

1 **Author's response to Interactive comment by Anonymous Referee #1 on**  
2 **"Evaluation of the IAGOS-Core GHG Package H<sub>2</sub>O measurements during**  
3 **the DENCHAR airborne inter-comparison campaign in 2011" by Annette**  
4 **Filges et al.**  
5

6 **About the paper**

7 This manuscript describes the new water vapor instrument for IAGOS commercial flight  
8 measurements, its calibration (in several ways), and some preliminary comparisons with other  
9 water vapor instruments during an aircraft-based inter-comparison campaign in 2011. The paper is,  
10 in many instances, very detailed and provides ample descriptions of the calibration and evaluation,  
11 but at times does not provide the information necessary to completely understand some of the  
12 findings. My finding is that the paper can be published in AMT if the shortcomings (noted below)  
13 are improved.  
14

15 **General Comments**

- 16 1) I don't agree with the measurement "repeatability" values determined in this paper for WV  
17 mixing ratios <100 ppm. On page 14 it is stated that an "upper limit for the measurement  
18 repeatability" .. is : : : "4 ppm or 5%, whichever is greater for WV < 100 ppm". If that is the case,  
19 why are there so many data points for WV < 100 ppm in Figure 7 that have standard deviations  
20 > 5 ppm (5%)? The upper limit estimate for WV > 100 ppm is much more in line with the Figure 7  
21 data.

22 Measurement repeatability values were revised. The new estimates are 4 ppm for H<sub>2</sub>O <10 ppm,  
23 20 % or 10 ppm (whichever is smaller) for 10 ppm < H<sub>2</sub>O <100 ppm, and 5 % or 30 ppm  
24 (whichever is smaller) for H<sub>2</sub>O >100 ppm

25 Changed/Added in Abstract, page 2 lines 1-2, in Section 4.1.1, page 14 lines 26-27, and in  
26 Conclusions, page 18 lines 4-5:

27 "... of 4 ppm for H<sub>2</sub>O <10 ppm, 20 % or 10 ppm (whichever is smaller) for 10 ppm < H<sub>2</sub>O <100  
28 ppm, and 5 % or 30 ppm (whichever is smaller) for H<sub>2</sub>O >100 ppm."

29 Added in Figure 7: line indicating a standard deviation of 20 %

30 Added in caption of Figure 7, line 6: "... of 4 and 30 ppm, and 5 and 20 % respectively."

- 31 2) The abstract is very, very detailed, and can be shortened by describing the overall work in more  
32 general terms. An example of too much detail is the inlet system description that is described  
33 again in similar (or even less) detail on page 6. Also, what were the quantitative goals of  
34 comparing the CRDS instrument to others? In other words, what levels of performance were you  
35 hoping to obtain from the CRDS in terms of precision, accuracy and stability?

36 Lines 23-28 in the Abstract ("The inlet system, .... within IAGOS") were deleted.

37 Added in Section 2, page 6 line 11: "operation of the instrument *with a controlled pressure in the*  
38 *sample cell of 186.65 hPa (140 Torr) throughout the aircraft altitude operating range..."*

39 Precision at low mixing ratios is known from previous studies to be around 6 ppm for a 2.3  
40 second integration time, but is related to white noise, such that the corresponding precision for  
41 an integration time of 2300 seconds drops to below 0.3 ppm (Filges et al, 2015). Regarding  
42 stability, there is not much known for the CRDS system, also regarding accuracy. For CO<sub>2</sub>  
43 measurements the system has shown excellent stability, but it is unclear if this is the case for  
44 water vapor. So rather than having clear targets, the aim was to see the results from the  
45 comparison against reference instruments, and then limit the use of the data to those areas  
46 where the quality is sufficient.

- 1 Added in Abstract, page 1 line 19: "... project. *Since the quantitative capabilities of the CRDS*  
2 *water vapor measurements were never evaluated and reviewed in detail in a publication before*  
3 ~~Thus, the ...~~"
- 4 Added and changed in Section 1, page 4 lines 28-29:  
5 "*In contrast to the CO<sub>2</sub> measurements from the CRDS, which have been studied thoroughly and*  
6 *have shown good performances, the quantitative capabilities of the CRDS water vapor*  
7 *measurements were never evaluated and reviewed in detail before. Precision in the laboratory is*  
8 *known from previous studies to be around 6 ppm for a 2.3 second integration time, but is related*  
9 *to white noise (Filges et al., 2015). Thus, sample averaging over 30 minutes can result in a*  
10 *precision of down to 0.3 ppm, which in principle can result in numerous scientific applications of*  
11 *the data. Each IAGOS aircraft is also equipped with ~~Since~~ the MOZAIC humidity device (Helten et*  
12 *al., 1998, Smit et al., 2008, Smit et al., 2013) ~~is deployed on each IAGOS aircraft, this~~*  
13 *combination, which provides ...*"
- 14 3) The frequent interchange between using ppm and % as units of mixing ratios (mole fractions)  
15 can lead to confusion, especially when it is desired to report relative differences or uncertainties  
16 (in percent). An example is Lines 5-6 on page 10 where "relative" is required to differentiate  
17 relative differences (in %) from mole fractions (in %). Mentally converting % to ppm is not  
18 terribly difficult, but by changing all the text and graph axes from % to ppm, you would  
19 completely avoid any potential confusion. Some of your 2-panel figures show mixing ratios in %  
20 (top panel) while differences or residuals are presented in ppm (bottom panel).  
21 Mole fractions in unit "%" were changed to "ppm" in the whole manuscript (text, tables, and  
22 figures). Ratios in "% (relative)" were changed to "%".
- 23 4) In many places uncertainties and error bars are presented without any indication if they are  
24 based on 1 or 2 standard deviations of the mean. It is helpful to report uncertainties, but only if  
25 the reader knows on which statistics they are based.  
26 Added in section 3.1.2, page 7 line 30: "... dew point (2-sigma) ..."  
27 Added in section 3.1.3, page 8 line 9: "... 4 % (1-sigma) by ..."  
28 Added in section 3.2.1, page 9 line 30: "... dew point (2-sigma). ..."  
29 Added in section 3.2.1, page 10 line 28: "... 1.3 %, 2-sigma)..."  
30 Added in section 3.2.3, page 11 line 28: "... 1.3 %, 2-sigma)..."  
31 Added in section 3.2.3, page 12 line 5: "... uncertainty (1-sigma) ..."  
32 Added in section 3.2.3, page 13 line 5: "... uncertainty (1-sigma) ..."  
33 Added in section 4, page 14 line6: "... accuracy of  $\pm 0.1^{\circ}\text{C}$  (1-sigma) dew point ..."  
34 Added in section 4, page 14 line9: "... accuracy of 6-8 % (1-sigma) in ..."  
35 Added in section 4.1.1, page 14 line26: "... measurement repeatability (1-sigma) of ..."  
36 Added in section 4.1.1, caption of table 2: "... repeatability estimates (1-sigma) ..."  
37 Added in Conclusions, page 17 line 23: "... 1.3 %, 2-sigma)..."  
38 Added in Conclusions, page 17 line 26: "... 4 % uncertainty range (1-sigma) of..."  
39 Added in Conclusions, page 17 line 29: "...uncertainty (1-sigma) ..."  
40 Added in Conclusions, page 18 line 3: "...precision (1-sigma) ..."  
41 Added in Conclusions, page 18 line 13: "Accuracy (1-sigma) of ..."
- 42 5) Please refrain from including Figure caption information in the main body text of the paper. It  
43 makes the paper more tedious to read. This occurs in numerous places throughout the  
44 manuscript.  
45 Changed in Section 3.2.1, page 10 line 19: "... of the 2014-comparison (~~black points~~)."  
46 Changed in Section 3.2.2, page 11 line 5: "... the data (~~blue line~~) shows ..."  
47 Changed in Section 3.2.3, page 11 line 25: "... the data (~~blue line~~) indicates ..."  
48 Changed in Section 4.1.1, page 14 line 27: "... ~~indicated by the three black lines in Fig. 7.~~"  
49 Changed in Section 4.1.2, page 15 line 9-10: "... ~~shown here as black points (30 seconds mean as~~  
50 ~~grey points)~~ and ... CR-2 (~~dark blue squares~~) and FISH (~~light blue triangles~~), ..."

1 Changed in Section 4.1.2, page 15 line 12-14: "... (Tatrai et al., 2015), ~~are shown as orange diamonds;~~ and flight data ... (Buchholz et al., 2013), ~~are shown as green triangles ...~~"

2

3 Changed in Section 4.1.2, page 15 line 16: "... water vapor (~~violet points~~) is added ..."

4 Changed in Section 4.1.3, page 15 line 26-28: "... water vapor (~~in black~~) and ... CR-2 (~~dark blue~~) and FISH (~~light blue~~), as well as ... pressure levels (~~in green~~), for all four flights. ~~For better comparison also the 30 seconds mean of the CRDS data is shown (in grey).~~"

5

6

7 Changed in Section 4.1.3, page line 2-7: "... in Fig. 11 (~~CRDS CR-2 as black points, CRDS FISH as dark blue diamonds, CR-2 FISH as light blue triangles~~). The water vapor measurements of CR-2 are chosen as x-axis, because they cover all flights in contrast to the FISH data. The differences are calculated from the 30 seconds mean data of all analyzers and are averaged over intervals of 1 ppm, 10 ppm, 100 ppm, 1000 ppm, and 10000 ppm water vapor in the corresponding water vapor ranges of 0-10 ppm, 10-100 ppm, 100-1000 ppm, 1000-10000 ppm, and >10000 ppm, respectively. The standard deviations of the averaged differences are shown as error bars. For plotting reasons all differences <1 ppm are set to 1 ppm. CRDS ..."

8

9

10

11

12

13

14

15 Added in caption of Figure 11: "... differences. The water vapor measurements of CR-2 are chosen as x-axis, because they cover all flights in contrast to the FISH data. For plotting reasons all differences <1 ppm are set to 1 ppm."

16

17

18 Changed in Section 4.1.3, page 16 line 16-20: "... water vapor (~~black points, 30 seconds mean as grey points~~) are plotted ... CR-2 (~~dark blue squares~~), FISH (~~light blue triangles~~), Wasul-Hygro (~~orange diamonds~~) and SEALDH-I (~~green triangles~~). The ... water vapor (~~violet points~~). ..."

19

20

21

22

- 23 6) During the comparison flights, did all instruments sample the airstream from a common Rosemount TAT inlet, or did each have its own inlet? I believe only the FISH inlet is described as being different. If there was a common inlet for the other instruments, how was the sample stream physically split between them? If each instrument had its own inlet, how much of a difference might the different types and/or locations of inlets play in the overall comparisons? The absence of information about the sampling inlet for each instrument makes it very difficult to understand the inlet-dependent discussion on Page 16. I therefore suggest that this discussion be omitted from the paper because (a) it requires adding much greater detail about inlets and (b) inlet influences aren't necessarily a part of the evaluation of the CRDS measurement capabilities for IAGOS.
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32 CR-2, Wasul-Hygro, and SEALDH I were connected to a common backward-faced inlet, FISH used a forward-faced inlet.
- 33
- 34 Added in Section 4, page 14 line 9: "... (Meyer et al., 2015). The CR-2 was connected to a backward-faced inlet to avoid sampling of cloud and ice particles. In contrast, FISH measured ..."
- 35
- 36 Added in Section 4.1.3, page 16 line 19: "... as the CR-2. This is in line with expectation, since all three shared the same backward-faced inlet, which prevented from sampling cloud droplets. However, ~~while~~ the CRDS shows ..."
- 37
- 38
- 39 7) The clarity of the presentation within this manuscript would be significantly improved through editing by a native English speaker. Hopefully that can be done before revisions are submitted.
- 40
- 41 Some proof reading was carried out.
- 42

43 **Specific Comments:**

- 44 1) Page 1, Line 16: "North-Germany" is not a proper noun, so "North" should not be capitalized and the hyphen should be omitted, i.e., "northern Germany". This occurs several times within the manuscript, e.g., Pg 4 L19.
- 45
- 46 Changed in Abstract (page 1, line 16): "...in northern Germany...".
- 47
- 48 Changed in Introduction (page 4, line 19):

1 ~~“...Four flights with a Learjet 35A took place in an area between North-Germany and South-~~  
2 ~~Norway, and North-Poland and the North-Sea respectively, reaching altitudes up to 12.5 km,~~  
3 ~~hence covering also the lower stratosphere...”~~ (See answer to Specific comment 8)

4 Changed in section 4 (page 14, line 4): “...from northern Germany ... to southern Norway...”.

- 5 2) Pg 1, L19: What (specifically) is meant by "assessment validation"? Are there targets for the  
6 measurement precision and/or accuracy of this new IAGOS instrument that are the basis for its  
7 evaluation as described in this paper?

8 Changed in Abstract, page 1 line 19: “... for an initial assessment ~~validation~~ of the ...”

9 See also answer to General Comment 2).

- 10 3) Pg 3, L6: "data of sufficient quality in the UTLS" - sufficient for what? This comment ties in with  
11 my previous comment about targets for "sufficient" measurement precision and/or accuracy.

12 Changed in Introduction, page 3 line 5-8:

13 ~~“Since there are only sparse measurement data of sufficient quality. However, due to only few~~  
14 ~~existing measurement data in the UTLS, and, as emphasized in Solomon et al. (2010),~~  
15 ~~limitations in prognostic models simulations of have difficulties to represent this region well~~  
16 ~~(Solomon et al., 2010), uncertainties in chemistry, transport processes, and trace gas~~  
17 ~~composition are relatively large., which This influences ...”~~

- 18 4) Pg 3, L12-17: why include this long list of satellite-based water vapor sensors when their data  
19 are not included in this paper?

20 Changed in Introduction, page 3 line 13-15: “... (Rind et al., 1993). ~~followed by Recent~~  
21 ~~observations are made by e.g. from e.g. the Halogen Occultation Experiment (HALOE) (Harries~~  
22 ~~et al., 1996), the Aura Microwave Limb Sounder (MLS) (Read et al., 2007; Lambert et al., 2007),~~  
23 ~~as well as the Michelson Interferometer ...”~~

24 Changed in References, page 20 line 10-12:

25 ~~Harries, J. E., Russell, J. M., Tuck, A. F., Gordley, L. L., Purcell, P., Stone, K., Bevilacqua, R. M.,~~  
26 ~~Gunson, M., Nedoluha, G., and Traub, W. A.: Validation of measurements of water vapor from~~  
27 ~~the Halogen Occultation Experiment (HALOE), J. Geophys. Res., 101(D6), 10,205-10,216, 1996.~~

28 Changed in References, page 20 line 29-34:

29 ~~Lambert, A., Read, W., Livesey, N., Santee, M., Manney, G., Froidevaux, L., Wu, D., Schwartz,~~  
30 ~~M., Pumphrey, H., Jimenez, C., Nedoluha, G., Cofield, R., Cuddy, D., Daffer, W., Drouin, B.,~~  
31 ~~Fuller, R., Jarnot, R., Knosp, B., Pickett, H., Perun, V., Snyder, W., Stek, P., Thurstans, R.,~~  
32 ~~Wagner, P., Waters, J., Jucks, K., Toon, G., Stachnik, R., Bernath, P., Boone, C., Walker, K.,~~  
33 ~~Urban, J., Murtagh, D., Elkins, J., and Atlas, E.: Validation of the Aura Microwave Limb Sounder~~  
34 ~~middle atmosphere water vapor and nitrous oxide measurements, J. Geophys. Res., 112,~~  
35 ~~D24S36, doi:10.1029/2007JD008724, 2007.~~

36 Changed in References, page 21 line 21-27:

37 ~~Read, W. G., Lambert, A., Bacmeister, J., Cofield, R. E., Christensen, L. E., Cuddy, D. T., Daffer,~~  
38 ~~W. H., Drouin, B. J., Fetzer, E., Froidevaux, L., Fuller, R., Herman, R., Jarnot, R. F., Jiang, J. H.,~~  
39 ~~Jiang, Y. B., Kelly, K., Knosp, B. W., Kovalenko, L. J., Livesey, N. J., Liu, H. C., Manney, G. L.,~~  
40 ~~Pickett, H. M., Pumphrey, H. C., Rosenlof, K. H., Sabounchi, X., Santee, M. L., Schwartz, M. J.,~~  
41 ~~Snyder, W. V., Stek, P. C., Su, H., Takacs, L. L., Thurstans, R. P., Vomel, H., Wagner, P. A.,~~  
42 ~~Waters, J. W., Webster, C. R., Weinstock, E. M., and Wu, D. L.: Aura Microwave Limb Sounder~~  
43 ~~upper tropospheric and lower stratospheric H<sub>2</sub>O and relative humidity with respect to ice~~  
44 ~~validation, J. Geophys. Res., 112, D24S35, doi:10.1029/2007JD008752, 2007.~~

- 45 5) Pg 3, L18-19: "insufficient spatial resolution" - what spatial resolution is sufficient? And don't  
46 you instead mean "vertical resolution" here when discussing satellite and remote-sensing  
47 observations?

48 Added in Introduction, page 3 line 19:

49 “... by clouds. As shown by Hoareau et al. (2013), vertical resolutions <500 m are needed for the  
50 simulation of cirrus clouds. To represent the very sharp gradient of 40 to 6 ppm water vapor  
51 within 0-2 km at the tropopause (Zahn et al., 2014), resolutions of even 400 m and higher have  
52 to be achieved (Poshyvailo et al., 2018). On the other hand ...”

- 1 Added in References, page 20 line 15:  
2 "*Hoareau, C., Keckhut, P., Noel, V., Chepfer, H., and Baray, J.-L.: A decadal cirrus clouds*  
3 *climatology from ground-based and spaceborne lidars above the south of France (43.9° N–5.7°*  
4 *E), Atmos. Chem. Phys., 13, 6951-6963, doi:10.5194/acp-13-6951-2013, 2013.*"
- 5 Added in References, page 21 line 21:  
6 "*Poshyvailo, L., Müller, R., Konopka, P., Günther, G., Riese, M., Podglajen, A., and Ploeger, F.:*  
7 *Sensitivities of modelled water vapour in the lower stratosphere: temperature uncertainty,*  
8 *effects of horizontal transport and small-scale mixing, Atmos. Chem. Phys., 18, 8505-8527,*  
9 *doi:10.5194/acp-18-8505-2018, 2018.*"
- 10 Added in References, page 23 line 32:  
11 "*Zahn, A., Christner, E., van Velthoven, P. F. J., Rauthe-Schöch, A., and Brenninkmeijer, C. A. M.:*  
12 *Processes controlling water vapor in the upper troposphere/lowermost stratosphere: An*  
13 *analysis of 8 years of monthly measurements by the IAGOS-CARIBIC observatory, J. Geophys.*  
14 *Res. Atmos., 119, 11505–11525, doi:10.1002/2014JD021687, 2014.*"
- 15 6) Pg 4, L6: "The central problematic of all these different ... are the remaining, unexplained  
16 discrepancies ...". Are the discrepancies "remaining"? From what? Are they a problem? It  
17 depends on the science you are investigating. Please improve the clarity of this statement.  
18 Changed in Introduction, page 4 line 6: "... are the ~~remaining~~, unexplained discrepancies ..."  
19 Added in Introduction, page 4 line 11: "Thus, possibilities e.g. to identify long-term trends in  
20 stratospheric water vapor or to study ice microphysical processes are limited (Rollins et al.,  
21 2014)."
- 22 7) Pg 4, L15-18: This is the type of general statement that belongs in the abstract instead of the  
23 very detailed descriptions.  
24 Lines 23-28 in the Abstract ("*The inlet system, .... within IAGOS*") were deleted.
- 25 8) Pg 4: I find a lot of the information presented in the abstract is duplicated here in the  
26 introduction. The introduction should explain why your investigation is important and how you  
27 intend to perform it. Specific information about instruments, including the MOZAIC humidity  
28 device, belongs in a separate section describing the various instruments.  
29 Changed in Introction, page 4 line 18-20: "... water vapor instruments. ~~Four flights with a~~  
30 ~~Learjet 35A took place in an area between North-Germany and South-Norway, and North-~~  
31 ~~Poland and the North Sea respectively, reaching altitudes up to 12.5 km, hence covering also~~  
32 ~~the lower stratosphere. As reference instruments served-The former were the Learjet version~~  
33 ~~...~~"  
34 Added in Section 1, page 4 lines 28-29: "*Therefore, a first impression of the quantitative*  
35 *capabilities of the CRDS water vapor measurements was needed, since they were never*  
36 *evaluated and reviewed in detail in a publication before...."*
- 37 9) Pg 5: There is a lot of repeated information on this page. Please ensure that you don't write  
38 the same sentence more than once in the paper, with the exception of abstract and  
39 conclusions.  
40 Changed in Introduction, page 5 line 2-4: "... long performance record ~~(FISH (Meyer et al.,~~  
41 ~~2015) and CR-2 Cryogenic Aircraft Hygrometer (Buck Research Instruments L.L.C., US,~~  
42 ~~www.hygrometers.com))~~ to evaluate ..."  
43 Changed in Introduction, page 4 line 21: "... Hygrometer (FISH) (Meyer et al., 2015) and a CR-2  
44 Cryogenic Aircraft ~~frost point~~-hygrometer (Buck Research Instruments L.L.C., Boulder, US,  
45 www.hygrometers.com), both ..."  
46 Changed in Introduction, page 5 line 9: "...flight tests, ~~including comparison with the reference~~  
47 ~~instruments, are ..."~~
- 48 10) Pg 5, L24: What "impact on gas density and spectroscopy" is being minimized? This statement  
49 is vague and needs clarification.  
50 Changed in Section 2, page 5 line 24: "*To minimize the impact of pressure and temperature on*  
51 *gas density and spectroscopy, ~~pressure and temperature both are kept constant~~ in the sample*  
52 *cell-are kept constant.*"

- 1 11) Pg 6, L4-6: This description was already given in great detail, in the abstract (where it likely  
2 does not belong).  
3 *Lines 23-28 in the Abstract ("The inlet system, .... within IAGOS") were deleted.*
- 4 12) Pg 6, L13: How are there only small humidity differences between the cabin and the outside  
5 air?  
6 *The cabin air is air conditioned. Therefore, differences in humidity between cabin and ambient  
7 air are small.*  
8 *Added in Section 2, page 6 line 13: "... differences in humidity between the air conditioned  
9 cabin and the ambient air, ..."*
- 10 13) Pg 6, L27-30: This would be the best place to explain that the instrument zero ("offset") is not  
11 stable (over what time scale?) and requires frequent adjustment, while the instrument gain is  
12 very stable and needs calibration only infrequently.  
13 *Added in Section 3.1.1, page 6 line 27: "... an offset correction is required once to improve ..."*  
14 *Added in Section 3.1.1, page 6 line 30: "... water vapor. The offset stability of different CRDS  
15 instruments was checked regularly over a time period of up to 10 years and no significant drift  
16 was observed."*
- 17 14) Pg 9, L30-32 and Figure 3: Are the reported uncertainties 1-sigma? 2-sigma? This comment  
18 applies to the entire paper, wherever uncertainties are given (text and Figures).  
19 *More details about the given uncertainties were added in the entire manuscript.*  
20 *See also answer to General Comment 3).*
- 21 15) Pg 10, L1: "0.0024±0.0021" is "not significant"? What does "significant" mean here? It can't  
22 possibly mean "not statistically different from zero" as the reader might expect.  
23 *Changed in section 3.2.1, page 10 line 1: "The fit-coefficient-impact of the quadratic term  
24 determined as 0.0024 ± 0.0021, on the result is small not-significant-compared to the overall  
25 uncertainty range. Thus, ..."*
- 26 16) Pg 10, L12-14, L20: The differences between calibration curves presented in ppm in Figures 3b  
27 and 4b and expressed in ppm and % (rel) in the text may cause confusion here. If you express  
28 all mixing ratios, differences and residuals in ppm it avoids any confusion. For the Figures 3b  
29 and 4b it would add value to include a right axis of relative differences/residuals. Given the  
30 near-linearity of the differences/residuals against mixing ratios, I think the relative data  
31 (different symbol/color) would be quite constant over the entire range of mixing ratios.  
32 *Mole fractions in unit "%" were changed to "ppm" in the whole manuscript (text, tables, and  
33 figures). Ratios in "% (relative)" were changed to "%".*  
34 *A right axis of relative differences/residuals was added in Figures 3b and 4b.*  
35 *Added in Section 3.2.1, page 10 line 5: "The relative difference of the two calibrations, shown  
36 on the right axis, increases ..."*  
37 *Added in Caption of Figure 3: "... calibration. The relative differences are plotted on the right  
38 axis as light blue triangles."*  
39 *Added in Section 3.2.1, page 10 line 20: "The relative difference (right axis in plot (b)) between  
40 the two experiments is <0.3 % for water vapors <8000 ppm. Since ..."*  
41 *Added in Caption of Figure 4: "...in plot (b). The relative residuals are plotted on the right axis  
42 as light blue triangles. Since the analyzers were not offset corrected before the experiment, the  
43 relative difference of 800 % of the data point at 4.5 ppm is less meaningful and is therefore not  
44 show in order to improve the clarity of the plot."*
- 45 17) Pg 10, L26-28: Do you have enough evidence to make this statement based on a conjecture  
46 that the instrument "was not calibrated well enough" in 2009? This seems like a hand-waving  
47 statement with little factual backing. Is there further evidence that allows you to select the  
48 2013 calibration as being correct?  
49 *Added and changed in Section 3.2.1, page 10 line 26: "... During the 2009 experiment the dew  
50 point mirror has not been calibrated for nearly nine years when it was used for the 2009  
51 experiment, but was not-calibrated well during the 2013 experiment. enough-and- Thus, ..."*

- 1 18) Pg 10, L33: I have to assume that the mixing ratio ranges compared with each other  
2 instrument were chosen based on the best measurement ranges of the other instruments. Is  
3 this the case? If so, please inform the reader of this earlier in the paper. I see the range  
4 information in Table 1, but the table is not mentioned until page 13. An earlier statement  
5 about the choices of comparison ranges for different instruments and a mention of Table 1 are  
6 needed, otherwise it appears the comparison ranges were chosen randomly.  
7 *Added in Section 3.1.2, page 7 line 24: "... dew point. Higher and lower water vapor levels*  
8 *could not be reached due to the environmental conditions in the laboratory."*  
9 *Added in Section 3.1.3, page 8 line 9: "... were measured. This corresponds to the standard*  
10 *calibration range of the FISH calibration bench and is a good addition to the dew point mirror*  
11 *calibration range. Maximum ..."*  
12 *Added in Section 3.1.4, page 9 line 12: "... constant. The adjustable water vapor levels were*  
13 *limited by the remaining humidity in the pressurized air on the one hand and the environmental*  
14 *conditions in the laboratory on the other. The ..."*  
15 *Added in Section 3.2.4, page 13 line 27: "... experiments. The water vapor ranges used in the*  
16 *comparison were determined by the experimental setups of the experiments and the standard*  
17 *calibration ranges of the instruments. The ..."*  
18 19) Pg 11, L5-13 and Figure 5: The residuals in Figure 5b (presented in ppm) are discussed in the  
19 text in relative (%) terms. Adding a right axis for relative residuals to Figure 5b (using different  
20 symbol/color) would strengthen the discussion of relative residuals in the text.  
21 *A right axis of relative differences/residuals was added in Figures 5b.*  
22 *Added in Section 3.2.2, page 11 line 8-10: "... residuals ... for water vapors > 100 ppm, which*  
23 *can be seen in plot (b), are small compared to the uncertainty range of the FISH calibration*  
24 *bench of 4 % indicated by the error bars. For the measurement point at 2 ppm water vapor the*  
25 *relative residuals are larger (6.2 %), due to the influence of the 12.2 ppm offset.*  
26 *Added in Caption of Figure 5: "...of 4 %. The relative residuals are plotted on the right axis as*  
27 *light blue triangles."*  
28 20) Pg 11, L30-31: It is odd that the residuals at 250 ppm are negative (Figure 6b) while at the  
29 lowest mixing ratio they are positive. With an offset of 16 ppm one would expect the residuals  
30 at both of the lowest mixing ratios to be of the same sign. Any idea why they are negative at  
31 250 ppm and positive at the lowest mixing ratio?  
32 *The residuals are negative at 2500 ppm, not 250 ppm, and positive at 300 ppm. Unfortunately*  
33 *there is only one data point at 300 ppm. With more measurements at this water vapor level*  
34 *we would expect the scatter/pattern of the residuals to be similar to the ones at other water*  
35 *vapor levels. Therefore, as described in Section 3.2.3, the uncertainty range of 1.6 ppm of the*  
36 *offset is likely unrealistic and follow on experiments should contain more measurements at*  
37 *low water vapor levels (<300 ppm).*  
38 21) Pg 12, L1: If "the lowest measurement was made at around 300 ppm", why are there data  
39 points below 300 ppm in Figure 6?  
40 *There are no data points below 300 ppm. 1 % water vapor mole fraction corresponds to 10000*  
41 *ppm, not 1000 ppm. For clarification all water vapor values were changed to ppm in the text*  
42 *and figures.*  
43 22) Pg 12, L6-7: "which was tried to exclude" is an example of where a native English speaker could  
44 help clean up the grammar in this manuscript.  
45 *Changed in Section 3.2.3, page 12 line 6-7: "..., which ~~was we~~ tried to avoid ~~exclude~~ by ..."*  
46 23) Pg 12, L10: It's not "zero air" if it contains 2 ppm CH<sub>4</sub>. You might want to use a different term.  
47 *Changed in Section 3.2.3, page 12, line 10: "...the pressurized ~~zero~~ air with 400 ppm CO<sub>2</sub> ..."*  
48 24) Pg 12, L20: Of course, "the pressure sensor is sensitive to water vapor", as it is to any gas-  
49 phase constituent. I think you are trying to convey that the pressure sensor responds in a non-  
50 linear way to increasing amounts of water vapor, creating a water vapor-dependent bias in the  
51 pressure readings. In L25, what does "assumed that the pressure changes linearly" mean here?

