the CCL on the WMO scale".

Done. Also, after Eq. (5).

R1.m25) p7, L16 "high-accuracy"

Done. Since "accuracy" is the noun and "accurate" is the adjective, we preferred to use "high-accurate".

R1.m26) p7, L17 should read: the calibration fits are generally performed using a limited range...

Done. Thanks.

R1.m27) p7 Is the virtual tank concentration its assigned value (x-axis) or measured value (y-axis)? Presumably, following the convention of equations 4-6, it is the x-axis, assigned value that is constant at 400 ppm and the raw response for that value is what is calculated using the equations. Perhaps just state this in the text when you mention the tank values.

The virtual tank concentration is the assigned value (X-axis). Agreed, we are going to state what you mentioned in the text of the revised manuscript.

R1.m28) p8-9: explain that the drift you are referring to is drift in the P, T sensors that then causes the cavity to be controlled at a slightly drifting temperature. P9, L1: Was this supposed to read .152 degrees C / year rather than Torr?

Agreed, in the new version of the manuscript we are going to state explicitly what you mentioned. No, it is 0.152 degrees C/Torr since the first step is to obtain the slope in the p-T space using Eq. (8), and then it is possible to obtain the dp/dt value using Eq. (9). We are going to state it more clearly in the new version and provide also the dT/dt value.

R1.m29) Fig 2: is the slope of the red line here equal to the left hand side of equation 8?

When multiplying that slope by 400/1850, we obtain the left-hand side of Eq. (8) (see the last paragraph of page 7 in the Discussion paper). We are going to state it more explicitly in the new version of the manuscript.

R1.m30) Eqn 8 & 9: A little more information on this would be good. Does this analysis assume that Temperature drift is a linear function of pressure drift? I guess if they are both linearly drifting with time, this makes sense.

Equations (8) and (9) are general relationships between partial derivatives that do not rely on any assumption. They can be directly applied to characterize the response drifts in our paper since Figs. 1 and 2 have shown that the drifts are linear. We are going to state it explicitly in the new version of the manuscript. Since the left-hand-side of Eq. (8) is constant (linear relationship shown in Fig. (2)), the slope dT/dp is constant (due to Eq. (8)). Since the derivative

in time of the fractional change of CH4 is constant (linear relationship shown in Fig. 1), the slope dp/dt is constant (due to Eq. (9)), and therefore p drifts linearly in time. Finally, since dT/dp and dp/dt are constant, dT/dt is also constant, and therefore T drifts linearly in time.

What is the slope? (Yver Kwok report a mean of 2.4). Perhaps a plot of the fractional change rather tahn the plot in Figure 2 would be more applicable to explain equations 8 & 9 (similar to figure 14 in Y-K).

As it was stated at the end of page 7 of the Discussion paper, in our case the slope is 4.495, whereas Yver Kwok obtained a mean value of 2.4 for the set of CRDSs they tested. As mentioned in the response R1.m29, we are going to mention in the new version of the manuscript that both plots are tightly connected simply by multiplying: the X-values by 1/400 and the Y-values by 1/1850.

R1.m31) p9: The singular of species is species, so I think it should be "for each species" (here and throughout the paper).

Done. Thanks.

R1.m32) Fig 3: the Y-axis here is odd in terms of units. Not sure what to do about it.

You are correct, we think this is the best option because it is informative.

R1.m33) p11: L14: explain "since CO standards use to drift upward"? When?

What we mean here is that when a CO standard drifts, this drift is generally positive. We are going to indicate it explicitly in the revised version of the manuscript, and cite a few references, for example: https://www.esrl.noaa.gov/gmd/ccl/co_scale.html .

R1.m34) p11 L15: why at different rates? (or perhaps different rates from the target gas). If the 4 tertiary standards are drifting at different rates, one would expect residuals to be increasing in time on the fits as well? It seems that without further evidence, it might be more appropriate to just state that is is likely that the standards and/or target tanks are drifting.

Yes, this is exactly the case for CO as Fig. 4 shows: the calibration-fit RMS residual increases in time for CO. This is mentioned in the discussion manuscript a few lines before.

R1.m35) Fig4: the labels should be on the Yaxis rather than as titles. Not sure if this was stated earlier, but each of these points represents an average of how long and every how often? This could be mentioned in the caption to remind the reader.