1 If the pressure changes linearly with changes in water vapor you don't have a problem with  
2 water vapor-dependent biases. I don't understand this assumption.

3 Added and changed in Section 3.2.3, page 12 line 20-22: "... the pressure sensor *has a non-*  
4 *linear dependence on is sensitive to* water vapor and thus, ... (Reum et al., 2017). A possible  
5 reason could be adsorption of water molecules on the sensor. To assess ... an incorrect pressure  
6 reading adjustment ..."

7 Added and changed in Section 3.2.3, page 12 line 25:

8 "Experiments with an additional independent pressure measurement presented by Reum et al.  
9 (2017), as well as analysis of the behavior of the proportional valve, which controls the pressure  
10 in the sample cell, show that if  $\Delta p$  is small enough it can be assumed that the pressure changes  
11 linearly with water vapor for mole fractions >2500 ppm (see Figure 2 in Reum et al., 2017). ~~and~~  
12 Therefore, the peak areas ... "

13 Changed in Section 3.2.3, page 13 line 3:

14 "~~Experiments with an additional independent pressure measurement (Reum et al., (2017), as~~  
15 ~~well as analysis of the behavior of the proportional valve, which controls the pressure in the~~  
16 ~~sample cell, show that Reum et al. (2017) determine  $\Delta p$  is around as about 0.7 mbar ... Note~~  
17 however, that the change in cell pressure with humidity is not fully-linear for water vapor mole  
18 fractions <2500 ppm (see Figure 2 in Reum et al., 2017), which could be the reason for the  
19 slightly systematic shape in the residuals at this low water vapor levels with lower values  
20 around 0.25 % water vapor mole fraction (Fig. 6 b).

- 21 25) Pg 12, L12-25: Don't the water vapor-dependent bias in the pressure readings also affect the  
22 in-flight data of the CRDS? Here you dwell on how this affects the CO<sub>2</sub> dilution calibration  
23 method, but don't discuss the effects on the actual in-flight measurements. If this also affects  
24 the flight data, why only discuss it here in the "Calibration by CO<sub>2</sub> Dilution" section?

25 For the in-flight measurements the error is removed, or rather corrected, by the calibration of  
26 the CRDS instrument with the dew point mirror. Moreover, the effect of the nonlinear part of  
27 the water dependent bias is small compared to the uncertainty range of the dew point mirror.  
28 The error due to the pressure deviation plays its role only during the dilution calibration, since  
29 here, the CO<sub>2</sub> measurements of the same instrument, which are affected by this error as well,  
30 are used for the calibration.

31 Added in Section 3.2.3, page 12 line 21:

32 "...While this error in the CRDS water vapor measurements is corrected by the calibration of the  
33 instrument with another hygrometer, it has to be considered for the dilution calibration, since  
34 the used CO<sub>2</sub> measurements are affected as well. To ..."

- 35 26) Pg 13, L26: It is strange to have a section entitled "Summary" in the middle of this paper. How  
36 about "Calibration Summary" instead?

37 Changed in Section 3.2.4, page 13 line 26: "3.2.4 Calibration Summary"

- 38 27) Pg 13, L30: "Note that both offsets, or rather their uncertainties, are likely not reliable." So, are  
39 the "offsets" (y-intercepts) likely not reliable, or their uncertainties, or both?

40 Changed in Section 3.2.4, page 13 line 30: "... both offsets, ~~or rather their~~ uncertainties, are ..."

- 41 28) Pg 14, L8-9: What is "a total accuracy" and what does "and down to 0.3 ppm for the lower  
42 mixing ratios" mean? How low is "lower"?

43 Changed in Section 4, page 14 line 8-9:

44 "... has an ~~total~~ accuracy of 6-8 % (1-sigma) in the range from 4 to 1000 ppm and ~~down to~~ 0.3  
45 ppm for lower mixing ratios down to 1 ppm (Meyer et al., 2015)."

- 46 29) Pg 14, L14-15: What are the "stable atmospheric conditions" mentioned here? This sentence  
47 needs a comma after "analysers".

48 Added in Section 4.1.1, page 14 line 14-15:

49 "analyzers, water ... atmospheric conditions, as pressure and temperature, were ..."

- 50 30) Pg 14, L13: Why use the term "repeatability" instead of precision? From the subsequent  
51 description it appears that you are looking at variations in the 0.4 Hz measurements from the  
52 60-second averages. I would call this "precision" and not repeatability, since repeatability can

1 imply stability in results obtained at different times, such as re-sampling the same air mass five  
2 minutes later and seeing how "repeatable" the measurements are.  
3 Changed in Section 4.1.1, page 14 line 13: "Measurement ~~repeatability-precision~~"  
4 Changed in Section 4.1.1, page 14 line 14: "... measurement ~~repeatability-precision~~ ..."  
5 Changed in Section 4.1.1, page 14 line 16: "... the ~~repeatability-precision~~ can be..."  
6 Changed in Section 4.1.1, page 14 line 26: "... the measurement ~~repeatability-precision~~ ..."  
7 Changed in Section 4.1.1, page 14 line 28: "For comparison ~~repeatability-precision~~ estimates ..."  
8 Changed in Section 4.1.1, page 15 line 1: "... a ~~repeatability-precision~~ of ..."  
9 Changed in Section 4.1.1, page 15 line 3: "... the ~~repeatability-precision~~ of the ..."  
10 Changed in Table 2, caption: "~~Repeatability-Precision~~ estimates ..."  
11 Changed in Table 2: "~~repeatability-precision~~ at ..."  
12 Changed in Section 3.2.1, page 10 line 10: "~~Repeatability-Precision of the CRDS analyzer~~ ..."  
13 31) Pg 14, L18-19: Isn't the 60-second average calculated from the 2.5-second measurements?  
14 Then why do you calculate the standard deviation of their differences when the standard  
15 deviation of the 60-second average directly provides a direct statistical measure of variability  
16 in the shorter-term measurements?  
17 The 60-second moving average is a measure for the long-periodic variability. Subtracting this  
18 long-periodic part of the signal from the measurements leaves the short-periodic variability,  
19 which is our measure for the short-term precision.  
20 32) Pg 14, L24: "results based on a larger number of data are highlighted". Aren't most 60-second  
21 averages based on 24 measurements made at 0.4 Hz?  
22 Only the flight measurement data from the four flights (total flight time: 14 h) are used for the  
23 analysis. Since not all water vapor ranges were represented equally during the flight  
24 measurements and, moreover, data periods with unstable cell pressure had to be neglected,  
25 the data availability in the different water vapor intervals varied significantly.  
26 33) Pg 14, L26-27: "upper limit" ... "4 ppm or 5%, whichever is greater for WV < 100 ppm". In  
27 Figure 7, why are there so many data points for WV < 80 ppm that have standard deviations >  
28 4 ppm (5%)? The upper limit estimate for WV > 100 ppm is much more in line with the Figure 7  
29 data: most of the data points for WV > 100 ppm lie below the 30 ppm standard deviation line  
30 and to the right of the 5% diagonal. This cannot be said about the WV < 100 ppm data, many of  
31 which lie above the 4 ppm standard deviation line and to the left of the 5% diagonal. Please  
32 explain the connection between this Figure and the numbers you have extracted from it that  
33 appear to underestimate the precision values for measurements at WV < 100 ppm.  
34 See answer to General Comment 1).  
35 34) Pg 15, L9-16: No need to describe the different symbols since that is done in both the Figure 8  
36 caption and legend. What is more important here is a description of what the Figure shows  
37 about response time of the CRDS. I think that showing only a short period of Figure 9 would  
38 better support this discussion of response times <10 seconds. Specifically, the flight segment  
39 from 12.35 to 12.55 UTC on 1 June 2011 when water vapor increases and falls an order of  
40 magnitude in about 3 minutes.  
41 See also answer to General Comment 5).  
42 Enlarged sections for selected shorter time periods are added in Figure 8 and 9.  
43 Added in Caption of Figure 8: "... flight on 31 May 2011. An enlarged section for the time period  
44 from 11:15 to 11:24 is shown in the lower left part of the plot."  
45 Added in Caption of Figure 9: "...flight on 1 June 2011. An enlarged section for the time period  
46 from 12:24:36 to 12:25:48 is shown in the upper left part of the plot. Here the water vapor  
47 increases in one minute about an order of magnitude from 200 to 1200 ppm."  
48 Added in Section 4.1.2, page 15 line 19: "... to wet. For water vapors <100 ppm the response  
49 time is comparable to the FISH instrument, as shown in the enlarged section of Fig. 8. During  
50 the increase of water vapor from 200 to 1200 ppm in about one minute, shown in the enlarged  
51 section in Figure 9, no significant delay can be detected. Thus, ..."

- 1 35) Pg 15, L22-25: This discussion of laboratory tests of response time is not very informative or  
2 conclusive because the reader has no idea how much water vapor is in the "wet and dry air".  
3 Also, "recovery time to 99% of a challenge" is not understandable.  
4 Added in Section 4.1.2, page 15 line 22: "... between wet (around 23000 ppm) and dry (around  
5 10 ppm) air, ..."  
6 Changed in Section 4.1.2, page 15 line 23: "... to 99 % of the final water vapor level ~~a challenge~~  
7 as ..."  
8 Changed in Section 4.1.2, page 15 line 24: "... from 23000 ppm (~~around 20°C dew point~~) to ..."  
9 Changed in Conclusions, page 18 line 7: "... to 99 % of the final water vapor level ~~a challenge~~:  
10 ..."  
11 36) Pg 15, L27-31 and Pg 16 L1-6: Please remove Figure 11 symbol descriptions and the  
12 explanation of data binning since all this information appears in the Figure caption.  
13 Descriptions and explanations were removed. See answer to General Comment 5.  
14 37) Pg 16, L1-2 and Figure 11 caption: "Absolute" differences steers my thinking towards absolute  
15 values, which are definitely not what Figure 10 shows. "Absolute" is unnecessary in both  
16 instances because the differences are clearly in units of ppm, i.e., not relative differences.  
17 Changed in Section 4.1.3, page 16 line 1: "The ~~absolute~~-water vapor differences ..."  
18 Changed in Figure 11, caption: "~~Absolute-d~~Differences ..."  
19 38) Pg 16, L4-6 and Figure 11: If the differences are binned by CR-2 water vapor in 1 ppm bins for  
20 CR-2 < 10 ppm, why are there multiple blue diamonds and multiple blue triangles in the lowest  
21 CR-2 mixing ratio bin (and some other bins)? Aren't the averages computed from all segments  
22 of all flights? If not, why are the average differences in each bin determined for subsets of the  
23 4 flights?  
24 Figure 11 was corrected. Now there is only one data point for each bin.  
25 39) Figure 11: The log-scale Y-axis is problematic for average differences < 1 ppm and for the error  
26 bars that appear asymmetric around the mean. Why don't you instead plot average relative  
27 (%) differences using a linear-scale Y axis? Average differences < 1 ppm will then be displayed  
28 correctly and the error bars will be symmetric around the mean values. I think the relative  
29 values will be much more like flat lines with CRDS-CR2 values above zero and the other  
30 instrument pairs below zero.  
31 We plotted the relative differences using a linear axis, too. The problem with this display is  
32 that for water vapors <100 ppm, and especially <10 ppm, the relative differences are very large  
33 (up to 50-100 %). For the interpretation of the results at <100 ppm a plot with differences in  
34 ppm is much clearer and easier to understand. So, both representations have their pros and  
35 cons. In the end we decided that the plot with the log-scales is a bit better suited for our  
36 purposes.  
37 40) Pg 16, L11-33: This section is difficult to understand because there are no descriptions of the  
38 various inlets for the different instruments. All that is known at this point is that the FISH inlet  
39 is different from the CRDS inlet, although there is now a hint that the other instruments do not  
40 share the CRDS inlet. Is it really important or necessary to include this section of the paper  
41 when it requires adding more details of the various instrument inlets? This section seems to  
42 cover a different topic from the main thrust of this paper and may even be worthy of a  
43 separate paper in the future.  
44 See answer to General Comment 6).

1 **Author's response to Interactive comment by Anonymous Referee #2 on**  
2 **"Evaluation of the IAGOS-Core GHG Package H<sub>2</sub>O measurements during**  
3 **the DENCHAR airborne inter-comparison campaign in 2011" by Annette**  
4 **Filges et al.**  
5

6 **About the paper**

7 This manuscript describes the calibration and intercomparison of a Picarro CRDS instrument for  
8 measurement of ambient water vapor. Multiple laboratory calibrations of the instrument by  
9 comparison with a NIST-traceable dewpoint hygrometer, the FZ Jülich calibration setup, and a new  
10 method using co-measured CO<sub>2</sub> dilution by the added H<sub>2</sub>O are described to assess the instrument  
11 accuracy and stability over time (up to years). Comparison of in-flight measurements with several  
12 research hygrometers are presented to assess the instrument performance. The CRDS instrument  
13 described is the model that has been selected for the IAGOS-Core payload for regular in-service  
14 measurements, and therefore the assessment provided will have bearing on a significant  
15 atmospheric data set. The manuscript topic and material presented are appropriate for publication  
16 in AMT. The manuscript would benefit from some changes to the analysis, structure and  
17 presentation.  
18

19 **General Comments**

- 20 1) The manuscript would benefit substantially from copyediting to correct usage and punctuation.  
21 Examples include: the use of commas after introductory prepositional phrases,  
22 [Some proof reading was carried out.](#)
- 23 2) The inclusion of water vapor mixing ratios in both ppm (ppmv) and % units is somewhat  
24 awkward, although I understand that those units are regularly used for water in different  
25 environments. Suggest consistently using ppmv for the manuscript with an initial parenthetical  
26 equivalence for the reader, e.g. on line 30 "7000 ppmv (0.7%) to 25,000 ppmv (2.5%)". This  
27 would have the additional benefit of simplifying the discussion of the uncertainty in % of  
28 measurement without needing to specify "(rel.)". The DENCHAR website uses ppmv even for  
29 high values mentioned.  
30 [Mole fractions in unit "%" were changed to "ppm" in the whole manuscript \(text, tables, and](#)  
31 [figures\). Ratios in "% \(relative\)" were changed to "%".](#)
- 32 3) The details of representations (e.g. blue line, dark blue triangle) of data in the plots should not  
33 be included in the text, only in the figure captions (where they should be fully described for  
34 clarity). The text should discuss the interpretation and implications of the data presented in the  
35 figures. The figure captions could be more descriptive.  
36 [Text and figure captions were changed accordingly. See also answer to General Comment 5\) of](#)  
37 [referee #1.](#)
- 38 4) The term "repeatability" is used in a number of places in the manuscript (e.g. section 4.1.1) that  
39 would appear to be better described as measurement precision. Repeatability is typically the  
40 measure of how consistently an instrument will produce the same response when challenged at  
41 different times with the same input related to drift. Precision is more typically used for the  
42 short time variation of measurement (as is the case here for the variation about the 60 s or 30 s  
43 mean value) or repeated consecutive measurements of a constant input.  
44 ["Repeatability" was replaced by "precision". See answer to Specific comment 30\) of referee #1.](#)
- 45 5) Some pieces of information are repeated numerous times within the manuscript out of a desire  
46 to be clear, but it seems to go too far. For example, it should be stated at the first reasonable

1 opportunity in the manuscript that 0.802 calibration factor from 2013 will be used going forward  
2 and then it need not be repeated every time.

3 "~~Changed in Section 3.2.2, page 11 line 1-2: " ... 2013 dew point mirror comparison (calibration~~  
4 ~~factor = 0.802 ± 0.010), ..."~~

5 "~~Changed in Section 3.2.3, page 11 line 25: " ... 2013 dew point mirror comparison (calibration~~  
6 ~~factor = 0.802 ± 0.010), ..."~~

7 "~~Changed in Section 4.1.1, page 14 line 16-17: " ... (according to the 2013 dew point mirror~~  
8 ~~comparison, calibration factor = 0.802 ± 0.010; in the following simply referred to as CRDS~~  
9 ~~measured water vapor) ..."~~

10 "~~Changed in Section 4.1.2, page 15 line 7-8: "In addition to the offset corrected and calibrated~~  
11 ~~(according to the 2013 dew point mirror comparison, calibration factor = 0.802 ± 0.010) water~~  
12 ~~vapor measurements of the CRDS analyzer, in the following simply referred to as CRDS measured~~  
13 ~~water vapor..."~~

#### 14 15 **Specific Comments:**

16 1) P1, L16: "North-Germany" should be "northern Germany"; similarly elsewhere, "South- Norway"  
17 and "North-Poland" should be "southern Norway" and "northern Poland"

18 ~~Changed in the whole manuscript. See answer to Specific Comment 1) of referee #1.~~

19 2) P1, L23: This paragraph seems inappropriate for the abstract—suggest removing and including  
20 the details in the instrument description section.

21 ~~Lines 23-28 in the Abstract ("The inlet system, .... within IAGOS") were deleted.~~

22 ~~Added in Section 2, page 6 line 11: "operation of the instrument with a controlled pressure in the~~  
23 ~~sample cell of 186.65 hPa (140 Torr) throughout the aircraft altitude operating range..."~~

24 3) P1, L30: "instruments" should be capitalized.

25 ~~Changed in Abstract, page 1 line 30: "...Michell #instruments Ltd. ..."~~

26 4) P1, L31+: "on the ground over the range from 2 to 600 ppmv against the dew point hygrometer  
27 used for calibration of the FISH instrument. A new, independent calibration method based on  
28 the dilution effect of water vapor on CO<sub>2</sub> was evaluated."

29 ~~Changed in Abstract, page 1 line 31-33: "...compared on the ground over the range from 2 to 600~~  
30 ~~ppm against the dew point hygrometer, which is used for calibration of the reference~~  
31 ~~instrument FISH, in the range from 2 to 600 ppm. Furthermore, a new, independent calibration~~  
32 ~~method, based on the dilution effect of water vapor on CO<sub>2</sub>, was evaluated tested."~~

33 5) P2, L3: "than" should be "then", but even better could just be deleted—"1% for the water vapor  
34 range from 25,000 ppmv down to 7000 ppmv, increasing to 5% at 50 ppm"

35 ~~Changed in Abstract, page 2 line 3: "... down to 7000 ppm, than increasing to ..."~~

36 6) P3, L6: sufficient quality for what?

37 ~~Changed in Introduction, page 3 line 5-8:~~

38 ~~"Since there are only sparse measurement data of sufficient quality. However, due to only few~~  
39 ~~existing measurement data in the UTLS, and ..."~~

40 7) P3, L7: "difficulty representing"; overall this sentence is awkward and difficult to parse.

41 ~~Changed in Introduction, page 3 line 5-8:~~

42 ~~"Since there are only sparse measurement data of sufficient quality. However, due to only few~~  
43 ~~existing measurement data in the UTLS, and, as emphasized in Solomon et al. (2010), limitations~~  
44 ~~in prognostic models simulations of have difficulties to represent this region well (Solomon et al.,~~  
45 ~~2010), uncertainties in chemistry, transport processes, and trace gas composition are relatively~~  
46 ~~large., which This influences ..."~~

47 8) P3, L18: for what purpose do the satellite measurements have insufficient spatial resolution?  
48 Horizontally? Vertically?

49 ~~See answer to Specific Comment 5) of referee #1.~~

50 9) P3, L32: What is meant by "beneath"?

51 ~~Changed in Introduction, page 3 line 32: "Beneath Besides frost point hygrometers ..."~~

- 1 10) P5, L28: “independent on ambient, respectively cabin pressure” is confusing. “independent of  
2 both ambient and cabin pressure”?  
3 *Changed in Section 2, page 5 line 28: “... ambient, ~~respectively and~~ cabin pressure ...”*
- 4 11) P5, L29: the filters in the sample line are likely contributing significantly to the long tail in the  
5 time constant.  
6 *The physical exchange time of the sample cell is only 3.6 s (volume = 35 ml, sample flow = 100  
7 ml/min, pressure = 186.65 hPa, sample temperature = 45°C). Surface effects on the walls of the  
8 inlet line and sample cell, as well as the filters can contribute to a longer time constant,  
9 especially when the change in water vapor is large. However, comparison of all instruments in  
10 Section 4.1.2 shows that the response time of the CRDS is similar to that of the other  
11 instruments.*
- 12 12) P7, L19: the use of  $\sim 125$  ppm as a metric for stability over a range that includes 3 ppm does not  
13 induce confidence—would be better to use a relative measure.  
14 *Changed in Section 3.1.2, page 7 line 19: “... ranging from around ~~3-5000 to 30000~~ ppm.”*  
15 *After correcting the offsets of all analyzers we expect the relative differences for water vapor  
16 <5000 ppm to also be around some percent.*
- 17 13) P8, L15: “the permeation is negligible” for 3500 sccm is technically incorrect since the  
18 permeation rate through the tubing is largely flow independent. The contribution of the  
19 permeation to the water vapor mixing ratio in the flow might be argued to be negligible.  
20 *Added in section 3.1.3, page 8 line 15: “... the *contribution of the permeation to the water vapor*  
21 *mole fraction in the flow is negligible ...”**
- 22 14) P9, L5: “gets only changed by” → “is only affected by”  
23 *Changed in Section 3.1.4, page 9 line 5: “... the peak area ~~gets only changed by~~ due to the  
24 dilution effect. ...”*
- 25 15) P9, L13: “whereby it required about” → “requiring”  
26 *Changed in Section 3.1.4, page 9 line 13: “...line, ~~whereby it required~~ requiring about 1.3 s ...”*
- 27 16) P9, L29: it would be helpful to the reader to convert the °C uncertainty of the Dewmet to mixing  
28 ratio (using the relevant operating pressure) for comparison with other uncertainties.  
29 *Added in Section 3.2.1, page 9 line 30: “... dew point (2-sigma), which corresponds to a relative  
30 uncertainty of 1.3 %.”*
- 31 17) P11, L11: the CRDS and FISH cal system can’t agree within 3% at mixing ratios below 400 ppm  
32 without first correcting for the 12.2 ppm offset. It seems reasonable that there was either an  
33 outgassing or small leak (it doesn’t take much) that led to the offset, but that should be argued  
34 and removed before stating the agreement in the response (slope).  
35 *Added in Section 3.2.2, page 11 line 11: “... up to 600 ppm after correcting for an offset of 12  
36 ppm. ...”*  
37 *Added in Section 3.2.2, page 11 line 21: “...than 1 ppm. Another possibility would be that the  
38 offset was caused by either an outgassing or a very small leak.”*
- 39 18) P12, L3: “likely unrealistic”—should be more descriptive; “the calculated error of 1.6 ppmv for  
40 the offset is likely a significant underestimate.”  
41 *Changed in Section 3.2.3, page 12 line 3: “...is likely a significant underestimate ~~unrealistic.~~”*
- 42 19) P12, L20: The section is unclear—are you saying that the pressure sensor response to H<sub>2</sub>O is  
43 different from that to other molecules, leading to an incorrect measurement of the wet flow cell  
44 pressure that depends on the H<sub>2</sub>O content?  
45 *See answer to Specific Comment 24) of referee #1.*
- 46 20) P13, L4: It would be better to state a best estimate of the CO<sub>2</sub> dilution method uncertainty  
47 rather than saying that a value “is achievable”. What do you estimate that you actually achieved,  
48 and is that consistent with the comparisons?  
49 *The conservative estimate for the presented experiment is 1 % (see page 13, line 5). For future  
50 experiments uncertainties at sub-percent level are achievable.*  
51 *Added in Section 3.2.3, page 13 line 4: “... is achievable for the dilution method in future  
52 experiments.”*

- 1 21) P14, L2: "Hohn, Germany, the flights"  
2 Changed in Section 4, page 14 line 2: "... Hohn, ~~Germany~~, the flights ..."
- 3 22) P14, L5: "Hence, also the lower stratosphere was reached." is an awkward construction, this  
4 could be better included within the previous sentence.  
5 Changed in Section 4, page 14 line 5: "... up to 12.5 km-~~Hence, so that also the lower ...~~"
- 6 23) P14, L10: "total water instead of only water vapor since the forward-oriented inlet resulted in  
7 sampling of cloud droplets and ice crystals when present." The sentence "Both instruments  
8 were..." seems unnecessary.  
9 Changed in Section 4, page 14 line 10: "...instead of *only* water vapor ..., since ~~the-it's forward-~~  
10 ~~oriented inlet system of FISH allowed for~~ resulted in sampling of cloud droplets and ice crystals  
11 ~~when present. Both instruments were operated by the research centre Jülich."~~
- 12 24) P14, L14: "CRDS analyzer, water vapor measurements during periods with stable"  
13 Changed in Section 4.1.1, page 14 line 14: "... CRDS analyzers, water vapor ..."
- 14 25) P14, L26: the value of 4 ppm/5% stated for the <100 ppmv precision seems to be at odds with  
15 the plotted standard deviations in the 10-100 ppmv range, which are typically higher than both.  
16 See answer to General Comment 1) of referee #1.
- 17 26) P15, L9: "and the flight data..." → "along with the flight data from the reference instruments  
18 CR-2 and FISH. Water vapor measurements from two additional analyzers that participated in  
19 the intercomparison campaign are also shown."  
20 Changed in Section 4.1.2, page 15 line 8-11:  
21 "CRDS measured water vapor *is shown along with and* the flight data *of* from the reference  
22 instruments CR-2 and FISH; ~~also~~ Water vapor measurements *from of* two additional analyzers;  
23 ~~which took that participated in the inter-comparison campaign;~~ *are also presented:- ...*"
- 24 27) P15, L17: "not slower than any other instrument." → "is similar to that of the other  
25 instruments."  
26 Changed in Section 4.1.2, page 15 line 17: "...~~is not slower than any~~ similar to that of the other  
27 instruments."
- 28 28) P15, L30: "CRDS data from the first flight" and "influenced" → "compromised"  
29 Changed in Section 4.1.3, page 15 line 30: "... CRDS data ~~of~~ from the first flight ... were *influenced*  
30 *compromised by ...*"
- 31 29) P16, L1: "can be seen" → "are shown" or "are presented" or "are plotted"  
32 Changed in Section 4.1.3, page 16 line 1: "... water vapor intervals ~~can be seen~~ *are plotted in ...*"
- 33 30) P16, L7: "are neglected." → "are omitted."  
34 Changed in Section 4.1.3, page 16 line 7: "...are ~~neglected~~ *omitted. ...*"
- 35 31) P16, L11: "sampled data only for two flights" → "operated successfully for only two of the four  
36 flights"  
37 Changed in Section 4.1.3, page 16 line 11: "... FISH *operated successfully for only two of the four*  
38 *flights sampled data only for two flights, ...*"
- 39 32) P16, L20: "This might indicate that also the CRDS samples cloud" → "This indicates that the  
40 CRDS sampling was likely also affected by cloud particles" Looking at the Rosemount inlet, there  
41 is a good likelihood that impaction of droplets or ice crystals (shattering) will produce particles  
42 small enough to be entrained in the sample flow and ingested into the instrument. The ice  
43 crystal fragments would then be small enough to sublime in the inlet and affect the  
44 measurement.  
45 Changed in Section 4.1.3, page 16 line 20-21:  
46 "~~This might indicate that also the CRDS samples cloud particles~~ *This indicates that the CRDS*  
47 *sampling was likely also affected by cloud particles, ...*"  
48 Added in Section 4.1.3, page 16 line 26: "*Such small enough particles could be produced e.g. by*  
49 *the shattering of water droplets or ice crystals in the Rosemount housing."*
- 50 33) P18, L1: It is unclear exactly what is meant by the statement "The dilution method can be used  
51 for other species...". The dilution method could be used to determined H2O using methane

1 instead of CO<sub>2</sub>, or are you trying to say that methane could be measured by the method by  
2 comparing to CO<sub>2</sub>?

3 Both options are meant. H<sub>2</sub>O can be calibrated using another species instead of CO<sub>2</sub>, and other  
4 species instead of H<sub>2</sub>O can be also calibrated by CO<sub>2</sub> or any another species, provided both used  
5 species are measurable by a CRDS analyzer and the dilution effect is large enough.

6 Changed in Conclusions, page 18 line 1: "... used for *the calibration of* other species, too,  
7 provided they *and the corresponding diluted species* are measurable by a CRDS analyzer and the  
8 ~~corresponding~~ dilution effect ..."

# Evaluation of the IAGOS-Core GHG Package H<sub>2</sub>O measurements during the DENCHAR airborne inter-comparison campaign in 2011

Annette Filges<sup>1</sup>, Christoph Gerbig<sup>1</sup>, Chris W. Rella<sup>2</sup>, John Hoffnagle<sup>2</sup>, Herman Smit<sup>3</sup>, Martina Krämer<sup>4</sup>, Nicole Spelten<sup>4</sup>, Christian Rolf<sup>4</sup>, Zoltán Bozóki<sup>5</sup>, Bernhard Buchholz<sup>6</sup>, Volker Ebert<sup>6</sup>

5 <sup>1</sup>Max Planck Institute for Biogeochemistry (MPI-BGC), Jena, 07747, Germany

<sup>2</sup>Picarro, Inc., Santa Clara, CA 95054, USA

<sup>3</sup>Research Centre Jülich, Institute for Energy and Climate Research Troposphere (IEK-8), Jülich, 52428, Germany

<sup>4</sup>Research Centre Jülich, Institute for Energy and Climate Research Stratosphere (IEK-7), Jülich, 52428, Germany

10 <sup>5</sup>Hungarian Academy of Sciences (MTA) and University of Szeged (SZTE) Research Group on Photoacoustic Spectroscopy, Szeged, H-6720, Hungary

<sup>6</sup>Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, 38116, Germany

*Correspondence to:* Annette Filges (annette.filges@bgc-jena.mpg.de)

15 **Abstract.** As part of the DENCHAR (Development and Evaluation of Novel Compact Hygrometer for Airborne Research) inter-comparison campaign in ~~North-northern~~ Germany in 2011, a commercial cavity ring-down spectroscopy (CRDS) based gas analyzer (G2401-m, Picarro Inc., US) was installed on a Learjet to measure atmospheric water vapor, CO<sub>2</sub>, CH<sub>4</sub>, and CO. The CRDS components were identical to those chosen for integration aboard commercial airliners within the IAGOS (In-service Aircraft for a Global Observing System) project. ~~Since the quantitative capabilities of the CRDS water vapor measurements were never evaluated and reviewed in detail in a publication before~~ Thus, the campaign allowed for an initial assessment-validation of the long-term IAGOS water vapor measurements by CRDS against reference instruments with a long performance record (Fast In-situ Stratospheric Hygrometer (FISH); and CR-2 frost point hygrometer (Buck Research Instruments L.L.C., US), both operated by research centre Juejlich).

20 ~~The inlet system, a 50 cm long 1/8" FEP tube connected to a Rosemount TAT housing (model 102BX, deiced) installed on a window plate of the aircraft, was designed to eliminate sampling of larger aerosols, ice particles, and water droplets, and provided additional ram pressure. In combination with a low sample flow of 100 sccm, corresponding to a 4 second response time, this ensured a fully controlled pressure in the sample cell of 186.65 hPa (140 Torr) throughout the aircraft altitude operating range up to 12.5 km without the need of an upstream sampling pump. This setup ensures full compatibility with the deployment of the analyzer within IAGOS.~~

30 For the initial water calibration of the instrument it was compared against a dew point mirror (Dewmet TDH, Michell Instruments Ltd., UK) in the range from ~~0.70000~~ to 2.5000 ppm% water vapor mole fraction. During the inter-comparison campaign the analyzer was compared on ~~the ground~~ over the range from 2 to 600 ppm against ~~thea~~ dew point hygrometer; ~~which is~~ used for calibration~~ng of~~ the reference instrument FISH, ~~in the range from 2 to 600 ppm. Furthermore, aA new,~~ independent calibration method; based on the dilution effect of water vapor on CO<sub>2</sub>; was ~~evaluated~~tested.

Comparison of the in-flight data against the reference instruments showed that the analyzer is reliable and has a good long-term stability. The flight data suggest a conservative precision estimate for measurements made at 0.4 Hz (2.5 seconds measurement interval) of 4 ppm ~~or 5 % (relative) (whichever is greater)~~ for H<sub>2</sub>O <-100 ppm, 20 % or 10 ppm (whichever is smaller) for 10 ppm < H<sub>2</sub>O <100 ppm. and 5 % ~~(relative)~~ or 30 ppm (whichever is smaller) for H<sub>2</sub>O >-100 ppm. Accuracy of the CRDS instrument was estimated, based on laboratory calibrations, as 1 % ~~(relative)~~ for the water vapor range from 2-5000 ppm% down to ~~0.7000 ppm-%, than~~ increasing to 5 % ~~(relative)~~ at 50 ppm water vapor. Accuracy at water vapor mole fractions below 50 ppm was difficult to assess, as the reference systems suffered from lack of data availability.

## 1 Introduction

Water vapor is a crucial factor for various atmospheric processes, weather, and climate. It is the most important greenhouse gas (Kiehl and Trenberth, 1997) and shows strong feedback to changes in the climate system (Dessler, 2008). Especially in the upper troposphere and lower stratosphere (UTLS) the amount of water vapor has large impact on the radiative balance of the atmosphere (e.g. Smith et al., 2001; Forster and Shine, 2002; Solomon et al., 2010). ~~Since there are only sparse measurement data of sufficient quality. However, due to only few existing measurement data~~ in the UTLS, and, ~~as emphasized in Solomon et al. (2010), limitations in prognostic models simulations of have difficulties to represent~~ this region ~~well~~(Solomon et al., 2010), uncertainties in chemistry, transport processes, and trace gas composition are relatively large, ~~which~~ This influences significantly the estimation of e.g. radiative effects (Riese et al., 2012; Kunz et al., 2013).

Water vapor observations covering the whole troposphere and at least lower parts of the stratosphere are achieved mainly by instruments based on balloons, aircraft or satellites, and from ground-based remote sensing techniques. The longest measurement time series was started in 1980 in Boulder (Colorado, US) with balloon-borne frost point hygrometers (Oltmans et al., 2000; Hurst et al., 2011). First long-term global satellite data were obtained in the mid-1980s as part of the Stratospheric Aerosol and Gas Experiment II (SAGE II) (Rind et al., 1993), ~~followed by Recent observations are made by e.g. from e.g. the Halogen Occultation Experiment (HALOE) (Harries et al., 1996), the Aura Microwave Limb Sounder (MLS) (Read et al., 2007; Lambert et al., 2007), as well as~~ the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) (Milz et al., 2005; von Clarmann et al., 2009) and the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) (Rozanov et al., 2011; Weigel et al., 2015), both aboard ENVISAT (Environmental Satellite). The main drawbacks of satellite data and remote sensing observations from ground (e.g. Schneider et al., 2006) are their insufficient ~~spatial-vertical~~ resolution in the troposphere and lower stratosphere and disturbances of the measurements by clouds. ~~As shown by Hoareau et al. (2013), vertical resolutions <500 m are needed for the simulation of cirrus clouds. To represent the very sharp gradient of 40 to 6 ppm water vapor within 0-2 km at the tropopause (Zahn et al., 2014), resolutions of even 400 m and higher have to be achieved (Poshyvailo et al., 2018).~~ On the other hand reliable radiosonde water vapor data up to stratospheric heights, e.g. from the GCOS (Global Climate Observing System) Reference Upper-Air Network (GRUAN) (Dirksen et al., 2014), as well as data sets from research aircraft are quite limited in time and space.

The use of commercial aircraft as cost-efficient platforms for dedicated instruments can at least partially bridge this gap providing regular measurements in the UTLS along major flight routes. For example, five Airbus A340 passenger aircraft were equipped with capacitive humidity sensors from 1994-2014 as part of the MOZAIC (Measurement of Ozone and Water Vapor by Airbus In-Service Aircraft) project (Marenco et al., 1998). The acquired data set is crucial for the study of chemical and dynamic processes in the upper troposphere and lower stratosphere (e.g. Gierens et al., 1999).

However, accurate and reliable airborne measurements of atmospheric water vapor are still a challenge. The large range from mole fractions of several percent at the ground to only a few parts per million ( $\text{ppm} = \mu\text{mol} \cdot \text{mol}^{-1}$ ) in the stratosphere and

the highly variable structures of water vapor in the atmosphere are demanding for analyzers regarding accuracy and time response.

Kley et al. (2000) gives a detailed summary of the most important water vapor instruments used onboard aircraft. ~~Besides~~neath frost point hygrometers (e.g. Vömel et al., 2007; Vömel et al., 2016; Hurst et al., 2011; Hall et al., 2016) these are mainly Lyman- $\alpha$  hygrometers, based on fluorescence techniques, for example the Harvard Water Vapor instrument (HWV) (Weinstock et al., 2009) and the Fast In-~~S~~situ Stratospheric Hygrometer (FISH) (Zöger et al., 1999; Meyer et al., 2015). More recently infrared absorption spectrometers like the Jet Propulsion Laboratory Laser Hygrometer (JLH) (May, 1998), the Integrated Cavity output Spectrometer (ICOS) (Sayres et al., 2009), or the Hygrometer for Atmospheric Investigations (HAI) (Buchholz et al., 2017), and the Atmospheric Ionization Mass Spectrometer (AIMS) (Kaufmann et al., 2016) have been deployed. The central problematic of all these different types of analyzers are the ~~remaining~~, unexplained discrepancies between water vapor measurements in the range below 10 ppm (e.g. Kley et al., 2000; Vömel et al., 2007; Weinstock et al., 2009). While the instruments compare well during static experiments (Fahey et al., 2014), they disagree significantly during airborne inter comparisons in the UTLS. For the recent Mid-latitude Airborne Cirrus Properties Experiment (MACPEX) in 2011 Rollins et al. (2014) estimated the differences to be of the order of 20 % at water vapor mixing ratios of 3-4 ppm, whereas the measurement uncertainties of the instruments account only for 5-10 %. Thus, possibilities e.g. to identify long-term trends in stratospheric water vapor or to study ice microphysical processes are limited (Rollins et al., 2014).

In this context the DENCHAR (Development and Evaluation of Novel Compact Hygrometer for Airborne Research) project was initiated by the European Facility for Airborne Research (EUFAR) to support the development and characterization of novel or improved compact airborne hygrometers for different airborne applications within EUFAR, including investigation of the sampling characteristics of different gas/ice inlets (cf. Tátrai et al., 2015). As part of the DENCHAR inter-comparison flight campaign in Hohn (Germany) in May-June 2011, a commercial cavity ring-down spectroscopy (CRDS) gas analyzer (G2401-m, Picarro Inc.,\_US), measuring CO<sub>2</sub>, CH<sub>4</sub>, CO and water vapor, was tested and compared against well-established reference hygrometers and newly developed water vapor instruments. ~~Four flights with a Learjet 35A took place in an area between North Germany and South Norway, and North Poland and the North Sea respectively, reaching altitudes up to 12.5 km, hence covering also the lower stratosphere. As reference instruments served~~The former were the Learjet version of the Fast In-situ Stratospheric Hygrometer (FISH) (Meyer et al., 2015) and a CR-2 Cryogenic Aircraft frost point-hygrometer (Buck Research Instruments L.L.C., Boulder, US, [www.hygrometers.com](http://www.hygrometers.com)), both operated by the research centre Jülich.

The same CRDS analyzer and corresponding inlet system components are scheduled for deployment onboard passenger aircraft within the IAGOS (In-service Aircraft for a Global observing System) project (Filges et al., 2015). IAGOS was launched in 2005 as the successor program of MOZAIC, but with modernized instrumentation and enhanced measurement capabilities (Volz-Thomas et al., 2009; Petzold et al., 2015). The current fleet of IAGOS-equipped aircraft as well as the spatial coverage of all flights can be found at the IAGOS-database ([www.iagos.org](http://www.iagos.org)). It is planned to equip five IAGOS aircraft with the CRDS system, as “IAGOS-core Greenhouse Gas (GHG) package”, in the next four years. In contrast to the

CO<sub>2</sub> measurements from the CRDS, which have been studied thoroughly and have shown good performances, the quantitative capabilities of the CRDS water vapor measurements were never evaluated and reviewed in detail before. Precision in the laboratory is known from previous studies to be around 6 ppm for a 2.3 second integration time, but is related to white noise (Filges et al., 2015). Thus, sample averaging over 30 minutes can result in a precision of down to 0.3 ppm, which in principle can result in numerous scientific applications of the data. Each IAGOS aircraft is also equipped with  
5 ~~Since the MOZAIC humidity device (Helten et al., 1998, Smit et al., 2008, Smit et al., 2013) is deployed on each IAGOS aircraft, this combination, which~~ provides the unique opportunity to compare both instrument types under the same conditions over a long-time period. IAGOS water vapor measurements include regular in situ data in the sensible UTLS region and vertical profiles of H<sub>2</sub>O in the troposphere and lower stratosphere for major parts of the globe. They are essential  
10 for validation of remote sensing based observations from satellites and ground, are needed for improving the performance of climate models and weather forecasts, and can be used for climate trend studies.

This paper presents the water vapor measurements made with the CRDS system during the DENCHAR inter-comparison flight campaign in 2011. The flight data are validated against reference instruments with a long performance record (~~FISH (Meyer et al., 2015) and CR 2 Cryogenic Aircraft Hygrometer (Buck Research Instruments L.L.C., US, www hygrometers.com)~~)  
15 ~~to evaluate the water vapor measurements made by the CRDS instrument. Furthermore, the analyzer was calibrated with the help of different hygrometers, and a newovel independent calibration method was tested. The corresponding results are analyzed and discussed regarding the feasibility of the different methods for the long-term operation of the analyzer within the IAGOS project.~~

The measurement system is introduced in Sect. 2, followed by an overview of the water vapor calibration approaches in Sect.  
20 3. Results from the flight tests, ~~including comparison with the reference instruments,~~ are presented in Sect. 4. Section 5 concludes the paper.

## 2 The measurement system

The measurements were conducted by a G2401-m wavelength scanned cavity ring-down spectroscopy analyzer from Picarro Inc. (US) (CFKB2004), which simultaneously measures CO<sub>2</sub>, CH<sub>4</sub>, CO and water vapor (Crosson, 2008; Chen et al., 2010).

25 The CRDS technique determines the mole fraction of a gas using the decay time of light intensity (“ring-down time”) due to absorption by the gas. Laser light of a specific set of wavelengths is injected into a mirrored sample cell (the “cavity”, 35 cm<sup>3</sup>, effective optical path length 15-20 km), which is flushed with the sample gas. When the light intensity reaches a predetermined threshold, the laser is turned off, after which the optical energy in the cavity decays with a characteristic exponential time constant of the light intensity in the cavity (the ringdown). The total absorption of the cavity (including  
30 both the absorption of the gas and the loss of the mirrors) is calculated directly from the exponential time constant. By tuning the wavelength of the laser, a specific spectral line of a species is scanned and analysis of the obtained spectrogram provides the peak height, which at constant pressure and temperature is proportional to the mole fraction of the species.

The analyzer uses selected spectral lines in the infrared for the measurements: at 1603 nm for  $^{12}\text{C}^{16}\text{O}_2$ , at 1651 nm for  $^{12}\text{CH}_4$  and  $\text{H}_2^{16}\text{O}$ , and at 1567 nm for  $^{12}\text{C}^{16}\text{O}$ .

To minimize the impact of pressure and temperature on gas density and spectroscopy, pressure and temperature both are kept constant in the sample cell ~~are kept constant~~. Pressure in the sample cell is controlled to  $186.65 \pm 0.04$  hPa (140 Torr) using a proportional valve (“inlet valve”) upstream of the cell, and the temperature is kept at  $45 \pm 0.02$  °C. Gas flow through the sample cell was controlled at 100 sccm with the help of a fixed needle valve, acting as flow-restricting orifice, downstream of the sample cell and upstream of the pump. Thus, the flow rate was independent on ambient and, respectively cabin pressure.

To protect the sample cell from contamination, two filters (Wafergard II F Micro In-Line Gas Filters, Entegris Inc.) are located in the sample line upstream of the sample cell. They also ensure thermal equilibration of the sample gas, as they are kept at the same temperature as the sample cell.

Each species was measured once every 2.5 seconds. The physical exchange time of the sample cell is 3.6 seconds (volume =  $35 \text{ cm}^3$ , sample flow = 100 sccm, pressure = 186.65 hPa, sample temperature = 45 °C), ensuring that the ambient air was continuously sampled given the shorter measurement interval of 2.5 seconds.

The instrument was equipped with a 50 cm long inlet line (3.18 mm (1/8”) OD, 1.58 mm (1/16”) ID, Fluorinated Ethylene Propylene (FEP) tube), which was connected to a Rosemount Total Air Temperature (TAT) housing (model 102B; Stickney et al., 1994) mounted on a window plate of the Learjet. The Rosemount probe acts as a virtual impactor since the inlet line is pointed orthogonal to the airflow through the housing (see Fig. 1), and thus prevents from sampling larger aerosols (larger than about  $2 \mu\text{m}$ ), ice particles, and water droplets (Volz-Thomas et al., 2005; Fahey et al., 2001, Smit et al., 2013).

Furthermore, it provides positive ram-pressure due to the reduction of the air velocity. This additional positive ram-pressure, together with the low sample gas flow of 100 sccm and the relatively short inlet line, ensured operation of the instrument with a controlled pressure in the sample cell of 186.65 hPa (140 Torr) throughout the aircraft altitude operating range up to 12.5 km without an upstream sampling pump. Diffusion effects of water vapor in the inlet line are minimal, given the short residence time of the sample gas, the small inner surface area, the small differences in humidity between the air conditioned cabin and the ambient air, and the low permeability of FEP. The sample flow was exhausted into the cabin of the aircraft.

The CRDS analyzer and the inlet system components are identical to those chosen for integration aboard commercial airliners within the IAGOS (In-service Aircraft for a Global observing System) project. This setup ensures full comparability with the deployment of the IAGOS-core Greenhouse Gas package (Filges et al., 2015).

### 3 Calibration

In contrast to calibration of the  $\text{CO}_2$ ,  $\text{CH}_4$  and  $\text{CO}$  measurements, for which traceability to the World Meteorological Organization (WMO) primary scales is ensured by measurement of gas standards traceable to the primary scale, calibration of the water vapor measurements of the instrument is not as straight forward. There is no globally valid primary scale, but several national standards exist (WMO, 2012, see Part I, Chapter 4). Calibration of an instrument is done by means of other

hygrometers that are traceable to one of the national standards, often gravimetric hygrometers. In the following different calibration methods for the CRDS analyzer are presented and compared.

### 3.1 Methods

#### 3.1.1 Offset ~~c~~Correction

5 Prior to calibration of the CRDS analyzer against e.g. a dew point mirror an offset correction is required once to improve measurements at low water levels (~~<0.1000 ppm%~~). This offset correction can be estimated by measuring dried ambient air from a high-pressure tank. At the MPI-BGC GasLab the tanks (volume: 50 liters) were filled with air, which was dried using magnesium perchlorate ( $Mg(ClO_4)_2$ ). The dew point of the air is around  $-75\text{ }^\circ\text{C}$  corresponding to 2.4 ppm water vapor. The offset stability of different CRDS instruments was checked regularly over a time period of up to 10 years and no  
10 significant drift was observed.

#### 3.1.2 Dew Point Mirror

The factory calibration of the Picarro Inc. CRDS analyzer consists of two parts: A self-broadening correction and a comparison with a dew point mirror.

15 Water vapor mole fraction is calculated using the peak height of the selected water absorption line. In this process self-broadening effects must be taken into account, which broaden the line shape and hence decrease the peak height, as the water vapor level increases. To avoid an underestimation of the water vapor mole fraction, a quadratic correction is implemented in the Picarro analyzer (Rella, 2010):

$$H_2O_{corrected} = H_2O_{reported} + 0.02525 * H_2O_{reported}^2 \quad (1)$$

20 Here,  $H_2O$  is water vapor mole fraction in %. In 2009 a G1301-m CRDS instrument from Picarro, measuring  $CO_2$ ,  $CH_4$  and  $H_2O$ , was calibrated at MPI-BGC Jena against a dew point mirror (Dewmet TDH, Cooled Mirror Dewpointmeter, Michell Instruments Ltd., UK, referenced to National Institute of Standards and Technology (NIST) primary scale) in the range from 0.7000 to 30000.0 ppm% water vapor mole fraction (Winderlich et al., 2010). The calibration constant obtained in this experiment was transferred to all greenhouse gas CRDS instruments manufactured by Picarro Inc. (Rella, 2010):

$$H_2O_{calibrated} = 0.772 * H_2O_{corrected} = 0.772 * (H_2O_{reported} + 0.02525 * H_2O_{reported}^2) \quad (2)$$

25 This calibration transfer from one to all other instruments is possible since their water vapor measurements agree within a sufficient range and are stable over time as shown by Rella et al. (2013). Here, three different analyzers (models G2401-m and EnviroSense 3000) were compared at different times against one selected standard instrument (CFADS37, model G1301-m). One of the comparisons was repeated after more than three years. All results ( $H_2O_{analyzer} - H_2O_{CFADS-37}$ ) lie within a range of  $\pm 125$  ppm for water vapor mole fractions ranging from around ~~35000 ppm~~ to 30000 ppm%. Hence, a good transferability  
30 and long-term stability of the analyzers water vapor measurements can be assumed.

In order to examine the robustness of the 2009 calibration it was repeated in 2013 using a G2401-m analyzer (CFKBDS2003) in comparison to the identical dew point mirror (Dewmet TDH, Michell Instruments Ltd.) that was used for the calibration in 2009. Both instruments measured simultaneously dried ambient air from a high-pressure tank, which was humidified by a dew point generator (Li-610 from Li-Cor) to specific water levels between 2°C and 20°C dew point. Higher and lower water vapor levels could not be reached due to the environmental conditions in the laboratory.

During the 2009 calibration the dew point mirror measurement was based on its original calibration, conducted by the manufacturer against test equipment traceable to the NIST primary standard in the end of 2000. In 2010 the dew point mirror was recalibrated by the manufacturer, however, no information was given on how the calibration factors changed. Another calibration by the manufacturer in 2014, shortly after the 2013 comparison of the CRDS instrument and the dew point mirror, showed no drift beyond the uncertainty range of the dew point mirror (given by manufacturer: 0.2°C at +20°C dew point, increasing linearly to 0.4°C at -60°C dew point (2-sigma)) compared to the calibration in 2010.