Agreed, we are going to include the labels in the Y-axis. Each point represents the assigned value to a target gas in a mole-fraction calibration performed using 4 WMO laboratory standards (every 3 or 4 weeks); except the red dot that represents the calibration-fit RMS residual. We are going to include this information at the end of the caption: "and mole-fraction calibration performed using 4 WMO laboratory standards (every 3 or 4 weeks)".

R1.m36) Figure 5: I don't find this figure particularly useful. If kept, labels should be put in the photo of the various components.

Agreed. We are going to put labels in the photo of the various components.

R1.m37) Figure 6: y-axis is unlabeled, should have the variable even though there are not units on this raw measurement.

You are correct. We are going to label the y-axis using the variable and removing it from the interior of the plot.

R1.m38) p14 L24: 39 minutes seems to be very long, when looking at figure 6, the water vapour could change significantly in that time, especially at the higher H2O values at the beginning of the test. Is such a long average really required, when you are achieving 1-minute precision of 0.87 ppb?

The experiment lasted many hours. We have checked that using a 39-minute running mean for CO has no significant impact in the accuracy of the data for that experiment. Using a 39-minute running mean instead of a 1-minute running mean, the random noise is reduced by a factor 6 approximately.

Also, given the complex equations presented here for CO water correction, how different is the final result from the internal correction (i.e. the variable "CO" in the raw output)? It would be necessary for the reader to evaluate this and how it depends on the level of water vapour, to see if it's worth implementing.

Agreed. We are going to include a new figure in the revised manuscript to show it.

R1.m39) P15-16, Regarding the -40C trap, can you relate this to a maximum H2O value you see in the data after that point? What is the magnitude of the water correction at those values, and how different is it to CO2_dry and CH4_dry? Rella et al. indicate that at low water vapour values, the internal correction works well. Again, the water vapour tests are cumbersome, especially for field analyzers, in many cases - it would be useful to the readers to know what kind of improvement is gained from the tests as opposed to using the factory correction in cases where there already is partial drying like this.

During the first year in which our CRDS was in operation, no drying was performed. This made necessary the H2O correction. We are going to include a short paragraph in the revised manuscript providing the information you asked for when partial drying is performed.

R1.m40) P16, L7-8: Awkward phrasing "for not discarding". Perhaps "We retain a pre-processed 30-second mean if the following conditions are met: ". Then rephrase the conditions accordingly (i.e. "the mean values of the following variables are within...") etc.

Done.

R1.m41) L9: condition 2 awkward: was exit supposed to be "exist"?

You are correct. Thanks. Changed.

R1.m42) P16 L17: clarify when 10 minutes are discarded (after a calibration or any sample path switch?).

After any sample path switch. This is going to be explicitly indicated in the revised version of the manuscript.

R1.m43) P18 L14: remove the final "in space", as earlier this sentence stated in space and time.

Done. Thanks.

R1.m44) Fig 11, caption should include the fact that these are 12-hour averages (if that is the case).

Yes, this is the case. We are going to include this information in the caption.

R1.m45) Table 4, first row, month says 2017? I wonder if the information in this table would be better expressed in a figure somehow, with dashed lines indicating the WMO compatibility goals?

Indeed, the first row shows the mean values for the period 2015-2017. We need to change the table to make it clearer. If possible, we would prefer to keep the table.

R1.m46) P21 L4: shown

Done.

R1.m47) In section 6.3, the reason for the CRDS-RGA differences for CO is ascribed as due to the CRDS standard drift, but here in Section 7, it is indicated that the drift did not improve the comparison. Can the authors give any explanation for this? Even if the RGA tanks are also drifting, the correction described in Section 7 would put both data sets relative to the same tanks after all.

Section 6.3 mentioned that only as a hypothesis. We have no explanation for the results obtained in Section 7 concerning the CO comparison. Perhaps the explanation might be in the problems detected recently by the CCL in the CO WMO-X2014A scale (https://www.esrl.noaa.gov/gmd/ccl/co_scale_update.html).

R1.m48) Finally, overall, given the large effort involved in achieving very high accuracy for this data set, do the authors envision prescribing an uncertainty to their measurements? [perhaps this is outside the scope of this paper, but something to think about maybe for the future].

That is outside the scope of this paper. Thanks for suggesting it as something to think about for the future.