### 3.1.3 Calibration Bench for FISH-Instrument

During the DENCHAR inter-comparison campaign in 2011 the CRDS analyzer CFKB2004 was compared against the laboratory calibration bench used regularly for calibration of the reference instrument FISH. This calibration bench consists of a humidifier, a mixing unit to mix dry and humid air, and a reference water vapor instrument, the MBW Dew Point instrument (model K-1806/DP30-SHSX- III, MBW Elektronik AG, Switzerland, www.mbw.ch) (Meyer et al., 2015). For the comparison the CRDS instrument was connected to the calibration bench via a ~~three~~ meter long 1/8" OD FEP-line. Since the calibration bench provided a flow of about 3500 sccm an open-split was installed in front of the FEP-line to allow the CRDS analyzer to sample at its low flow rate of 100 sccm. During the comparison four humidity steps covering the range of 2 to around 600 ppm were measured. This corresponds to the standard calibration range of the FISH calibration bench and is a good addition to the dew point mirror calibration range. Maximum uncertainty of the calibration bench is given as  $\pm 4\%$  (1-sigma) (~~relative~~) by Meyer et al. (2015).

Due to the low sample flow (100 sccm) through the analyzer and the large difference in water vapor mole fraction between the measured air and the outside air during the comparison, permeation of water vapor through the FEP-tube (3 m length in the calibration setup) has to be considered. To provide information from which a correction factor for the permeation effect could be determined, a dry tank air stream (~2 ppm water vapor mole fraction) at different flow rates (100 sccm and 3500 sccm) was provided through the FEP-tube. Assuming that for a flow of 3500 sccm the contribution of the permeation to the water vapor mole fraction in the flow is negligible the correction factor was computed as the difference in the calibrated CRDS H<sub>2</sub>O mole fraction between these two measurements.

### 3.1.4 Calibration by CO<sub>2</sub> dilution effect

In addition to the standard calibrations by different hygrometers a newovel, and completely independent calibration method was tested, which takes advantage of the high precision CO<sub>2</sub> measurements by the CRDS analyzer. Specifically, the dilution

effect of water vapor on the CO<sub>2</sub> mole fraction is used: If water vapor is added to dry air, while total pressure and temperature of the gas remain unchanged, the mole fractions of the residual air components decrease. The mole fraction of CO<sub>2</sub> in dry air is

$$X_{CO_2}^{dry} = \frac{n_{CO_2}}{n_{air} + n_{CO_2}}, \quad (3)$$

5 whereas the CO<sub>2</sub> mole fraction in wet air is given as

$$X_{CO_2}^{wet} = \frac{n_{CO_2}}{n_{air} + n_{CO_2} + n_{H_2O}} \quad (4)$$

Together with the wet air mole fraction of water,

$$X_{H_2O}^{wet} = \frac{n_{H_2O}}{n_{air} + n_{CO_2} + n_{H_2O}}, \quad (5)$$

Eqs. (3) and (4) yield:

$$10 \frac{X_{CO_2}^{wet}}{X_{CO_2}^{dry}} = 1 - X_{H_2O}^{wet} \quad (6)$$

Thus, the amount of water vapor in air is directly linked to the ratio of the CO<sub>2</sub> wet and dry air mole fraction of the air.

15 However, the measured CO<sub>2</sub> mole fraction ~~of from~~ the CRDS instruments in wet air is not only influenced by the dilution effect, but also through pressure broadening effects of the water vapor. To separate both effects the measurement software of the analyzer was modified to allow for a fine scan of the CO<sub>2</sub> and water vapor absorption line. While the peak height, which is normally used for the measurement, is influenced by both effects, the peak area ~~gets~~ only changes ~~ed by~~ due to the dilution effect. The fine scan, combined with spectral models and fitting procedures optimized for this purpose, provides the peak areas with sufficient precision.

20 To test the concept, pressurized zero air with 3000 ppm CO<sub>2</sub> from a high pressure tank was split into two paths, as can be seen in Fig. 2. The air in one path was humidified in a bubbler. Afterwards the dry and wet gas stream were recombined and then measured by a CRDS-analyzer (CFADS2196, model G2301) in fine scan mode. With the help of mass flow controllers in both paths the water concentration of the combined stream could be varied without changing the CO<sub>2</sub> dry mole fraction by changing the flow of each path while the total flow was kept constant. The adjustable water vapor levels were limited by the remaining humidity in the pressurized air on the one hand and the environmental conditions in the laboratory on the other.

25 The measurements alternated between the water line and the CO<sub>2</sub> line, ~~whereby it required~~ requiring about 1.3 s to make one pair of measurements. Since the pressure and temperature in the sample cell were kept constant the measured peak areas were proportional to the mole fractions:

$$X_i = C_i * A_i \quad (7)$$

Where A is the peak area and C the proportionality or “calibration” factor. With Eq. (6) this yields:

$$A_{CO_2}^{wet} = A_{CO_2}^{dry} (1 - C_{H_2O} * A_{H_2O}) \quad (8)$$

30 If the measured area of the CO<sub>2</sub> line is plotted as a function of the measured area of the water line, the calibration factor for water vapor C<sub>H<sub>2</sub>O</sub> is just the ratio of the slope and the intercept. The scan of the water line provided also the conventional

water vapor measurement using the peak height of the absorption line, which allowed for comparing the two water vapor estimates.

## 3.2 Results and Discussion

### 3.2.1 Dew Point Mirror

5 Plot (a) in Fig. 3 shows the self-broadening and offset corrected, but uncalibrated, water vapor mole fraction measured by the Picarro CRDS instrument (in the following referred to as “ $\text{H}_2\text{O}_{\text{uncalibrated}}$ ”) against the measurements ~~from~~ the dew point mirror (using its factory calibration from 2010, which was confirmed in 2014) during the comparison in 2013. The Dewmet measurements were converted from dew point to wet air mole fraction based on the Goff-Gratch equation (Goff, 1957). The corresponding fit can be seen as blue line, the grey line indicates the calibration curve of the 2009 experiment (Eq. (2)). The  
10 uncertainty of the fitted slope is composed of the fitting error and the uncertainty of the dew point mirror. Uncertainty of the Dewmet is given by the manufacturer as  $0.2^\circ\text{C}$  at  $+20^\circ\text{C}$  dew point, increasing linearly to  $0.4^\circ\text{C}$  at  $-60^\circ\text{C}$  dew point ~~(2-sigma), which corresponds to a relative uncertainty of 1.3 %~~. In order to check the linearity of the CRDS instrument the CRDS and Dewmet data were also fitted using a quadratic fit. The slope of the quadratic fit was determined as  $0.807 \pm 0.011$ , which agrees well with the slope of the linear fit (0.802) taking account of the uncertainty range. The ~~impact~~  
15 ~~coefficient~~ of the quadratic term, ~~(determined as  $0.0024 \pm 0.0021$ )~~, ~~on the result~~ is ~~not significant small compared to the overall uncertainty range~~. Thus, the CRDS analyzer can be considered as linear.

Figure 3 b) shows the difference between the 2009 and the 2013 calibration. The error bars demonstrate the uncertainty range, which comprises the dew point mirror uncertainty during the 2009 as well as during the 2013 experiment. The relative difference of the two calibrations, ~~shown on the right axis~~, increases from ~~1.72-2 % (relative)-at 0.89800 ppm% CRDS water vapor mole fraction~~ ~~(2.2 % with regard to the measured mole fraction from the dew point mirror)~~ up to ~~4.43.6 % (relative)-at 2.530600 ppm% (4.4 %)~~, indicating significant differences for water vapor mole fractions above about ~~10000 ppm%~~.

This difference between the 2009 and the 2013 calibrations is much larger than the uncertainties of the instruments and the calibration transfer give reason to expect. The largest source of uncertainty is the dew point mirror with ~~1.3 % (rel.)~~ uncertainty. ~~Repeatability-Precision~~ of the CRDS analyzer is given as  $<14$  ppm by the manufacturer. Uncertainty of the calibration transfer between different Picarro analyzers, which has to be considered since the two calibrations were done with different CRDS instruments, is  $<125$  ppm (or  $0.5\%$  relative at ~~2.5000 %ppm~~ water vapor mole fraction) (Rella et al., 2013). Since the difference between the two calibrations is up to ~~4.43.6 % (4.4 %)~~, ~~(rel.)~~ the dew point mirror, the Picarro analyzer, or both instruments must have been drifting.

30 In order to test the stability of the water vapor measurements of the CRDS analyzer, the CFKBDS2003 instrument was compared to another Picarro CRDS analyzer (CFADS37, model G3101-m) once in 2011 and again in 2014. During the experiments both instruments measured in parallel pressurized, dried ambient air from a high-pressure tank, which was

humidified by a dew point generator (Li-610 from Li-Cor) to specific water levels between 2°C and 20°C dew point. Plot (a) in Fig. 4 shows the result of the 2014-comparison (~~black points~~). The blue line is the quadratic fit of the 2011-comparison. The relative difference (right axis in plot (b)) between the two experiments is are <0.3 % (~~(rel.)~~) for water vapors <8000 ppm. Since it is unlikely that both instruments drifted in the same way, this strongly suggests that the CFKBDS2003 analyzer did not drift significantly in the three years between the two comparisons against the dew point mirror.

This conclusion together with the calibration history of the dew point mirror (see Sect. 3.1) suggests that the large differences between the two calibrations of the CRDS instrument in 2009 and 2013 are caused by drift of the dew point mirror calibration. During the 2009 experiment the dew point mirror has not been calibrated for nearly nine years when it was used for the 2009 experiment, but was ~~not~~ calibrated well during the 2013 experiment, enough and ~~t~~ Thus, only the results of the 2013 experiment, corresponding to a calibration factor of 0.802 for the CRDS water vapor measurements, are considered reliable. Accuracy of the calibration is limited by the uncertainty range of the dew point mirror (1.3 % (~~(rel.)~~), 2-sigma). For water vapor levels <0.7000 ppm% the calibration is only extrapolated based on the measurements between 0.7000 – 2.5000 ppm%, which has to be accounted for in the uncertainty estimate.

### 15 3.2.2 Calibration Bench for FISH-Instrument

Plot (a) in Fig. 5 shows the result of the comparison between the CRDS analyzer and the FISH calibration bench, during which four different water levels in the range 2-600 ppm were measured. The water vapor measurements off from the CRDS analyzer are offset corrected and calibrated according to the 2013 dew point mirror comparison (~~calibration factor = 0.802 ± 0.010~~), and additionally corrected for another 3.5 ppm resulting from permeation of water vapor from air surrounding the 3 m FEP inlet line. Subsequently, the CRDS measurements were converted from wet to dry air mole fractions according to:

$$H_2O_{dry} = \frac{H_2O_{wet}}{1 - H_2O_{wet}} \quad (9)$$

A linear fit of the data (~~blue line~~) shows that the dew point mirror-calibration of the CRDS was within 3 % of the FISH calibration bench and showed an offset of 12.2 ppm. Uncertainties of the fit coefficients (slope: ±0.04, offset: ±0.5 ppm) were estimated assuming a worst-case scenario by including 4 % (~~(rel.)~~) bias of the FISH calibration bench and 1.3 % (~~(rel.)~~) uncertainty of the dew point mirror calibration. The residuals (difference between the Fish calibration bench and the fit) for water vapors > 100 ppm, which can be seen in plot (b), are small compared to the uncertainty range of the FISH calibration bench of 4 % (~~(rel.)~~) indicated by the error bars. For the measurement point at 2 ppm water vapor the relative residuals are larger (6.2 %), due to the influence of the 12.2 ppm offset.

The CRDS analyzer and the FISH calibration bench agree within 3 % (~~(rel.)~~) in the water vapor range up to 600 ppm after correcting for an offset of 12 ppm. This confirms that extrapolation of the dew point mirror calibration to water vapor levels below 0.7000 ppm% is appropriate, at least within the uncertainty of 4 % (~~(rel.)~~) assumed for the calibration bench. Regarding the offset of 12.2 ppm it has to be considered that the measured air at the lowest water level, which has the largest effect on

the estimation of the offset, was perhaps not completely in equilibrium with the inner surface of the connection line between the CRDS instrument and the calibration bench and the tubing inside the analyzer. If the water vapor mole fraction in the gas stream decreases water molecules adsorbed at the surface are released until a new equilibrium with the air is reached. Due to the large length of the connection line (3 m) the inner surface is relatively large and thus, the ~~balance~~ingequilibration process takes relatively long time. For the higher measured water levels the ~~balance~~equilibrium is reached faster. ~~and f~~Furthermore, except for the highest measured level, memory effects were canceled out by measuring the water level twice: once from low to high, and once going from high to low mole fractions. The differences in the two measurement sequences are smaller than 1 ppm. Another possibility would be that the offset was caused by either an outgassing or a very small leak.

### 3.2.3 Calibration by CO<sub>2</sub> dilution effect

Figure 6 a) shows the comparison of the water vapor mole fraction determined with the help of the CO<sub>2</sub> dilution method and the conventional water vapor measurements, which are offset corrected and calibrated according to the 2013 dew point mirror comparison (~~calibration factor = 0.802 ± 0.010~~), during the fine scan experiment. A linear fit of the data (~~blue line~~) indicates a bias of 2.9 % (~~rel.~~) of the dilution-based water vapor compared to the dew point mirror calibration and an offset of 16.1 ppm. The uncertainty estimates (slope: ±0.013, offset: ±1.6 ppm) are based on the uncertainty of the dew point mirror calibration (1.3 % ~~rel.~~, 2-sigma). Residuals can be seen in plot (b) of Fig. 6.

The water vapor mole fraction calculated with the CO<sub>2</sub> dilution method and the conventional water vapor measurements calibrated according to the 2013 dew point mirror comparison agree within 2.9 % (~~rel.~~) in the water vapor range from 300 ~~ppm~~ to 2.7000 ~~ppm~~%. The residuals (difference between water vapor mole fraction based on dilution method and the linear fit) are small, but show a slight systematic shape depending on the water vapor level. An offset was determined as 16.1 ppm, however the lowest measurement was made at around 300 ppm and thus, the offset is based on extrapolation. Higher scatter in the residuals at low water vapor (~~< 0.2500 %ppm~~) might indicate a different behavior for this range. Hence, the estimated error of 1.6 ppm for the offset is likely ~~unrealistic~~ a significant underestimate.

Estimating the uncertainty of the CO<sub>2</sub> dilution method is not straightforward. The repeatability of the peak area measurements accounts only for less than 0.1 % (~~rel.~~) uncertainty (1-sigma), whereas systematic errors can have a larger influence on the accuracy. One potential error is the direct spectroscopic interference of either water on CO<sub>2</sub> or vice versa, which ~~was~~ tried to ~~avoid~~exclude by careful selection of the used absorption lines and detailed spectral models. To check for remaining influences an additional test was conducted: Since a direct spectroscopic interference would affect the measurements differently for different CO<sub>2</sub> concentrations, the fine scan experiment was repeated with 400 ppm CO<sub>2</sub> instead of the original 3000 ppm CO<sub>2</sub>. Unfortunately, the pressurized ~~zero~~ air with 400 ppm CO<sub>2</sub> also contained 2 ppm of methane, whereas the 3000 ppm CO<sub>2</sub> air was pure CO<sub>2</sub> in zero air. Thus, a neighboring methane absorption line had to be considered, which added another variable to the analysis. In future experiments this should be excluded by preparing a set of high-pressure tanks of exactly the same air composition but different CO<sub>2</sub> concentrations. The calibration constant C<sub>H<sub>2</sub>O</sub> of the water measurements (see Eq. (7) and (8)) measured for 400 ppm CO<sub>2</sub> was 0.6 % (~~rel.~~) larger than the calibration constant

measured for 3000 ppm CO<sub>2</sub>. Another systematic error can arise if the spectroscopic models and fitting procedures do not perfectly account for the changes in the absorption line shapes during varying water vapor mole fractions. For this experiment the absorption line shape model was carefully tested over the range of conditions in the analyzer, and it was found that the corresponding error can be neglected compared to the other sources of uncertainty.

5 Recently, a potentially serious source of systematic error regarding the pressure control in the sample cell was discovered: Observations suggest that the pressure sensor ~~has a non-linear dependence on~~ ~~is sensitive to~~ water vapor and thus, the pressure in the sample cell is stabilized to a humidity dependent value instead of the fixed 186.65 hPa (= 140 Torr) (Reum et al., 2017). A possible reason could be adsorption of water molecules on the sensor. While this error in the CRDS water vapor measurements is corrected by the calibration of the instrument with another hygrometer, it has to be considered for the  
 10 dilution calibration, since the used CO<sub>2</sub> measurements are affected as well. To assess the quantitative effect of such an incorrect pressure ~~reading adjustment~~ we assume the pressure in the measurement cell to be

$$p = p_0 + \Delta p, \quad (10)$$

where  $p_0$  is the actual set point at 186.65 hPa (140 Torr) and  $\Delta p$  is a pressure difference depending on the water vapor mole fraction. Experiments with an additional independent pressure measurement presented by Reum et al. (2017), as well as  
 15 analysis of the behavior of the proportional valve, which controls the pressure in the sample cell, show that ~~If  $\Delta p$  is small enough it can be assumed that the pressure~~ changes linearly with water vapor for mole fractions >2500 ppm (see Figure 2 in Reum et al., 2017). ~~and~~ Therefore, the peak areas of the absorption lines follow

$$A(p) = A(p_0) * \left(1 + \frac{\Delta p}{p_0}\right). \quad (11)$$

Substituting  $A_{CO_2}(p_0)$  and  $A_{H_2O}(p_0)$  according to Eq. (11) in Eq. (8) yields:

$$20 \quad A_{CO_2}^{wet}(p) = A_{CO_2}^{dry} \left(1 - C_{H_2O}(p_0) * A_{H_2O}(p) + \frac{\Delta p}{p_0}\right). \quad (12)$$

Thus, the bias in the sample cell pressure introduces an error to the calibration constant  $C_{H_2O}(p_0)$ , which is proportional to the relative pressure change  $\Delta p / p_0$ . ~~Experiments with an additional independent pressure measurement (Reum et al., (2017), as well as analysis of the behavior of the proportional valve, which controls the pressure in the sample cell, show that~~ determine  
 25  $\Delta p$  is around as about 0.7 mbar (0.5 Torr) for a water vapor mole fraction of 30000 ppm-%. Hence, the pressure bias causes an error of <0.4 % ~~(rel.)~~ to the CO<sub>2</sub> dilution calibration method. Note however, that the change in cell pressure with humidity is not ~~fully~~ linear for water vapor mole fractions <2500 ppm(see Figure 2 in Reum et al., 2017), which could be the reason for the slightly systematic shape in the residuals at this low water vapor levels ~~with lower values around 0.25 % water vapor mole fraction~~ (Fig. 6 b).

In summary it can be said, that an uncertainty at percent or even sub-percent level is achievable for the dilution method in  
 30 future experiments. Using a conservative estimate of 1 % ~~(rel.)~~ uncertainty (1-sigma) for assessing water vapor from the CO<sub>2</sub> dilution experiment presented here, added to the 1.3 % ~~(rel.)~~ uncertainty of the dew point mirror calibration, comparison of the dilution based estimate ( $H_2O_{dilution} = (1.029 \pm 0.023) * H_2O_{dewmet2013}$ ) with the FISH calibration bench ( $H_2O_{CalBench} = (0.97 \pm 0.04) * H_2O_{dewmet2013}$ ) (neglecting the offsets) shows an overlap within their combined uncertainty. Note that this also means

that the H<sub>2</sub>O calibration via dilution of CO<sub>2</sub> is statistically consistent with the classical calibration using dew point or frost point hygrometers. This is a promising result for this experiment, especially when considering that different CRDS instruments were used and the comparisons took place two years apart.

Follow on experiments can achieve better and more reliable results for the water calibration by CO<sub>2</sub> dilution, if also low water vapor levels (<300 ppm) are measured, the sample cell pressure is corrected for deviations due to different water vapor levels, optimized spectral models and fitting procedures are applied, and sample air with a CO<sub>2</sub> mole fraction in the atmospheric range is used. To determine a calibration factor for the water vapor estimates based on peak height measurements, which is the standard measurement method of the CRDS analyzers at the moment, since it provides better short-term precision than the peak area measurements, the experiment can be simplified. As can be seen in Eq. (8) the water vapor mole fraction ( $C_{H_2O} \cdot A_{H_2O}$ ) can be calculated, if the dry and wet peak areas of the CO<sub>2</sub> absorption line are known. Thus, the measurement of the water vapor peak area can be skipped, which reduces the overall uncertainty. On the other hand, for low water vapor mole fractions (<10 ppm) a wrong pressure reading (as described above) has a higher impact since it affects the wet peak area, but not the dry peak area measurement. By looking at the deviation of the ratio between wet and dry peak area to one the error gets enhanced even more.

Obviously the dilution method can be applied to other species, too, and is not limited to CO<sub>2</sub> and water vapor. The same principle can be used for any species measurable by a CRDS analyzer, provided that the corresponding dilution effect is large enough to be detectable with sufficient precision.

### 3.2.4 Calibration Summary

Table 1 shows in summary the results of the different calibration experiments. The water vapor ranges used in the comparison were determined by the experimental setups of the experiments and the standard calibration ranges of the instruments. The uncertainties of the coefficients for the FISH calibration bench comparison result from the dew point mirror calibration uncertainty and the uncertainty of the calibration bench. For the CO<sub>2</sub> dilution effect it is the dew point mirror calibration uncertainty plus a conservative estimate of the dilution method uncertainty. Note that both offsets, ~~or rather their~~ uncertainties, are likely not reliable.

Based on this experiments the calibration constant of  $0.802 \pm 0.010$  from the dew point mirror comparison in 2013 is recommended for the water vapor measurements ~~off~~ from the CRDS instrument.

## 4 Analysis of the flight data and comparison with the reference instruments

During the DENCHAR flight campaign between 23 May and 1 June 2011 four inter-comparison flights with a total flight time of about 14 hours were conducted with a Learjet 35A. Starting from an airbase in Hohn, ~~(Germany)~~, the flights covered a region ranging from ~~North~~ northern Germany and Poland to ~~South~~ southern Norway, the North and the Baltic Sea, and altitudes up to 12.5 km. ~~Hence,~~ so that also the lower stratosphere was reached. Two instruments served as reference

instruments for the water vapor measurements. The first was CR-2, a frost point hygrometer with an accuracy of  $\pm 0.1^\circ\text{C}$  (1-sigma) dew point (manufacturer data, Buck Research Instruments L.L.C., US, www.hygrometers.com). The second reference instrument was the Fast In-situ Stratospheric Hygrometer (FISH), which is based on the Lyman- $\alpha$  photofragment fluorescence technique and has ~~an total~~ accuracy of 6-8 % (1-sigma) in the range from 4 to 1000 ppm and ~~down to~~ 0.3 ppm for lower mixing ratios down to 1 ppm (Meyer et al., 2015). The CR-2 was connected to a backward-faced inlet to avoid sampling of cloud and ice particles. In contrast, FISH measured total water instead of only water vapor during the campaign, since ~~their's forward-oriented inlet system of FISH allowed for~~ resulted in sampling of cloud droplets and ice crystals when present. ~~Both instruments were operated by the research centre Jülich.~~

## 4.1 Results and ~~D~~discussion

### 4.1.1 Measurement ~~repeatability~~precision

To assess the measurement ~~repeatability~~precision of the CRDS analyzers, water vapor measurements during the periods with stable atmospheric conditions, as pressure and temperature, were selected. Of course, there are still natural variations left in the data, therefore only upper limits of the ~~repeatability~~precision can be estimated. After correcting for offset and calibration (according to the 2013 dew point mirror comparison: in the following simply referred to as CRDS measured water vapor, calibration factor =  $0.802 \pm 0.010$ ), the standard deviation of the difference between the 0.4 Hz data and the 60 seconds moving averages is calculated as a measure of short-term fluctuations. In order to avoid additional noise from variations in sample cell pressure, periods with unstable sample cell pressure were neglected. Deviations of the sample cell pressure from its set point of 186.65 hPa can occur during sudden, fast changes in altitude for which the pressure adjustment is too slow to adapt. Figure 7 shows the resulting short-term fluctuations (i.e. the standard deviations of the difference between the 0.4 Hz data and the 60 second moving averages) for different water vapor ranges. The significance of the results certainly depends highly strongly on the number of data, which were available to calculate the standard deviations in each water vapor interval. Thus, in order to find a reliable estimate for the measurements, results based on a larger number of data are highlighted. Although high scatter of the data between 30 and 100 ppm makes it difficult to find a reliable estimate, the flight data suggest an upper limit for the measurement ~~repeatability~~precision (1-sigma) of 4 ppm for  $\text{H}_2\text{O} < 10$  ppm, 20 % or 10 ppm (whichever is smaller) for  $10 \text{ ppm} < \text{H}_2\text{O} < 100$  ppm, or 5 % (rel.) (whichever is greater) for water vapor  $< 100$  ppm, and 5 % (rel.) or 30 ppm (whichever is smaller) for water vapor  $> 100$  ppm, ~~indicated by the three black lines in Fig. 7.~~

For comparison ~~repeatability~~precision estimates of the CRDS water vapor measurements determined under laboratory conditions, at 2.5 s time resolution and for an integration time of 30 s, are shown in Table 2. They were derived from experiments during which the CRDS analyzer measured pressurized, dried ambient air from a high-pressure tank, which was humidified by a dew point generator (Li-610 from Li-Cor) to specific water levels. For water vapor  $< 100$  ppm the results of the flight and laboratory data are in good agreement. For water vapor  $> 1000$  ppm the laboratory data indicate that a

~~repeatability-precision~~ of 30 ppm for the flight data is a very conservative estimate, which is most likely due to natural variations in the atmosphere.

Compared to the reference instruments the ~~repeatability-precision~~ of the CRDS analyzer is worse at low water vapor levels (<100 ppm), but comparable at higher levels.

#### 5 4.1.2 Response time

Figure 8 and Figure 9 show selected time periods of the third flight on 31 May and the fourth flight on 1 June, respectively.

~~In addition to the offset corrected and calibrated (according to the 2013 dew point mirror comparison, calibration factor =  $0.802 \pm 0.010$ ) water vapor measurements of the CRDS analyzer, in the following simply referred to as CRDS measured water vapor, shown here as black points (30 seconds mean as grey points) is shown along with~~ the flight data ~~off~~from the reference instruments CR-2 (dark blue squares) and FISH (light blue triangles); ~~also w~~Water vapor measurements ~~from~~of two additional analyzers, ~~which took that~~ participated in the inter-comparison campaign; are also presented. Flight data of WaSul-Hygro, a tunable diode laser-based dual-channel photoacoustic humidity measuring system (Tátrai et al., 2015), ~~are shown as orange diamonds; and~~ flight data of the Selective Extractive Airborne Laser Diode Hygrometer (SEALDH-1), which is based on tunable diode laser absorption spectroscopy (Buchholz et al., 2013), ~~are shown as green triangles~~ (please note: not to be confused with the currently used new instrument SEALDH-II, which has a much better performance and smaller uncertainties). Furthermore, the saturated water vapor (~~violet points~~) is added to point out that the measurements were taken outside of clouds.

The flight data of all analyzers in Fig. 8 and Fig. 9 indicate that the response time of the CRDS is not slower than any similar to that of the other instruments. This applies for the whole water vapor range and for both transition directions: from wet to dry conditions as well as from dry to wet. For water vapors <100 ppm the response time is comparable to the FISH instrument, as shown in the enlarged section of Fig. 8. During the increase of water vapor from 200 to 1200 ppm in about one minute, shown in the enlarged section in Figure 9, no significant delay can be detected. Thus, the low sample gas flow of 100 sccm and the ~~one meter~~50 cm long inlet line cause no disadvantages. As expected, the slowest response is shown by the CR-2, whose measurement signal tends to overshoot and oscillate after fast changes in water vapor.

Results of a simple laboratory test, where a three way valve was used to switch between wet (around 23000 ppm) and dry air (around 10 ppm), allowed to estimate the 10-90 % rise and 90-10 % fall times as 6-7 seconds and the recovery time to 99 % of the final water vapor level ~~a challenge~~ as 25 seconds. For a step from 2-3000 % (~~around 20°C dew point~~) to 10 ppm water vapor mole fraction the measurement takes about 200 seconds to get down. The times are pretty much identical whether or not the 50 cm long inlet line is included.

#### 30 4.1.3 Comparison to reference instruments

Figure 10 shows the in-flight CRDS measured water vapor (~~in black~~) and the reference instruments CR-2 (~~dark blue~~) and FISH (~~light blue~~), as well as the corresponding atmospheric pressure levels (~~in green~~), for all four flights. ~~For better~~

~~comparison also the 30 seconds mean of the CRDS data is shown (in grey).~~ Due to an internal leak FISH could not deliver reliable data for the first two flights. CRDS data ~~offrom~~ the first flight after around 1 pm were ~~influenced~~compromised by icing of the inlet, since the deicing of the Rosemount inlet was accidentally not switched on during that flight.

The ~~absolute~~ water vapor differences between the three instruments for different water vapor intervals ~~can be seen~~are plotted in Fig. 11\_ ~~(CRDS CR 2 as black points, CRDS FISH as dark blue diamonds, CR 2 FISH as light blue triangles).~~ The water vapor measurements of CR 2 are chosen as x axis, because they cover all flights in contrast to the FISH data. The differences are calculated from the 30 seconds mean data of all analyzers and are averaged over intervals of 1 ppm, 10 ppm, 100 ppm, 1000 ppm, and 10000 ppm water vapor in the corresponding water vapor ranges of 0-10 ppm, 10-100 ppm, 100-1000 ppm, 1000-10000 ppm, and >10000 ppm, respectively. The standard deviations of the averaged differences are shown as error bars. For plotting reasons all differences <1 ppm are set to 1 ppm. CRDS data influenced by icing during the first flight are ~~neglected~~omitted. Likewise, measurement data of all instruments in the presence of clouds are excluded, since FISH measured total water. Based on observations during the flights, which are recorded in the flight logs, this concerns in particular all measurements made between 11:13 am and 11:40 am on flight four (1 June).

A reliable evaluation is hard to make as the reference instrument FISH operated successfully for only two of the four ~~flightssampled data only for two flights,~~ and for flight 4 the measurements diverge significantly from the CR-2 data to a large extent. Moreover, the slow response of the CR-2 and the oscillations of the signal after sudden changes in water vapor are problematic for the comparison. Furthermore, flight data between 11:13 am and 11:40 am for flight 4 could not be used, due to the occurrence of clouds. However, it is interesting that ~~also~~ the CRDS water vapor measurements deviate from the CR-2 during that period, as can be seen in Fig. 10. Figure 12 shows a closer look at this cloud-affected flight section. The CRDS measured water vapor ~~(black points, 30 seconds mean as grey points)~~ are is plotted together with flight data of CR-2 ~~(dark blue squares), FISH (light blue triangles), Wasul-Hygro (orange diamonds) and SEALDH-I (green triangles).~~ The latter two show approximately the same behavior as the CR-2, This is in line with expectation, since all three shared the same backward-faced inlet, which prevented from sampling cloud droplets. However, while the CRDS shows a behavior similar to that of FISH (measuring total water) with H<sub>2</sub>O mole fractions within clouds larger than that corresponding to saturated water vapor ~~(violet points).~~ This might indicate that also the CRDS samples cloud particles. This indicates that the CRDS sampling was likely also affected by cloud particles, i.e. the separation in the Rosemount air inlet of ice particles and water droplets from the sample air is not fully efficient. In fact relative humidity measurements ~~offrom~~ the MOZAIC humidity device, which uses the same type of Rosemount Inlet housing, occasionally show similar artifacts, when measuring within clouds, ~~that containing~~ liquid water ~~partieles~~ (air temperature > -40°C) (Smit et al., 2013). Most likely some small ice particles and water droplets are able to follow the sharp right angle turn of the minor air flow into the inner part of the Rosemount housing, instead of flying straight through the main channel of the housing (see Figure 2.6 in Smit et al., 2013). Such small enough particles could be produced e.g. by the shattering of water droplets or ice crystals in the Rosemount housing. However, due to the very short time period the sample air stays inside the housing until it passes the sensor elements and leaves again through a small outlet, only the liquid water droplets can evaporate fast enough to be observed by

the humidity device. In contrast, as can be seen in Fig. 12, the CRDS measurements do show cloud artifacts also at air temperatures ~~(black line)~~ below  $-40^{\circ}\text{C}$ , i.e. in pure ice clouds. Most likely the reason for this is that water droplets and ice particles enter the inlet line of the CRDS and are evaporated within the inlet line or at the heated inlet filter of the CRDS. Meaningful statistics about how often droplets and ice particles are measured in clouds can be obtained as soon as more flight data ~~from~~ the CRDS analyzer are available within the IAGOS project, since every IAGOS aircraft equipped with the GHG-package is also equipped with the MOZAIC humidity device and a cloud probe.

The absolute differences in Fig. 11 indicate a positive difference between the CRDS and CR-2 of  $<10\%$  or  $10\text{ ppm}$  (whichever is greater) for water vapor ranges  $>10\text{ ppm}$ . FISH has a negative deviation to both instruments in that range. For water vapor  $>100\text{ ppm}$  the data imply a difference of  $10\text{-}20\%$ . For the interval of  $10\text{-}100\text{ ppm}$  water vapor the difference to the CRDS is around  $10\%$ , to CR-2 about  $10\text{ ppm}$ . At very low water vapor ( $<10\text{ ppm}$ ) the reference instruments show a good agreement during flight 3, but disagree strongly during flight 4. On average the CR-2 has a positive bias  $<2\text{ ppm}$  against FISH. For the CRDS the water vapor data suggest a positive bias  $<2\text{-}3\text{ ppm}$  to the CR-2, but the measurements are highly affected by the slow response of the CR-2. Comparison to FISH likewise indicates a positive difference  $<2\text{-}3\text{ ppm}$ . During comparison against the FISH calibration bench the CRDS analyzer showed a positive bias of  $12.2\text{ ppm}$  (see Sect. 3.2), which strengthens the presumption that the air hasn't been in equilibrium for the lowest water vapor level measured during the experiment.

Meyer et al. (2015) report an agreement of FISH with other in situ and remote sensing hygrometers under field conditions of about  $\pm 5\text{-}20\%$  ~~@~~  $<10\text{ ppm}$  and  $\pm 0\text{-}15\%$  ~~@~~  $>10\text{ ppm}$ . Thus, results of the comparison between CRDS and FISH during the DENCHAR inter-comparison campaign are at the upper end of that range.

## 20 Conclusions

During the DENCHAR inter-comparison flight campaign in Hohn (Germany) in May-June 2011 a commercial cavity ring-down spectroscopy (CRDS) based gas analyzer (G2401-m, Picarro Inc.,\_US) was installed on a Learjet to measure atmospheric water vapor,  $\text{CO}_2$ ,  $\text{CH}_4$  and  $\text{CO}$ . The components of the instrument and the inlet system are identical to those chosen for the IAGOS-core Greenhouse Gas package.

For the calibration of the water vapor measurements three different methods were tested. The standard calibration of the CRDS analyzer is the comparison against a dew point mirror (Dewmet TDH, Cooled Mirror Dewpointmeter, Michell Instruments Ltd., UK) in the range from about  ~~$0.8000\%$~~  to  $3.0000\%$  ~~ppm~~ water vapor mole fraction. If the dew point mirror is calibrated regularly by the manufacturer, the accuracy of this calibration method is limited by the uncertainty range of the dew point mirror ( $1.3\%$  ~~(rel.)~~, 2-sigma). A comparison against the FISH calibration bench, during the DENCHAR flight campaign, in the range from  $2\text{-}600\text{ ppm}$  water vapor, confirmed that the extrapolation of the dew point mirror calibration down to low water vapor levels is possible, and that the standard calibration of the CRDS analyzer is in agreement with the FISH calibration within the  $4\%$  uncertainty range (1-sigma) of the FISH calibration bench. Furthermore, a new and

completely independent calibration method, which is based on measurement of the dilution effect of water vapor on the CO<sub>2</sub> mole fraction, was presented. This new method was found to agree with the dew point mirror calibration within 2.9 % ~~(rel.)~~ in the water vapor range from 300 ppm to 2.7000 ppm%. Assuming a conservative 1 % ~~(rel.)~~ uncertainty (1-sigma) for the CO<sub>2</sub> dilution method, comparison of the dilution based estimate with the FISH calibration bench showed an overlap within their combined uncertainty. Thus, the water vapor calibration via dilution of CO<sub>2</sub> is statistically consistent with the classical calibration using dew point or frost point hygrometers. The dilution method can be used for the calibration of other species, too, provided they and the corresponding diluted species are measurable by a CRDS analyzer and the ~~corresponding~~ dilution effect is large enough to be within the detection limits.

An upper limit of the precision (1-sigma) of the water vapor measurements by the CRDS was determined from flight data of the DENCHAR inter-comparison campaign, as 4 ppm for H<sub>2</sub>O <10 ppm, 20 % or 10 ppm (whichever is smaller) for 10 ppm < H<sub>2</sub>O <100 ppm, 4 ppm or 5 % (rel.) (whichever is greater) for water vapor <100 ppm, and 5 % ~~(rel.)~~ or 30 ppm (whichever is smaller) for water vapor >100 ppm. A more reliable estimate will be possible as soon as more H<sub>2</sub>O flight data are available. During the four DENCHAR flights the CRDS analyzer showed a good time response (10-90 % rise and 90-10 % fall times: 6-7 s, recovery time to 99 % of ~~a challenge~~ the final water vapor level: 25 s) and long-term stability for the water vapor measurements. Comparison against the reference instruments was difficult, due to lack of data availability of FISH, the slow response of CR-2, the exclusion of data, which were affected by clouds, and the partly poor agreement between FISH and CR-2. However, for water vapor levels >10 ppm the flight data imply a negative difference between the CRDS and FISH from about 10-20 % and a positive difference between the CRDS and CR-2 of <10 % or 10 ppm (whichever is greater). For water vapor <10 ppm the flight data suggest a positive bias of <2-3 ppm to both, FISH and CR-2.

Accuracy (1-sigma) of the CRDS instrument was estimated, based on the laboratory calibrations, as 1 % ~~(relative)~~ for the water vapor range from 2.5000 ppm% down to ~~0.7000 ppm%~~, than increasing to 5 % ~~(relative)~~ at 50 ppm water vapor. Accuracy at water vapor mole fractions below 50 ppm was difficult to assess, as the reference systems suffered from lack of data availability.

Future deployment of the CRDS system within IAGOS will help to further evaluate the performance, via better statistics and long-term comparison to the MOZAIC humidity device, which is deployed on each IAGOS aircraft. Thus, essential water vapor measurements, including regular in situ data in the sensible UTLS region and vertical profiles of H<sub>2</sub>O in the troposphere and lower stratosphere for major parts of the globe, are expected to be delivered for validation of remote sensing based observations from satellites and ground, for the improvement of the performance of climate models and weather forecasts, or for climate trend studies.

### 30 **Competing interests**

The authors declare that they have no conflict of interest.

## Acknowledgements

This work was supported by the European Commission through the FP6 project IAGOS Design Study (contract number 011902-DS) and IAGOS-ERI, an FP7 project (grant agreement no. 212128). The DENCHAR flight campaign was funded by the European Commission's FP7 project EUFAR. Funding for IAGOS Deutschland was provided by the Bundesministerium für Bildung und Forschung (BMBF), Germany. Additional funding from the German Max Planck Society to support the instrument development is greatly acknowledged. Development and operation of Wasul-Hygro was supported by the Hungarian Research and Technology Innovation Fund (OTKA), project no. NN109679.

## References

- Buchholz, B., Kühnreich, B., Smit, H.G.J., and Ebert, V.: Validation of an extractive, airborne, compact TDL spectrometer for atmospheric humidity sensing by blind intercomparison, *Appl. Phys. B*, 110, 249-262, doi:10.1007/s00340-012-5143-1, 2012.
- Buchholz, B., Afchine, A., Klein, A., Schiller, C., Krämer, M., and Ebert, V.: HAI, a new airborne, absolute, twin dual-channel, multi-phase TDLAS-hygrometer: background, design, setup, and first flight data, *Atmos. Meas. Tech.*, 10, 35-57, doi:10.5194/amt-10-35-2017, 2017.
- Chen, H., Winderlich, J., Gerbig, C., Hofer, A., Rella, C. W., Crosson, E. R., Van Pelt, A. D., Steinbach, J., Kolle, O., Beck, V., Daube, B. C., Gottlieb, E. W., Chow, V. Y., Santoni, G. W., and Wofsy, S. C.: High-accuracy continuous airborne measurements of greenhouse gases (CO<sub>2</sub> and CH<sub>4</sub>) using the cavity ringdown spectroscopy (CRDS) technique, *Atmos. Meas. Tech.*, 3, 375–386, doi:10.5194/amt-3-375-2010, 2010.
- Crosson, E. R.: A cavity ring-down analyzer for measuring atmospheric levels of methane, carbon dioxide, and water vapor, *Appl. Phys. B*, 92, 403–408, doi:10.1007/s00340-008-3135-y, 2008.
- Dessler, A. E., Zhang, Z., and Yang, P.: Water-vapor climate feedback inferred from climate fluctuations, 2003–2008, *Geophys. Res. Lett.*, 35, L20704, doi:10.1029/2008GL035333, 2008.
- Dirksen, R. J., Sommer, M., Immler, F. J., Hurst, D. F., Kivi, R., and Vömel, H.: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde, *Atmos. Meas. Tech.*, 7, 4463–4490, doi:10.5194/amt-7-4463-2014, 2014.
- Fahey, D. W., Gao, R.S., Carslaw, K. S., Kettleborough, J., Popp, P. J., Northway, M. J., Holecek, J. C., Ciciora, S. C., McLaughlin, R. J., Thompson, T. L., Winkler, R. H., Baumgardner, D. G., Gandrud, B., Wennberg, P. O., Dhaniyala, S., McKinney, K., Peter, Th., Salawitch, R. J., Bui, T. P., Elkins, J.W., Webster, C. R., Atlas, E. L., Jost, H., Wilson, J. C., Herman, R. L., Kleinböhl, A., and von König, M.: The Detection of Large HNO<sub>3</sub>-Containing Particles in the Winter Arctic Stratosphere, *Science*, 291, 1026-1031, 2001.
- Fahey, D. W., Gao, R.-S., Möhler, O., Saathoff, H., Schiller, C., Ebert, V., Krämer, M., Peter, T., Amarouche, N., Avallone, L. M., Bauer, R., Bozóki, Z., Christensen, L. E., Davis, S. M., Durrý, G., Dyroff, C., Herman, R. L., Hunsmann, S., Khaykin,

- S. M., Mackrodt, P., Meyer, J., Smith, J. B., Spelten, N., Troy, R. F., Vömel, H., Wagner, S., and Wienhold, F. G.: The AquaVIT-1 Intercomparison of Atmospheric Water Vapor Measurement Techniques, *Atmos. Meas. Tech.*, 7, 3177-3213, doi:10.5194/amt-7-3177-2014, 2014.
- Filges, A., Gerbig, C., Chen, H., Franke, H., Klaus, C., and Jordan, A.: The IAGOS-core greenhouse gas package: a measurement system for continuous airborne observations of CO<sub>2</sub>, CH<sub>4</sub>, H<sub>2</sub>O and CO, *Tellus B*, 67,27989, doi:10.3402/tellusb.v67.27989, 2015.
- Forster, P. M. de F. and Shine, K. P.: Assessing the climate impact of trends in stratospheric water vapor, *Geophys. Res. Lett.*, 29(6), 1086, doi:10.1029/2001GL013909, 2002.
- Gierens, K., Schumann, U., Helten, M., Smit, H., and Marengo, A.: A distribution law for relative humidity in the upper troposphere and lower stratosphere derived from three years of MOZAIC measurements, *Ann. Geophys.*, 7, 1218-1226, 1999.
- Goff, J. A.: Saturation pressure of water on the new Kelvin temperature scale, Transactions of the American society of heating and ventilating engineers, presented at the semi-annual meeting of the American society of heating and ventilating engineers, 347-354, Murray Bay, Que. Canada, 1957.
- Hall, E. G., Jordan, A. F., Hurst, D. F., Oltmans, S. J., Vömel, H., Kühnreich, B., and Ebert, V.: Advancements, measurement uncertainties, and recent comparisons of the NOAA frostpoint hygrometer, *Atmos. Meas. Tech.*, 9, 4295-4310, doi:10.5194/amt-9-4295-2016, 2016.
- ~~Harries, J. E., Russell, J. M., Tuck, A. F., Gordley, L. L., Purcell, P., Stone, K., Bevilacqua, R. M., Gunson, M., Nedoluha, G., and Traub, W. A.: Validation of measurements of water vapor from the Halogen Occultation Experiment (HALOE), *J. Geophys. Res.*, 101(D6), 10,205-10,216, 1996.~~
- Helten, M., Smit, H. G. J., Straeter, W., Kley, D., Nedelec, P., Zöger, M., and Busen, R.: Calibration and performance of automatic compact instrumentation for the measurement of relative humidity from passenger aircraft, *J. Geophys. Res.*, 103, 25,643–25,652, 1998.
- ~~Hoareau, C., Keckhut, P., Noel, V., Chepfer, H., and Baray, J.-L.: A decadal cirrus clouds climatology from ground-based and spaceborne lidars above the south of France (43.9° N–5.7° E), *Atmos. Chem. Phys.*, 13, 6951-6963, doi:10.5194/acp-13-6951-2013, 2013.~~
- Hurst, D. F., Oltmans, S. J., Vömel, H., Rosenlof, K. H., Davis, S. M., Ray, E. A., Hall, E. G., and Jordan, A. F.: Stratospheric water vapor trends over Boulder, Colorado: Analysis of the 30 year Boulder record, *J. Geophys. Res.*, 116, D02306, doi:10.1029/2010JD015065, 2011.
- Kaufmann, S., Voigt, C., Jurkat, T., Thornberry, T., Fahey, D. W., Gao, R.-S., Schlage, R., Schäuble, D., and Zöger, M.: The airborne mass spectrometer AIMS – Part 1: AIMS-H<sub>2</sub>O for UTLS water vapor measurements, *Atmos. Meas. Tech.*, 9, 939–953, doi:10.5194/amt-9-939-2016, 2016.
- Kiehl, J. T. and Trenberth, K. E.: Earth’s annual global mean energy budget, *Bull. Am. Meteorol. Soc.*, 78(2), 197-208, doi:10.1175/1520-0477, 1997.

- Kley, D., Russell III, J. M., and Phillips, C., SPARC assessment of upper tropospheric and stratospheric water vapour, World Climate Research Programme Rep. 113, WMO/TD No. 1043, SPARC Rep. No. 2, 324 pp, Geneva, Switzerland, 2000.
- Kunz, A., Mueller, R., Homonnai, V., Janosi, I. M., Hurst, D., Rap, A., Forster, P. M., Rohrer, F., Spelten, N., and Riese, M.: Extending water vapor trend observations over Boulder into the tropopause region: trend uncertainties and resulting radiative forcing, *J. Geophys. Res. Atmos.*, 118, 11269-11284, doi:10.1002/jgrd.50831, 2013.
- ~~Lambert, A., Read, W., Livesey, N., Santee, M., Manney, G., Froidevaux, L., Wu, D., Schwartz, M., Pumphrey, H., Jimenez, C., Nedoluha, G., Cofield, R., Cuddy, D., Daffer, W., Drouin, B., Fuller, R., Jarnot, R., Knosp, B., Pickett, H., Perun, V., Snyder, W., Stek, P., Thurstans, R., Wagner, P., Waters, J., Jucks, K., Toon, G., Stachnik, R., Bernath, P., Boone, C., Walker, K., Urban, J., Murtagh, D., Elkins, J., and Atlas, E.: Validation of the Aura Microwave Limb Sounder middle atmosphere water vapor and nitrous oxide measurements, *J. Geophys. Res.*, 112, D24S36, doi:10.1029/2007JD008724, 2007.~~
- Marengo, A., Thouret, V., Nédélec, P., Smit, H., Helten, M., Kley, D., Karcher, F., Simon, P., Law, K., Pyle, J., Poschmann, G., Von Wrede, R., Hume, C., and Cook, T.: Measurement of ozone and water vapor by Airbus in-service aircraft: The MOZAIC airborne program, An overview, *J. Geophys. Res.*, 103(D19), 25,631–25,642, 1998.
- May, R. D.: Open-path, near-infrared tunable diode laser spectrometer for atmospheric measurements of H<sub>2</sub>O, *J. Geophys. Res.*, 103, 19,161-19,172, 1998.
- ~~J.-Meyer, J., Rolf, C., Schiller, C., Rohs, S., Spelten, N., Afchine, A., Zöger, M., Sitnikov, N., Thornberry, T. D., Rollins, A. W., Bozóki, Z., Tátrai, D., Ebert, V., Kühnreich, B., Mackrodt, P., Möhler, O., Saathoff, H., Rosenlof, K. H., and Krämer, M.: Two decades of water vapor measurements with the FISH fluorescence hygrometer: a review, *Atmos. Chem. Phys.*, 15, 8521–8538, doi:10.5194/acp-15-8521-2015, 2015.~~
- Milz, M., von Clarmann, T., Fischer, H., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Mengistu Tsidu, G., Steck, T., Stiller, G. P., Funke, B., López-Puertas, M., and Koukouli, M. E.: Water vapor distributions measured with the Michelson Interferometer for Passive Atmospheric Sounding on board Envisat (MIPAS/Envisat), *J. Geophys. Res.*, 110, D24307, doi:10.1029/2005JD005973, 2005.
- Oltmans, S. J., Vömel, H., Hofmann, D. J., Rosenlof, K. H., and Kley, D.: The increase in stratospheric water vapor from balloonborne, frostpoint hygrometer measurements at Washington, D.C., and Boulder, Colorado, *Geophys. Res. Lett.*, 27(21), 3453–3456, doi:10.1029/2000GL012133, 2000.
- ~~Petzold, A., Thouret, V., Gerbig, C., Zahn, A., Brenninkmeijer, C. A.M., Gallagher, M., Hermann, M., Pontaud, M., Ziereis, H., Boulanger, D., Marshall, J., Nédélec, P., Smit, H. G.J., Friess, U., Flaud, J.-M., Wahner, A., Cammas, J.-P., and Volz-Thomas, A.: Global-scale atmosphere monitoring by in-service aircraft - current achievements and future prospects of the European Research Infrastructure IAGOS, *Tellus B*, 67, 28452, doi:10.3402/tellusb.v67.28452, 2015.~~
- ~~Poshyvailo, L., Müller, R., Konopka, P., Günther, G., Riese, M., Podglajen, A., and Ploeger, F.: Sensitivities of modelled water vapour in the lower stratosphere: temperature uncertainty, effects of horizontal transport and small-scale mixing, *Atmos. Chem. Phys.*, 18, 8505-8527, doi:10.5194/acp-18-8505-2018, 2018.~~

~~Read, W. G., Lambert, A., Baemeister, J., Cofield, R. E., Christensen, L. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Fetzer, E., Froidevaux, L., Fuller, R., Herman, R., Jarnot, R. F., Jiang, J. H., Jiang, Y. B., Kelly, K., Knosp, B. W., Kovalenko, L. J., Livesey, N. J., Liu, H. C., Manney, G. L., Pickett, H. M., Pumphrey, H. C., Rosenlof, K. H., Sabounchi, X., Santee, M. L., Schwartz, M. J., Snyder, W. V., Stek, P. C., Su, H., Takaes, L. L., Thurstans, R. P., Vomel, H., Wagner, P. A., Waters, J. W., Webster, C. R., Weinstock, E. M., and Wu, D. L.: Aura Microwave Limb Sounder upper tropospheric and lower stratospheric H<sub>2</sub>O and relative humidity with respect to ice validation, *J. Geophys. Res.*, 112, D24S35, doi:10.1029/2007JD008752, 2007.~~

5 Rella, C.W.: Accurate Greenhouse Gas Measurements in Humid Gas Streams Using the Picarro G1301 Carbon Dioxide/Methane/Water Vapor Gas Analyzer, Picarro, Inc., Santa Clara, CA, USA, <http://www.picarro.com/resources/whitepapers>, 2010.

Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., Karion, A., Miles, N. L., Richardson, S. J., Steinbacher, M., Sweeney, C., Wastine, B., and Zellweger, C.: High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air, *Atmos. Meas. Tech.*, 6, 837-860, doi:10.5194/amt-6-837-2013, 2013.

10 Reum, F., Gerbig, C., Lavric, J. V., Rella, C. W., Göckede, M.: An improved water correction function for Picarro greenhouse gas analyzers, *Atmos. Meas. Tech. Discuss.*, <https://doi.org/10.5194/amt-2017-174>, 2017.

15 Riese, M., Ploeger, F., Rap, A., Vogel, B., Konopka, P., Dameris, M., and Forster, P.: Impact of uncertainties in atmospheric mixing on simulated UTLS composition and related radiative effects, *J. Geophys. Res.*, 117, D16305, doi:10.1029/2012JD017751, 2012.

20 Rind, D., Chiou, E. W., Chu, W., Oltmans, S., Lerner, J., Larsen, J., McCormick, M.P., and McMaster, L.: Overview of the Stratospheric Aerosol and Gase Experiment II water vapour observations: Method, validation, and data characteristics, *J. Geophys. Res.*, 98(D3), 4835-4856, 1993.

Rollins, A. W., Thornberry, T. D., Gao, R. S., Smith, J. B., Sayres, D. S., Sargent, M. R., Schiller, C., Krämer, M., Spelten, N., Hurst, D. F., Jordan, A. F., Hall, E. G., Vömel, H., Diskin, G. S., Podolske, J. R., Christensen, L. E., Rosenlof, K. H., Jensen, E. J., and Fahey, D. W.: Evaluation of UT/LS hygrometer accuracy by intercomparison during the NASA MACPEX mission, *J. Geophys. Res.*, 119, 1915-1935, doi:10.1002/2013JD020817, 2014.

25 Rozanov, A., Weigel, K., Bovensmann, H., Dhomse, S., Eichmann, K.-U., Kivi, R., Rozanov, V., Vömel, H., Weber, M., and Burrows, J. P.: Retrieval of water vapor vertical distributions in the upper troposphere and the lower stratosphere from SCIAMACHY limb measurements, *Atmos. Meas. Tech.*, 4, 933-954, doi:10.5194/amt-4-933-2011, 2011.

30 Russell, J. M., Gordley, L. L., Park, J., Drayson, S. R., Tuck, A. F., Harries, J. E., Cicerone, R. J., Crutzen, P. J., and Frederick, J. E.: The Halogen Occultation Experiment, *J. Geophys. Res.*, 98, 10,777-10,797, 1993.

Sayres, D. S., Moyer, E. J., Hanisco, T. F., St. Clair, J. M., Keutsch, F. N., O'Brien, A., Allen, N. T., Lapson, L., Demusz, J. N., Rivero, M., Martin, T., Greenberg, M., Tuozzolo, C., Engel, G. S., Kroll, J. H., Paul, J. B., and Anderson, J. G.: A new cavity based absorption instrument for detection of water isotopologues in the upper troposphere and lower stratosphere, *Rev. Sci. Instrum.*, 80, 044102, doi:10.1063/1.3117349, 2009.

- Schneider, M., Hase, F., and Blumenstock, T.: Water vapour profiles by ground-based FTIR spectroscopy: study for an optimised retrieval and its validation, *Atmos. Chem. Phys.*, 6, 811-830, doi:10.5194/acp-6-811-2006, 2006.
- Smit, H. G. J., Volz-Thomas, A., Helten, M., Pätz, H.-W., and Kley, D.: An in-flight Calibration Method for Near-Real-Time Humidity Measurements with the Airborne MOZAIC sensor, *Journal of atmospheric and oceanic technology*, 25, 656-666, doi:10.1175/2007JTECHA975.1, 2008
- Smit, H.G.J., Kivi, R., Vömel, H., and Paukkunen, A.: Thin Film Capacitive Sensors, in: *Monitoring Atmospheric Water Vapour: Ground-Based Remote Sensing and In-situ Methods*, Kämpfer, N. (Ed. ), ISSI Scientific Report Series 10, 11-38, Springer Science+Buisness Media, New York, doi: 10.1007/978-1-4614-3909-7, 2013.
- Smith, C. A., Haigh, J. D., and Toumi, R.: Radiative forcing due to trends in stratospheric water vapour, *Geophys. Res. Lett.*, 28, 179-182, 2001.
- Solomon, S., Rosenlof, K. H., Portmann, R. W., Daniel, J. S., Davis, S. M., Sanford, T. J., and Plattner, G.-K.: Contributions of stratospheric water vapor to decadal changes in the rate of global warming, *Science*, 327, 1219-1223, doi:10.1126/science.1182488, 2010.
- Tátrai, D., Bozóki, Z., Smit, H., Rolf, C., Spelten, N., Krämer, M., Filges, A., Gerbig, C., Gulyás, G., and Szabó, G.: Dual-channel photoacoustic hygrometer for airborne measurements: background, calibration, laboratory and in-flight intercomparison tests, *Atmos. Meas. Tech.*, 8, 33-42, doi:10.5194/amt-8-33-2015, 2015.
- Volz-Thomas, A., Berg, M., Heil, T., Houben, N., Lerner, A., Petrick, W., Raak, D., Pätz, H.-W.: Measurements of total odd nitrogen (NO<sub>y</sub>) aboard MOZAIC in-service aircraft: instrument design, operation and performance, *Atmos. Chem. Phys.*, 5(3), 583-595, 2005.
- Volz-Thomas, A., Cammas, J.-P., Brenninkmeijer, C. A. M., Machida, T., Cooper, O., Sweeney, C., Waibel, A.: Civil Aviation Monitors Air Quality and Climate, *Journal of the Air & Waste Management Association*, October 2009, 16-19. 2009.
- von Clarmann, T., Höpfner, M., Kellmann, S., Linden, A., Chauhan, S., Funke, B., Grabowski, U., Glatthor, N., Kiefer, M., Schieferdecker, T., Stiller, G. P., and Versick, S.: Retrieval of temperature, H<sub>2</sub>O, O<sub>3</sub>, HNO<sub>3</sub>, CH<sub>4</sub>, N<sub>2</sub>O, ClONO<sub>2</sub> and ClO from MIPAS reduced resolution nominal mode limb emission measurements, *Atmos. Meas. Tech.*, 2, 159-175, doi:10.5194/amt-2-159-2009, 2009.
- Vömel, H., David, D. E., and Smith, K.: Accuracy of tropospheric and stratospheric water vapor measurements by the cryogenic frost point hygrometer: Instrumental details and observations, *J. Geophys. Res.*, 112, D08305, doi:10.1029/2006JD007224, 2007.
- Vömel, H., Naebert, T., Dirksen, R., and Sommer, M.: An update on the uncertainties of water vapor measurements using cryogenic frost point hygrometers, *Atmos. Meas. Tech.*, 9, 3755-3768, doi:10.5194/amt-9-3755-2016, 2016.
- Weigel, K., Rozanov, A., Azam, F., Bramstedt, K., Damadeo, R., Eichmann, K.-U., Gebhardt, C., Hurst, D., Kraemer, M., Lossow, S., Read, W., Spelten, N., Stiller, G. P., Walker, K. A., Weber, M., Bovensmann, H., and Burrows, J. P.: UTLS

water vapour from SCIAMACHY limb measurements V3.01 (2002–2012), *Atmos. Meas. Tech.*, 9, 133-158, doi:10.5194/amt-9-133-2016, 2016.

Weinstock, E. M., Smith, J. B., Sayres, D. S., Pittman, J. V., Spackman, J. R., Hints, E. J., Hanisco, T. F., Moyer, E. J., St. Clair, J. M., Sargent, M. R., and Anderson, J. G.: Validation of the Harvard Lyman- $\alpha$  in situ water vapor instrument: Implications for the mechanisms that control stratospheric water vapor, *J. Geophys. Res.*, 114, D23301, doi:10.1029/2009JD012427, 2007.

Winderlich, J., Chen, H., Gerbig, C., Seifert, T., Kolle, O., Lavrič, J. V., Kaiser, C., Höfer, A., and Heimann, M.: Continuous low maintenance CO<sub>2</sub>/CH<sub>4</sub>/H<sub>2</sub>O measurements at the Zotino Tall Tower Observatory (ZOTTO) in Central Siberia, *Atmos. Meas. Tech.*, 3, 1113-1128, doi:10.5194/amt-3-1113-2010, 2010.

10 World Meteorological Organization (WMO): Guide to Meteorological Instruments and Methods of Observation, 2008 edition, updated in 2010, WMO-No. 8, Geneva, 2012.

Zahn, A., Christner, E., van Velthoven, P. F. J., Rauthe-Schöch, A., and Brenninkmeijer, C. A. M.: Processes controlling water vapor in the upper troposphere/lowermost stratosphere: An analysis of 8 years of monthly measurements by the IAGOS-CARIBIC observatory, *J. Geophys. Res. Atmos.*, 119, 11505–11525, doi:10.1002/2014JD021687, 2014.

15 Zöger, M., Afchine, A., Eicke, N., Gerhards, M.-T., Klein, E., McKenna, D. S., Mörschel, U., Schmidt, U., Tan, V., Tuitjer, F., Woyke, T., and Schiller, C.: A new family of balloon-borne and airborne Lyman- $\alpha$  photofragment fluorescence hygrometers, *J. Geophys. Res.*, 104, 1807-1816, doi:10.1029/1998JD100025, 1999.

**Table 1: Overview of the different calibration methods.**

method	water vapor mole	result
	fraction range [ppm]	
dew point mirror (Dewmet) from 2013	<del>0-7000</del> - <del>2-5000</del> %	$H_2O_{\text{Dewmet2013}} = (0.802 \pm 0.010) * H_2O_{\text{uncalibrated}}$
FISH calibration bench	2-600- <del>ppm</del>	$H_2O_{\text{calBench}} = (0.97 \pm 0.04) * H_2O_{\text{Dewmet2013}} - (12.2 \pm 0.5) \text{ ppm}$
CO <sub>2</sub> dilution effect	300 <del>ppm</del> - <del>2-7000</del> -%	$H_2O_{\text{dilution}} = (1.029 \pm 0.023) * H_2O_{\text{Dewmet2013}} - (16.1 \pm 1.6) \text{ ppm}$

Table 2: Repeatability Precision estimates (1-sigma) of the CRDS water vapor measurements derived from laboratory experiments.

water vapor mole fraction [ppm]	<u>repeatability-precision</u> at	<u>repeatability-precision</u> at
	2.5 s time resolution [ppm]	30 s integration time [ppm]
3	<6	<2
30	<10	<5
5000	<9	<5
8000	<10	<2
12000	<10	<4
19000	<12	<6

5

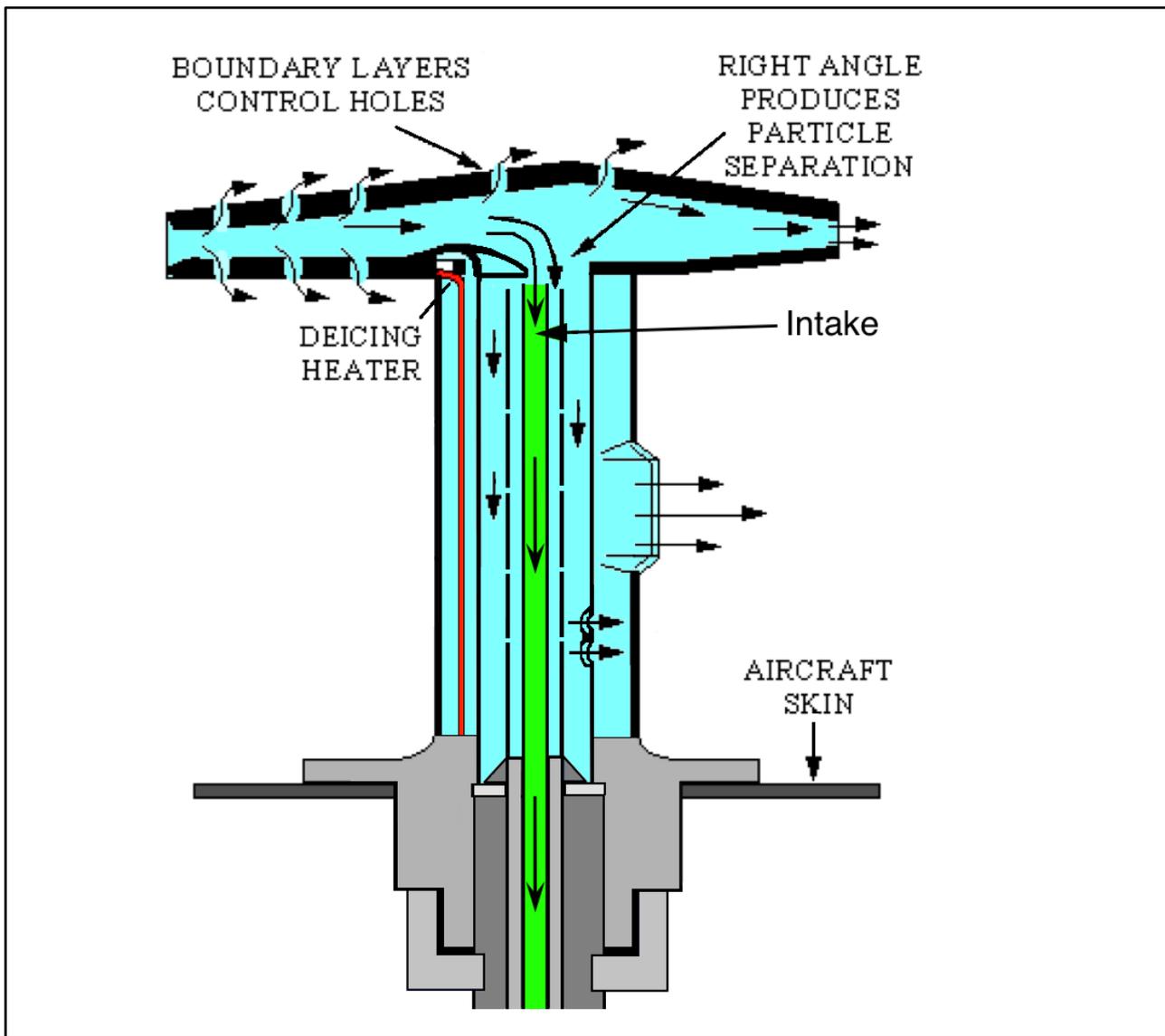


Figure 1: Cross section of the inlet line (green) mounted into a Rosemount Total Air Temperature housing (model 102B, adapted from Stickney et al. (1994)). The inlet line is pointed orthogonal to the airflow through the housing to prevent from sampling larger aerosols, ice particles, and water droplets.

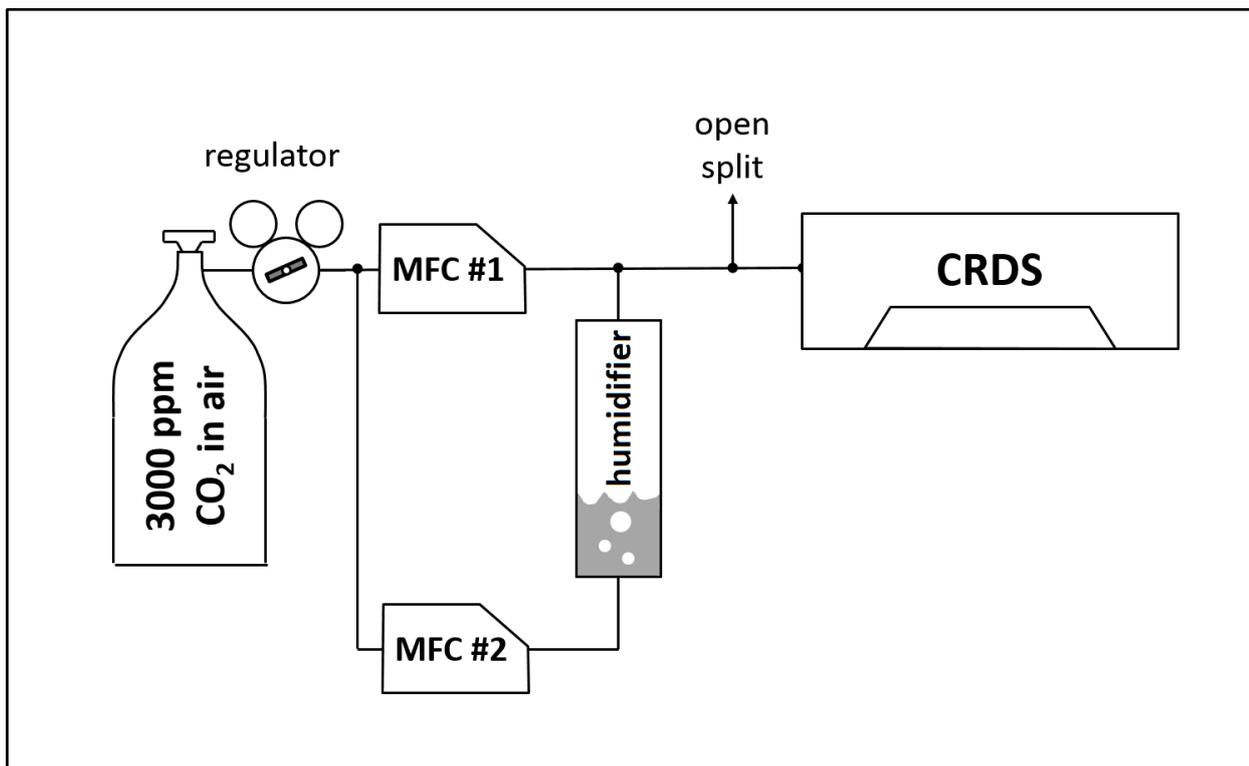
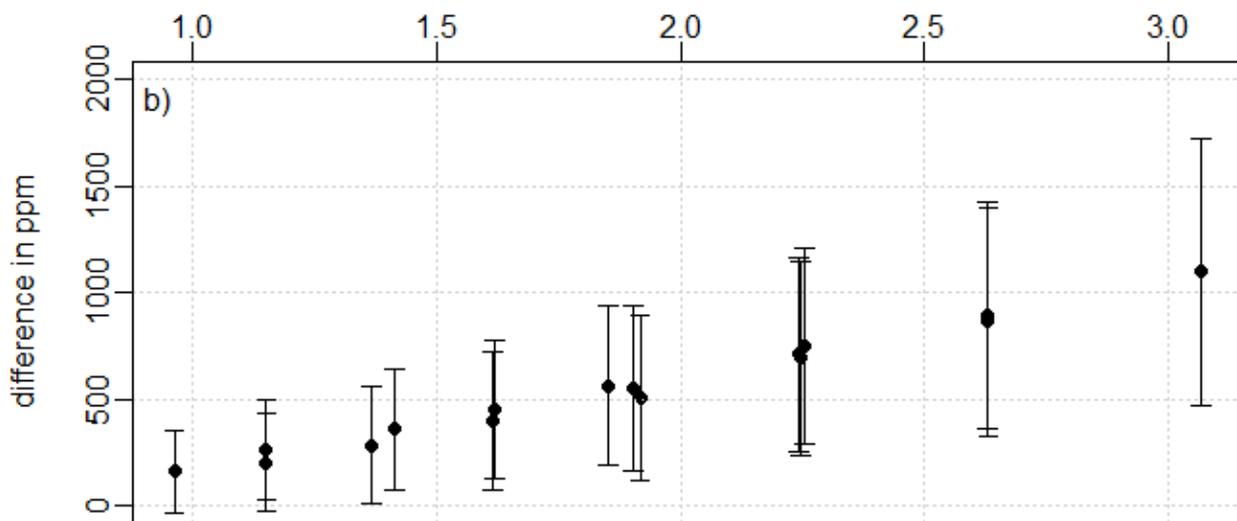
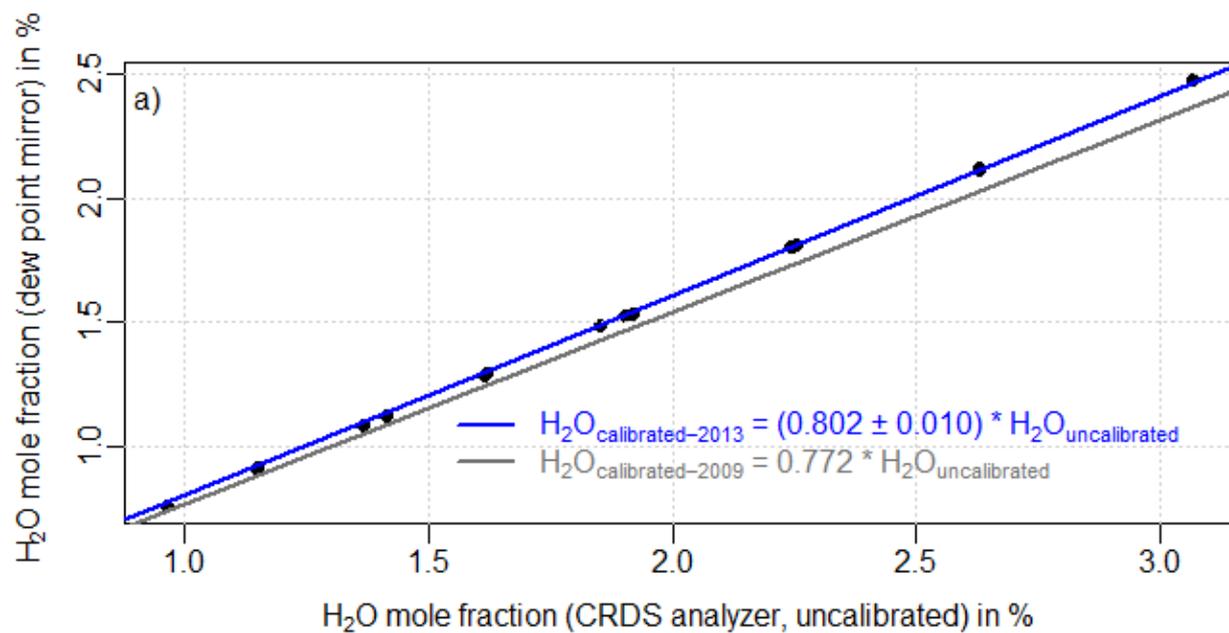


Figure 2: Experimental setup for the water vapor calibration by the CO<sub>2</sub> dilution effect. MFC #1 and 2 are mass flow controllers.



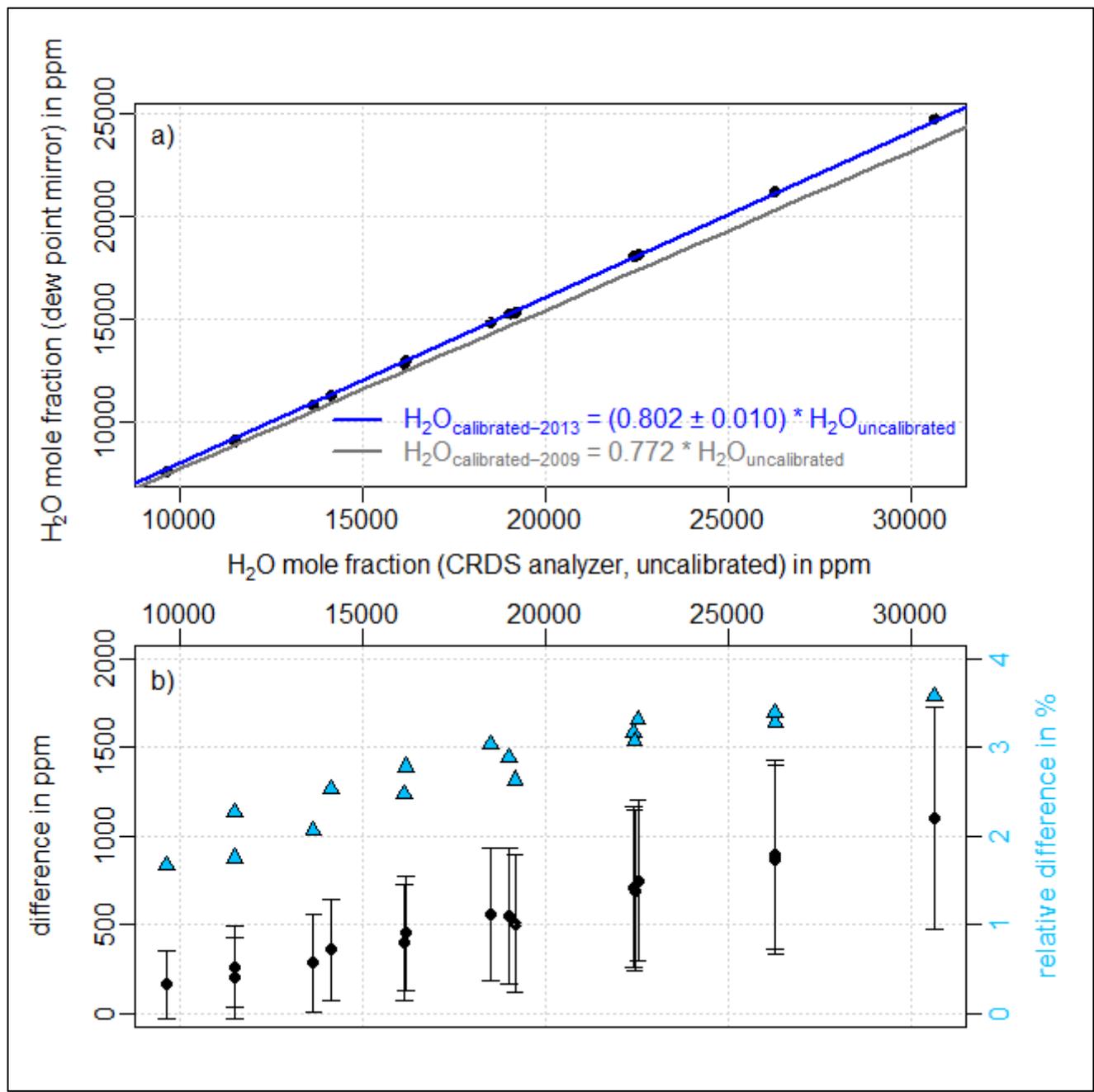
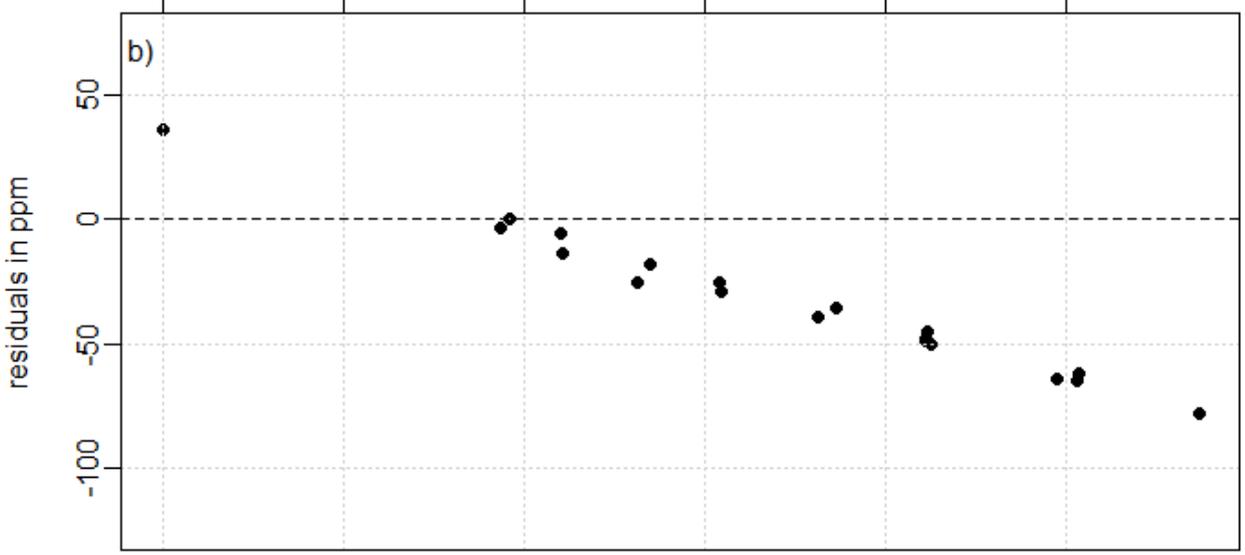
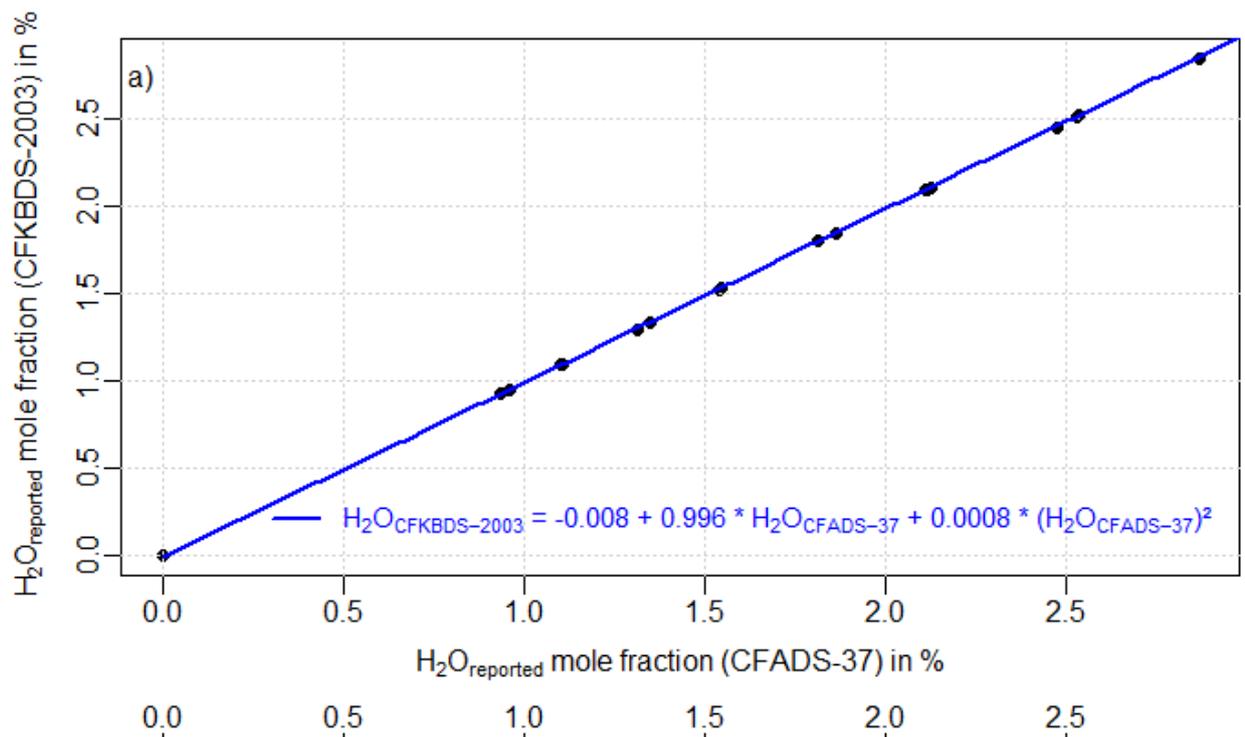


Figure 3: a) Uncalibrated water vapor measurements offfrom the Picarro CRDS analyzer (CFKBDS2003) against the measurements offfrom the dew point mirror (Dewmet TDH) during the 2013 calibration. The corresponding fit is shown as blue line, the calibration curve of the 2009 comparison as a grey line. The uncertainty of the fitted slope is composed of the fitting error and the uncertainty of the dew point mirror. b) Water vapor difference between the 2009 and 2013 experiments. The error bars indicate the uncertainty range, which results from the combination of the dew point mirror uncertainties during the 2009 and the 2013 calibration. The relative differences are plotted on the right axis as light blue triangles.



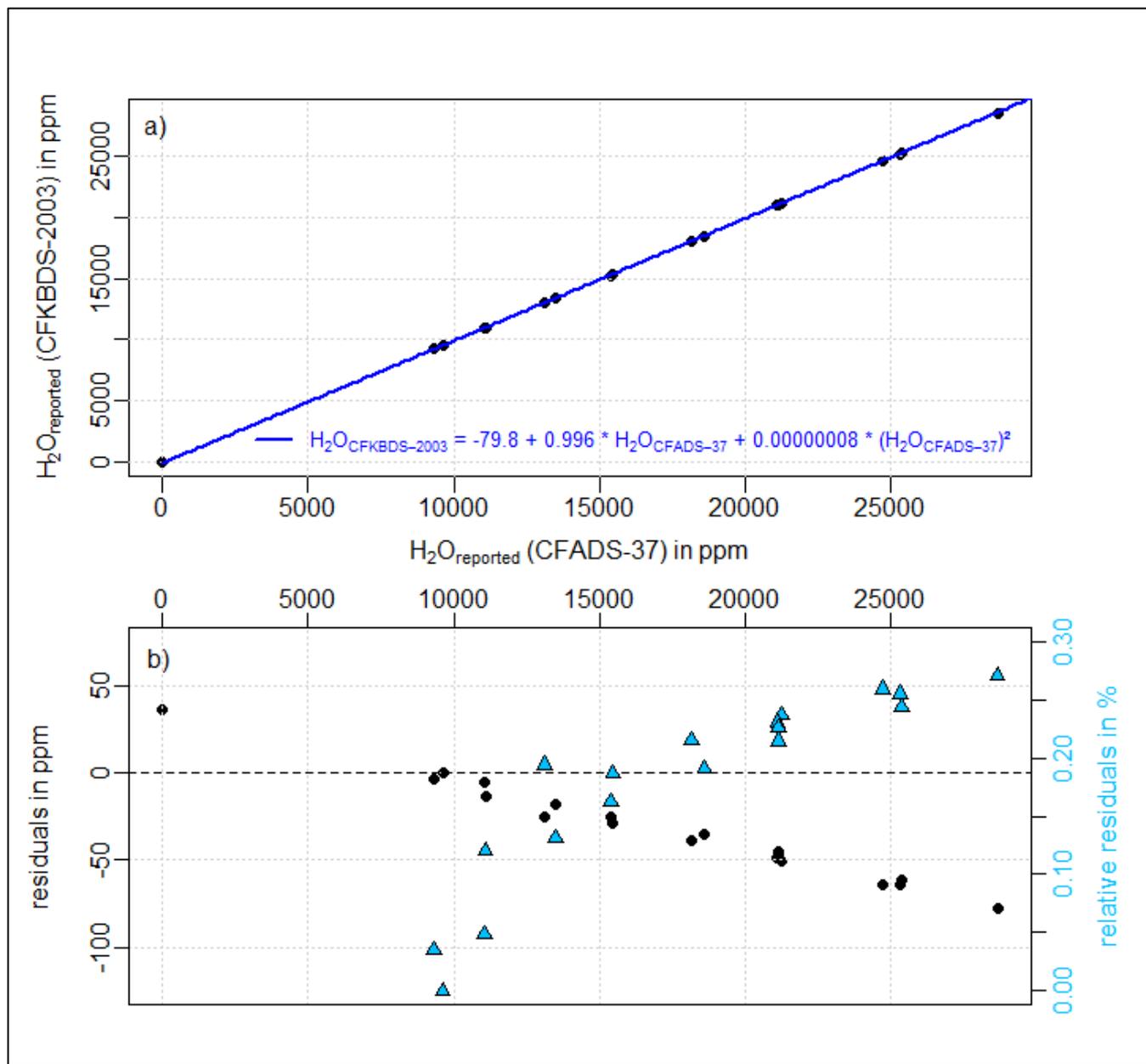
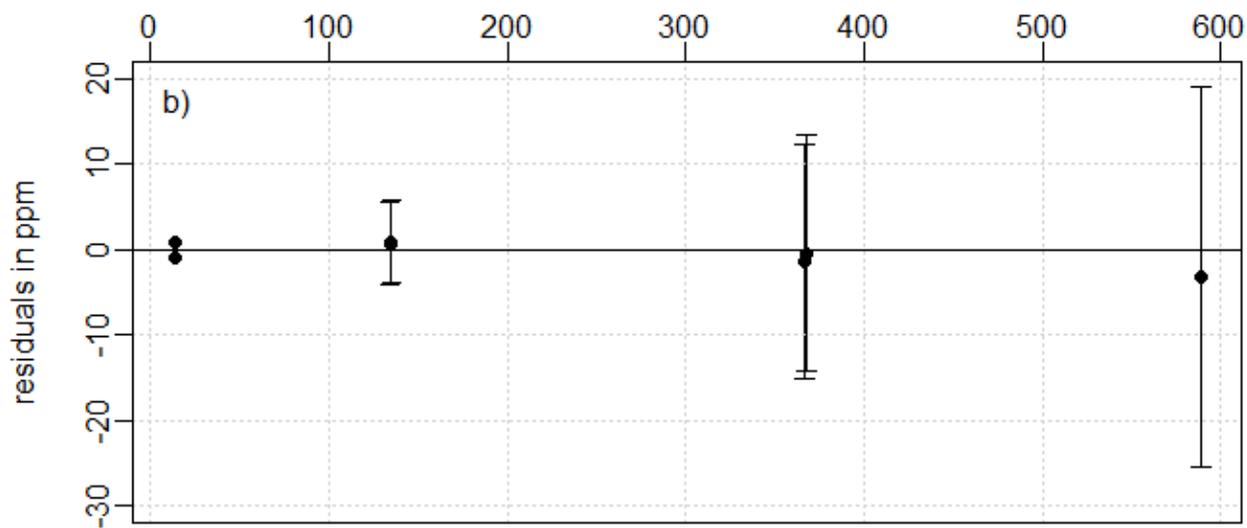
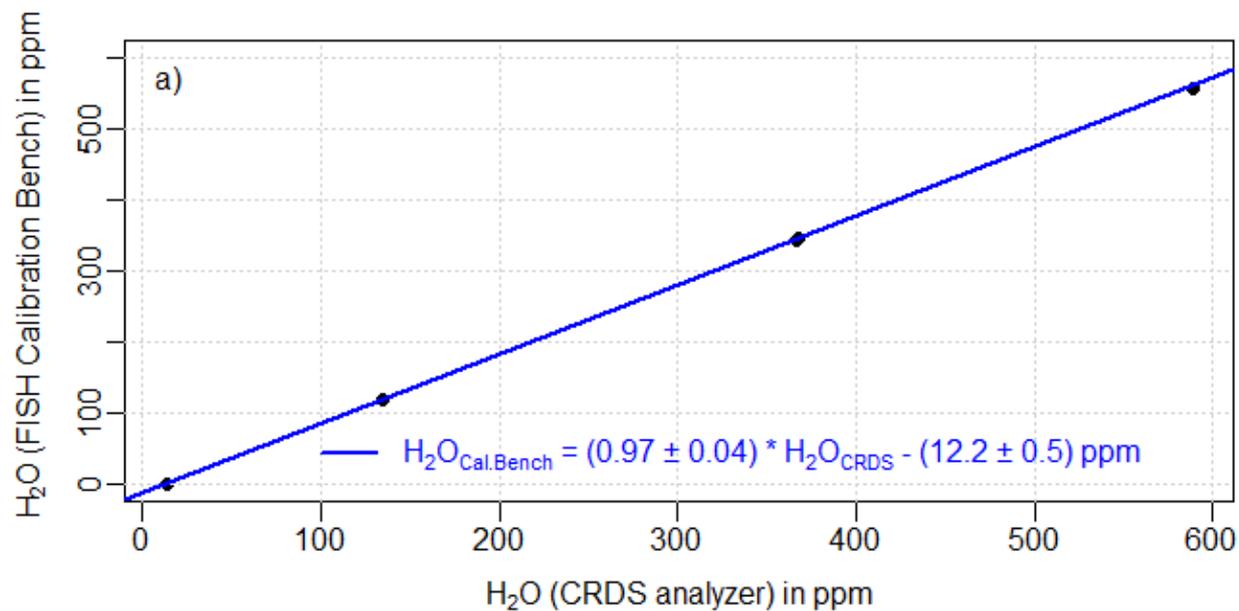


Figure 4: a) Uncalibrated water vapor mole fractions from the two Picarro CRDS analyzers CFADS37 and CFKBDS2003 during a comparison experiment in 2014 (black points). The blue line indicates the quadratic fit of an earlier comparison between the same instruments in 2011. The differences between the two comparisons, thus the drift of the two analyzers over the three years from 2011 to 2014, are shown in plot (b). The relative residuals are plotted on the right axis as light blue triangles. Since the analyzers were not offset corrected before the experiment, the relative difference of 800 % of the data point at 4.5 ppm is less meaningful and is therefore not shown in order to improve the clarity of the plot.



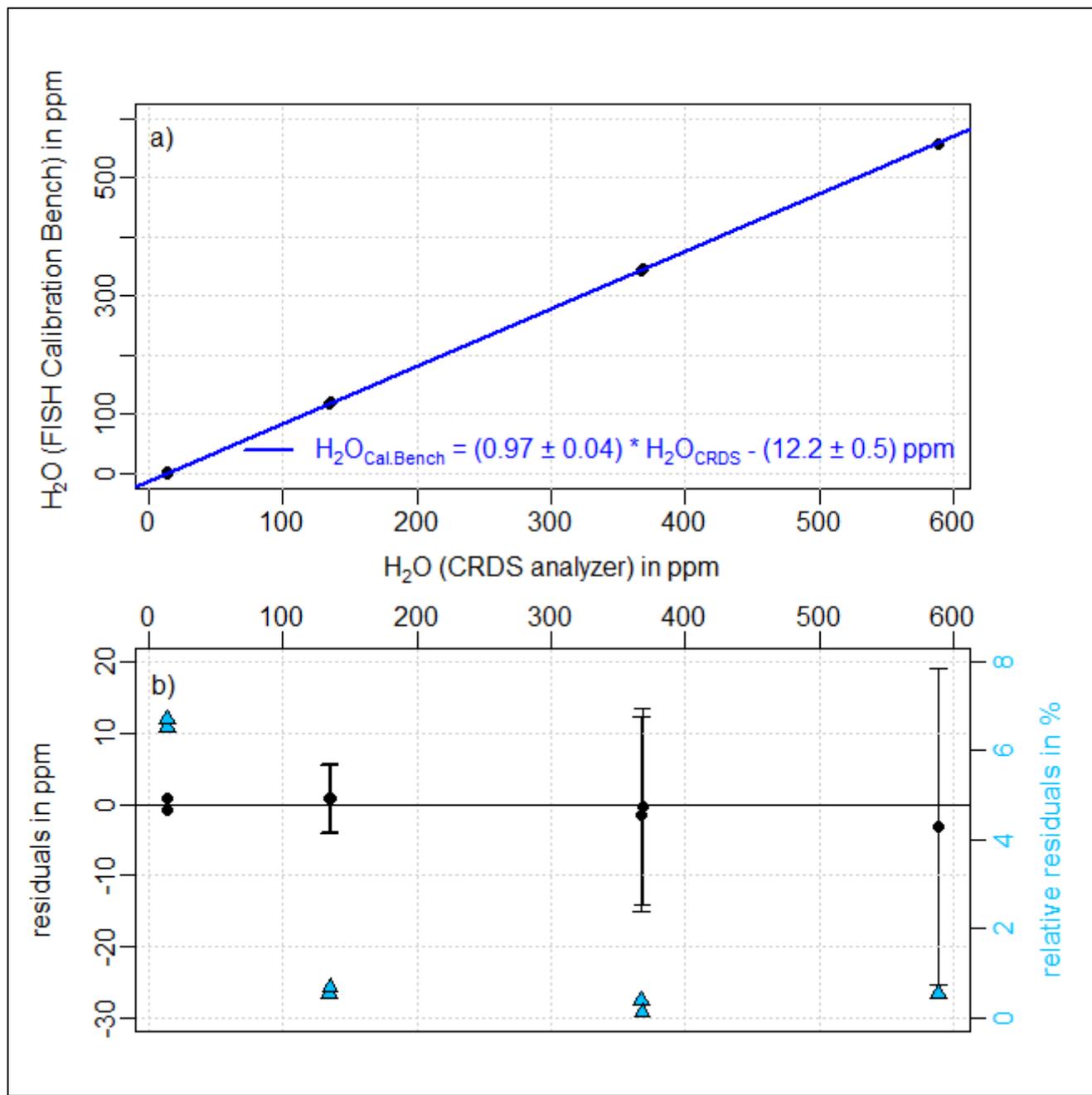
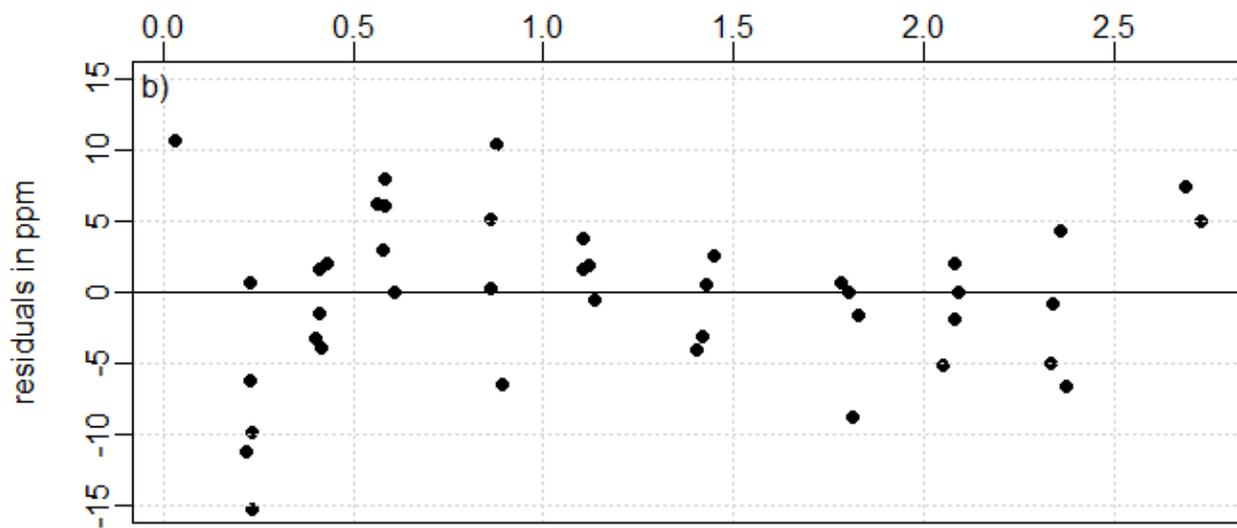
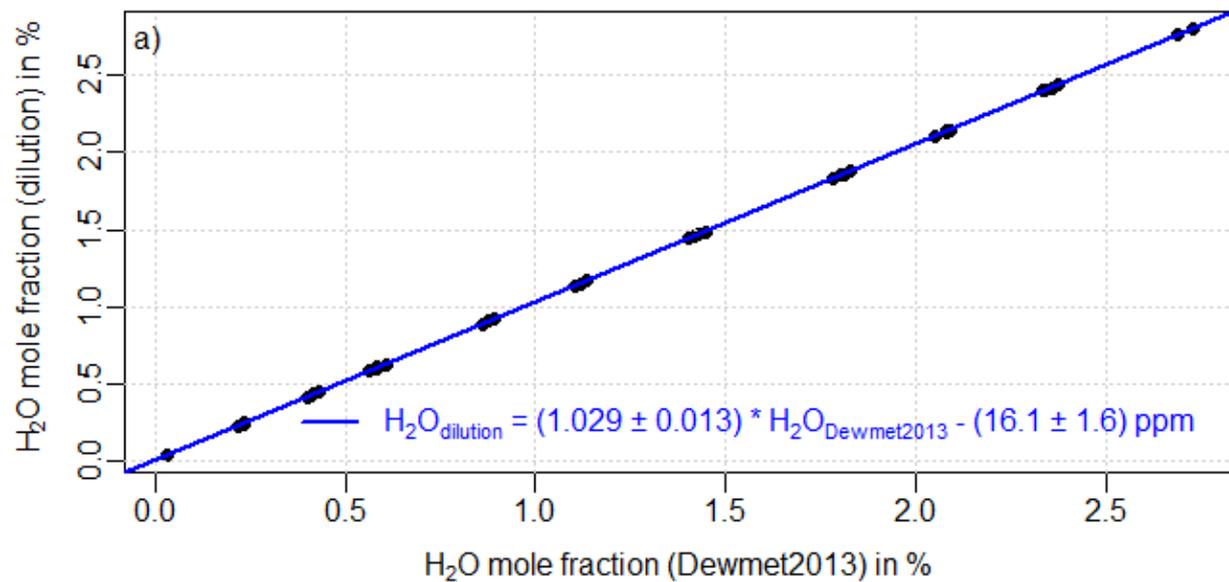
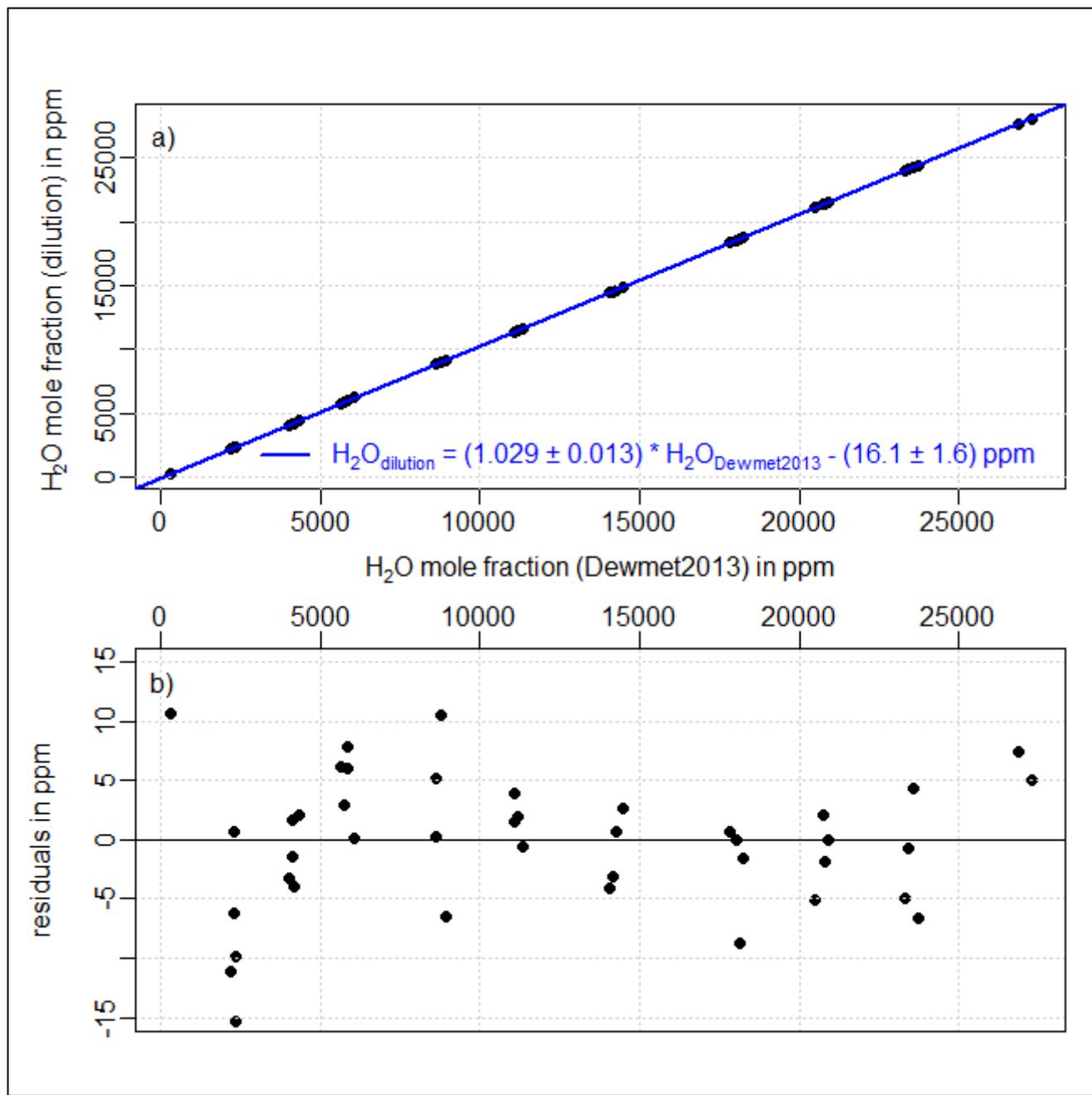
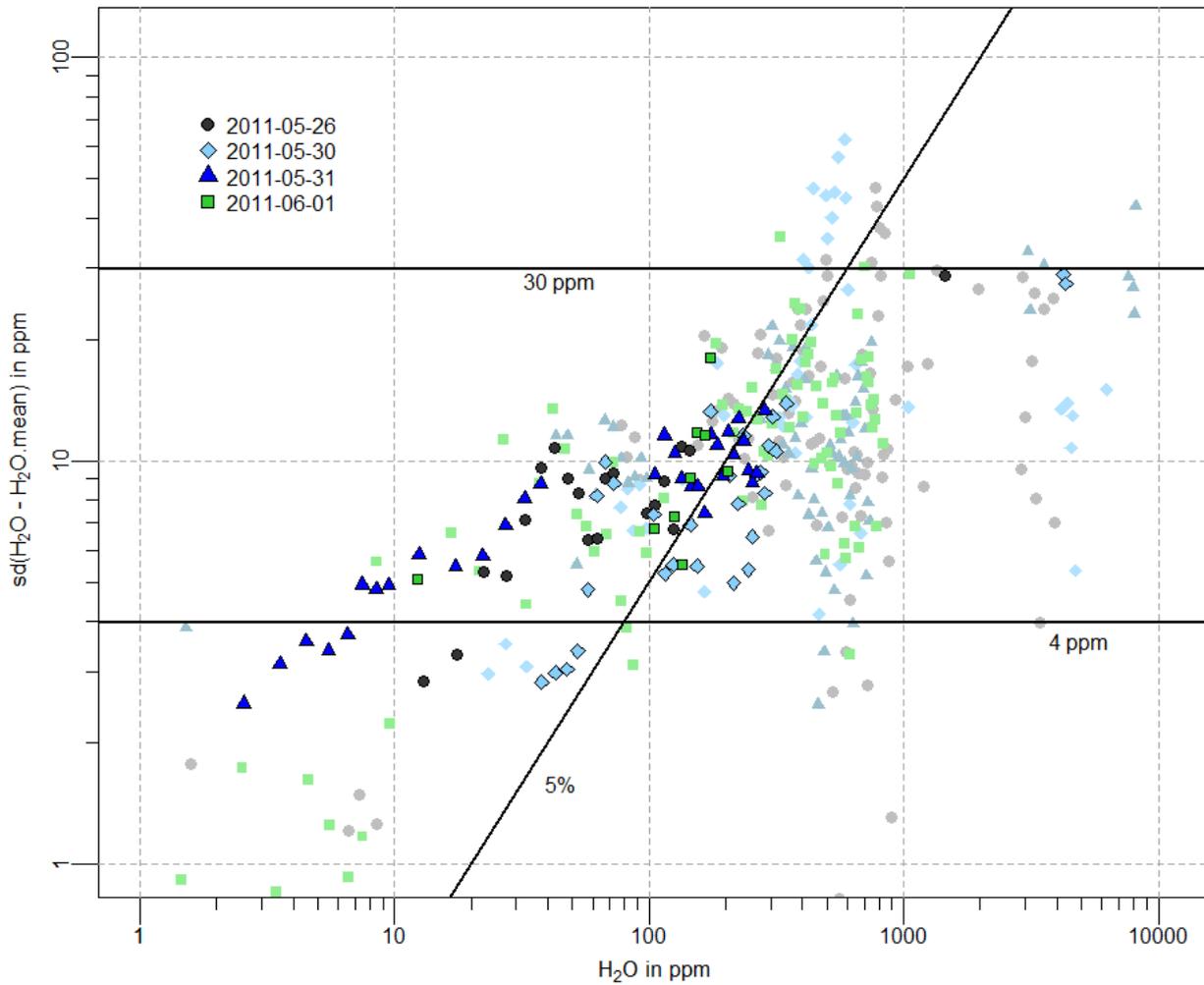


Figure 5: a) Dry air water vapor mole fractions measured by the CFKB2004 CRDS analyzer and the FISH calibration bench during a comparison in 2011. The CRDS data are **offset corrected and** calibrated according to the 2013 comparison against a dew point mirror. The corresponding fit is shown in blue. Residuals (difference between water vapor mole fractions measured by FISH calibration bench and the linear fit) can be seen in the plot (b). Error bars indicate the uncertainty range of the calibration bench of 4 % ~~(rel.)~~. **The relative residuals are plotted on the right axis as light blue triangles.**





5 Figure 6: a) Water vapor mole fraction based on the CO<sub>2</sub> dilution method plotted against the water vapor mole fraction measurement effrom the CRDS analyzer (offset corrected and calibrated according to the comparison against a dew point mirror in 2013) during the fine scan experiment. The corresponding fit is shown in blue. Residuals (difference between water vapor mole fraction based on dilution method and the linear fit) can be seen in plot (b).



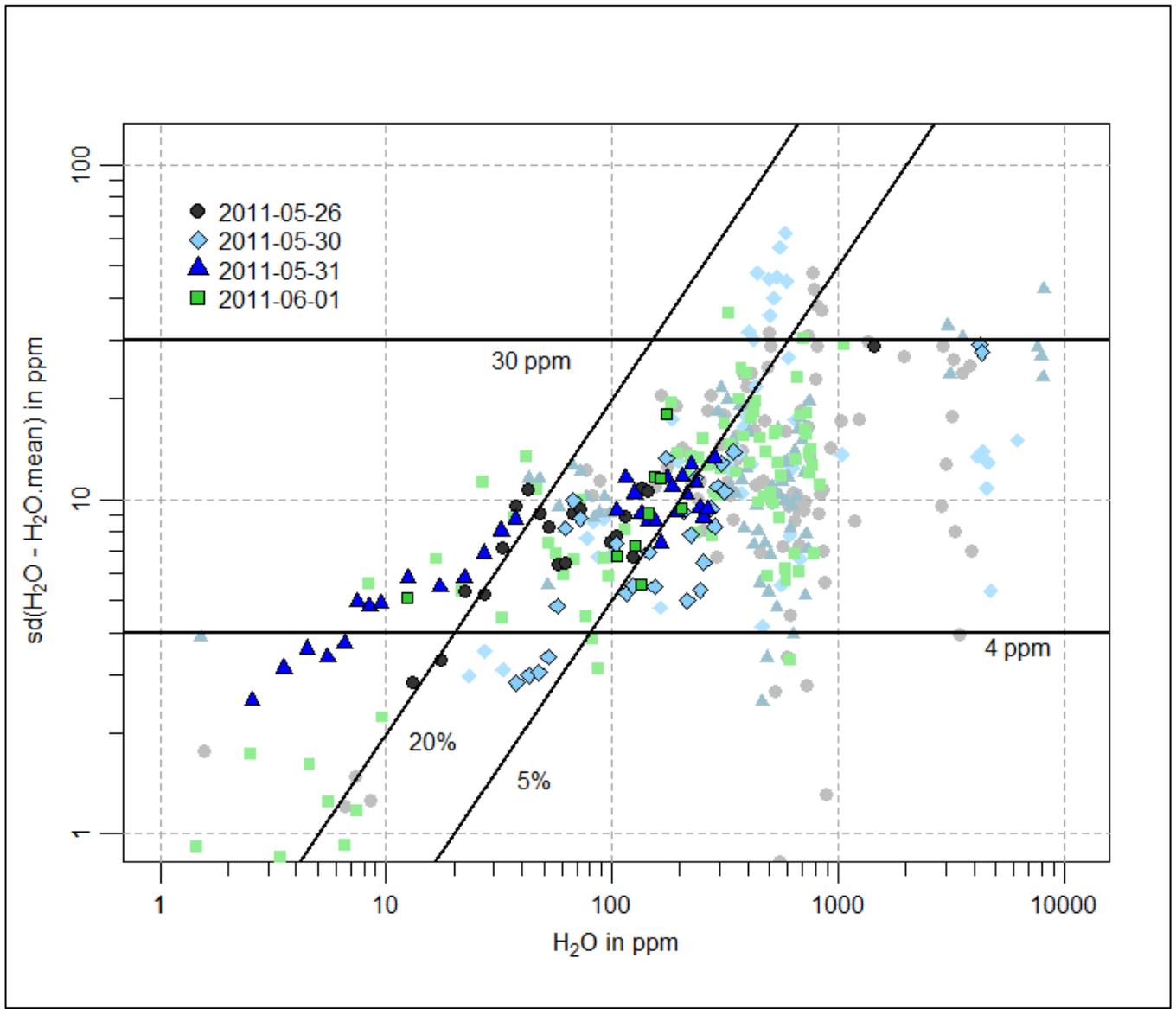
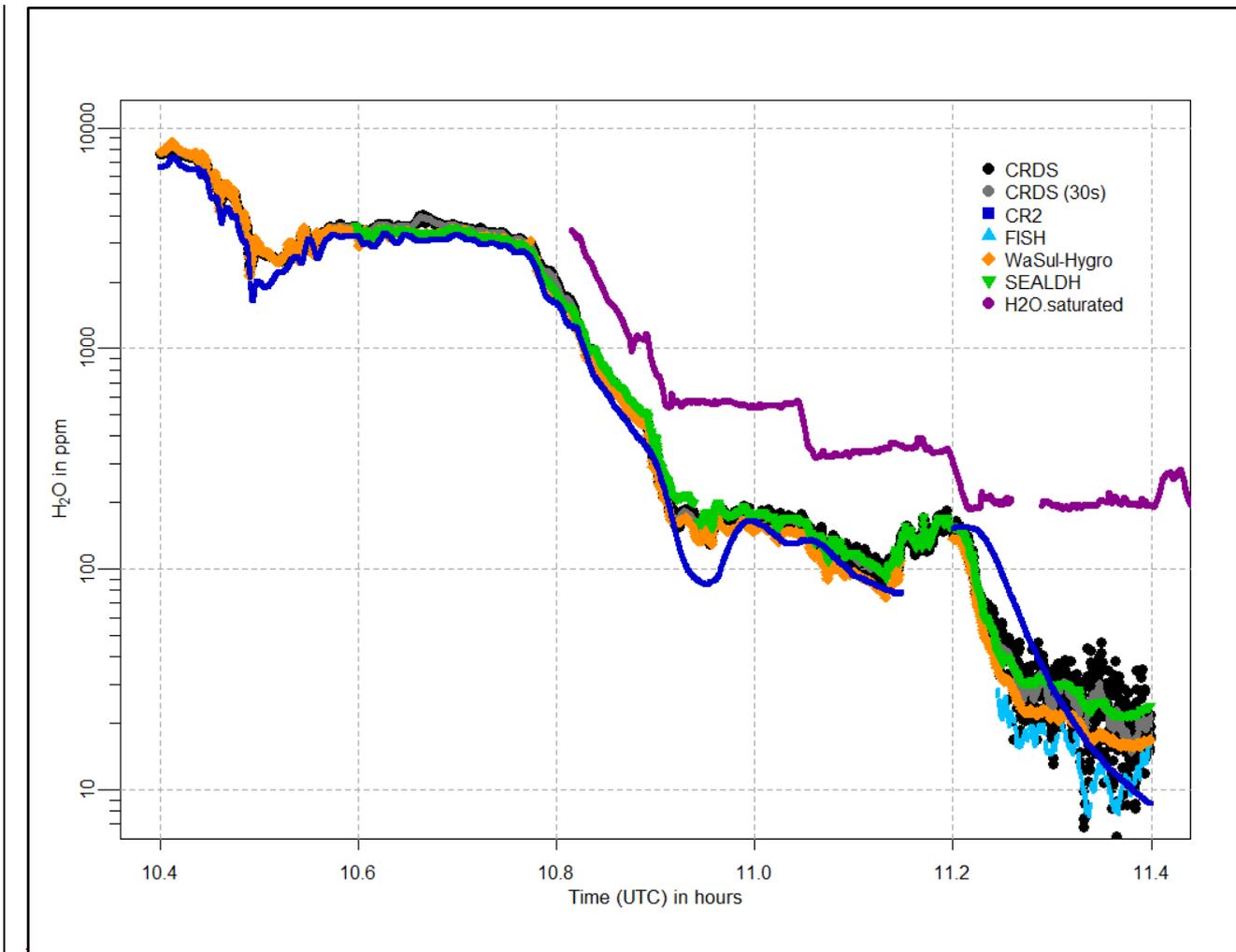


Figure 7: Standard deviation of the difference between the 0.4 Hz CRDS flight data (offset-corrected and calibrated according to the 2013 dew point mirror comparison) and the 60 second averages, averaged for intervals of 1 ppm, 5 ppm, 10 ppm, and 100 ppm water vapor in the corresponding water vapor ranges of 0-10 ppm, 10-100 ppm, 100-1000 ppm, 1000-10000 ppm, respectively. Different colors and symbols indicate different flights. Results with higher priority are highlighted. The horizontal and diagonal black lines indicate standard deviations of 4 ppm, and 30 ppm, and 5 and 20 % (rel.) respectively.



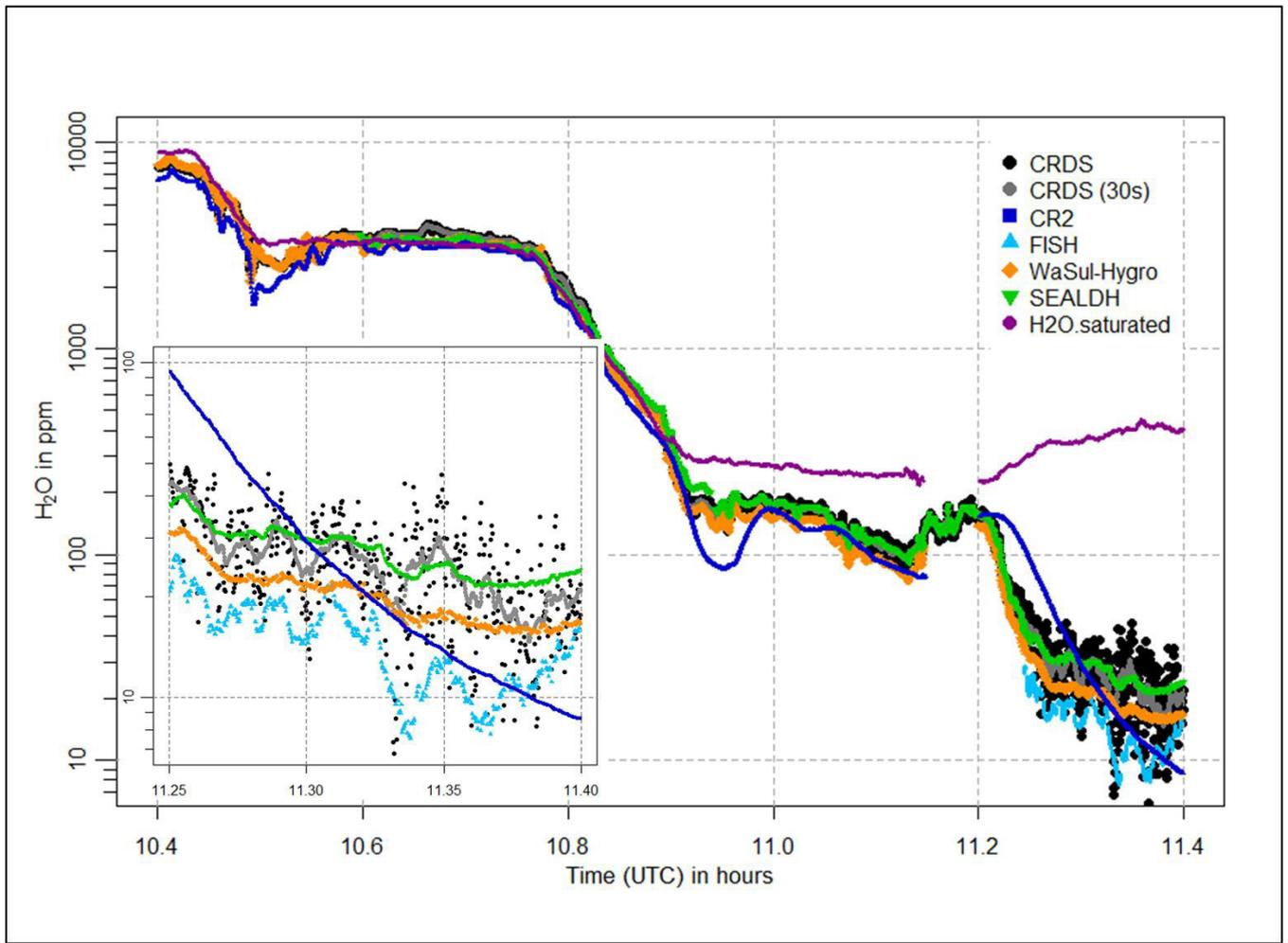
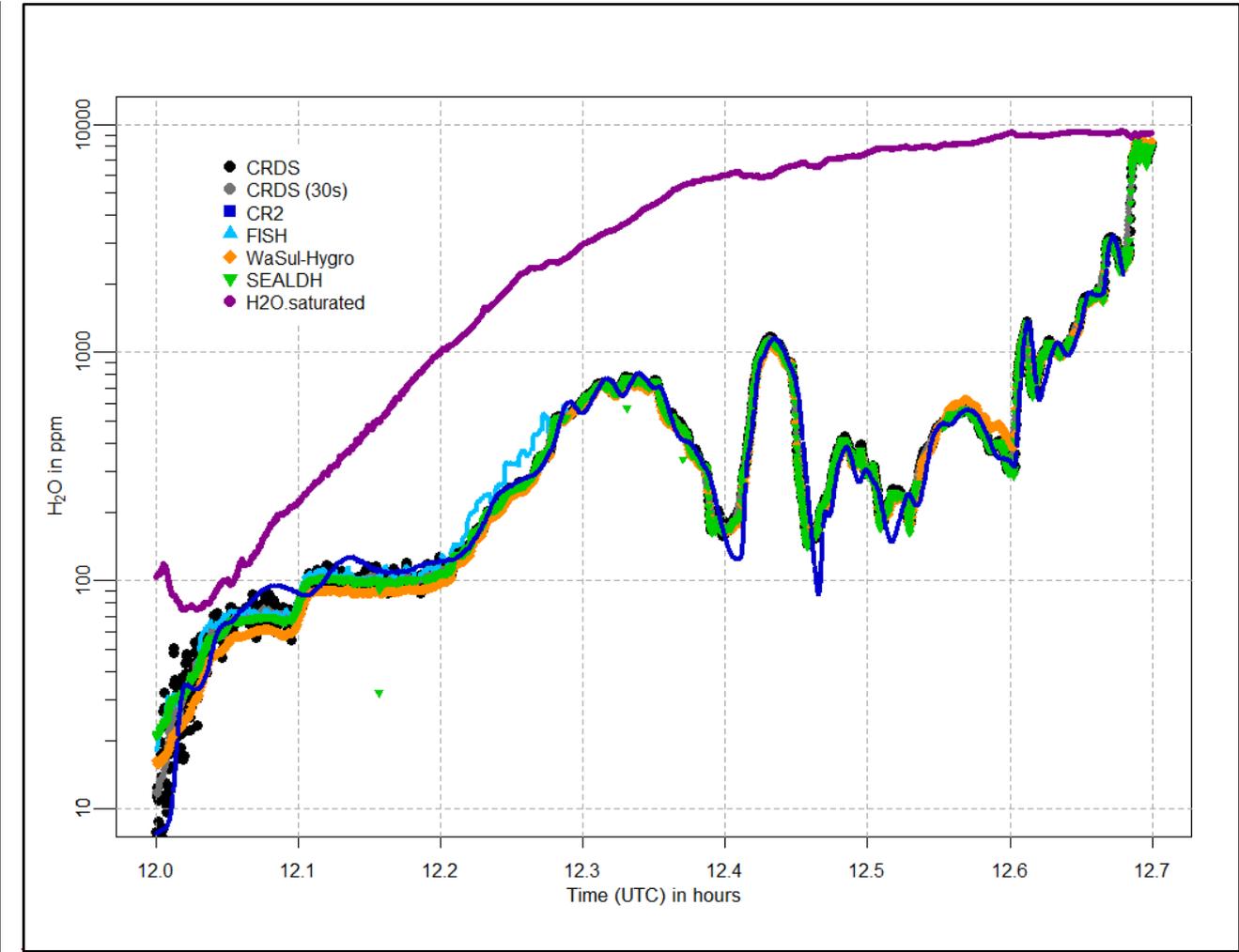


Figure 8: Water vapor mole fractions measured by of the CRDS analyzer (~~offset corrected and calibrated according to the 2013 dew point mirror comparison~~) (black points, 30 seconds mean as grey points), the CR-2 (dark blue squares) and FISH (light blue triangles) instruments, as well as the WaSul-Hygro (orange diamonds) and SEALDH-I (green triangles) analyzers, for a time period during the flight on 31 May 2011. An enlarged section for the time period from 11:15 to 11:24 is shown in the lower left part of the plot.



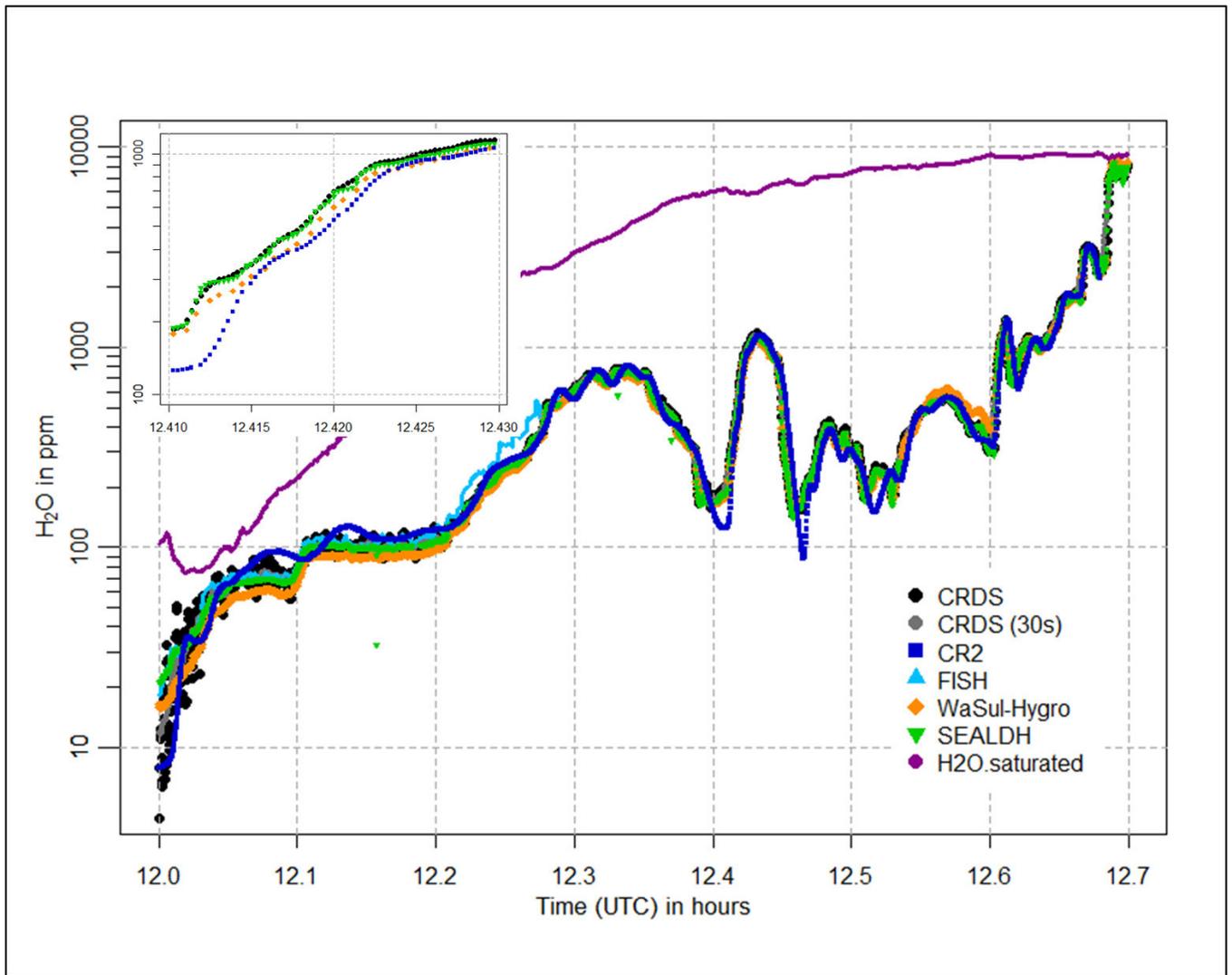


Figure 9: Same as Fig. 8, but for a time period during the flight on 1 June 2011. An enlarged section for the time period from 12:24:36 to 12:25:48 is shown in the upper left part of the plot. Here the water vapor increases in one minute about an order of magnitude from 200 to 1200 ppm.

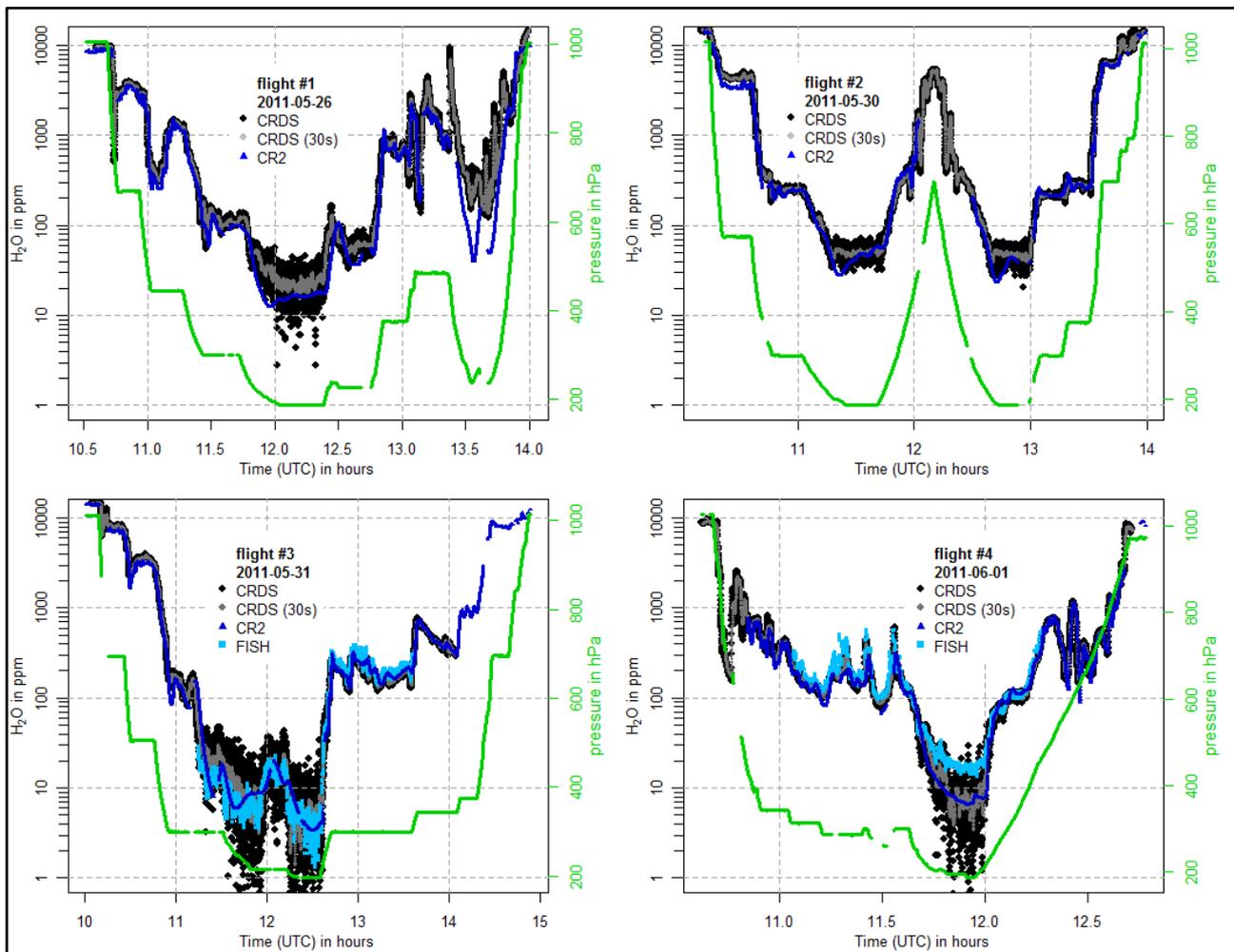
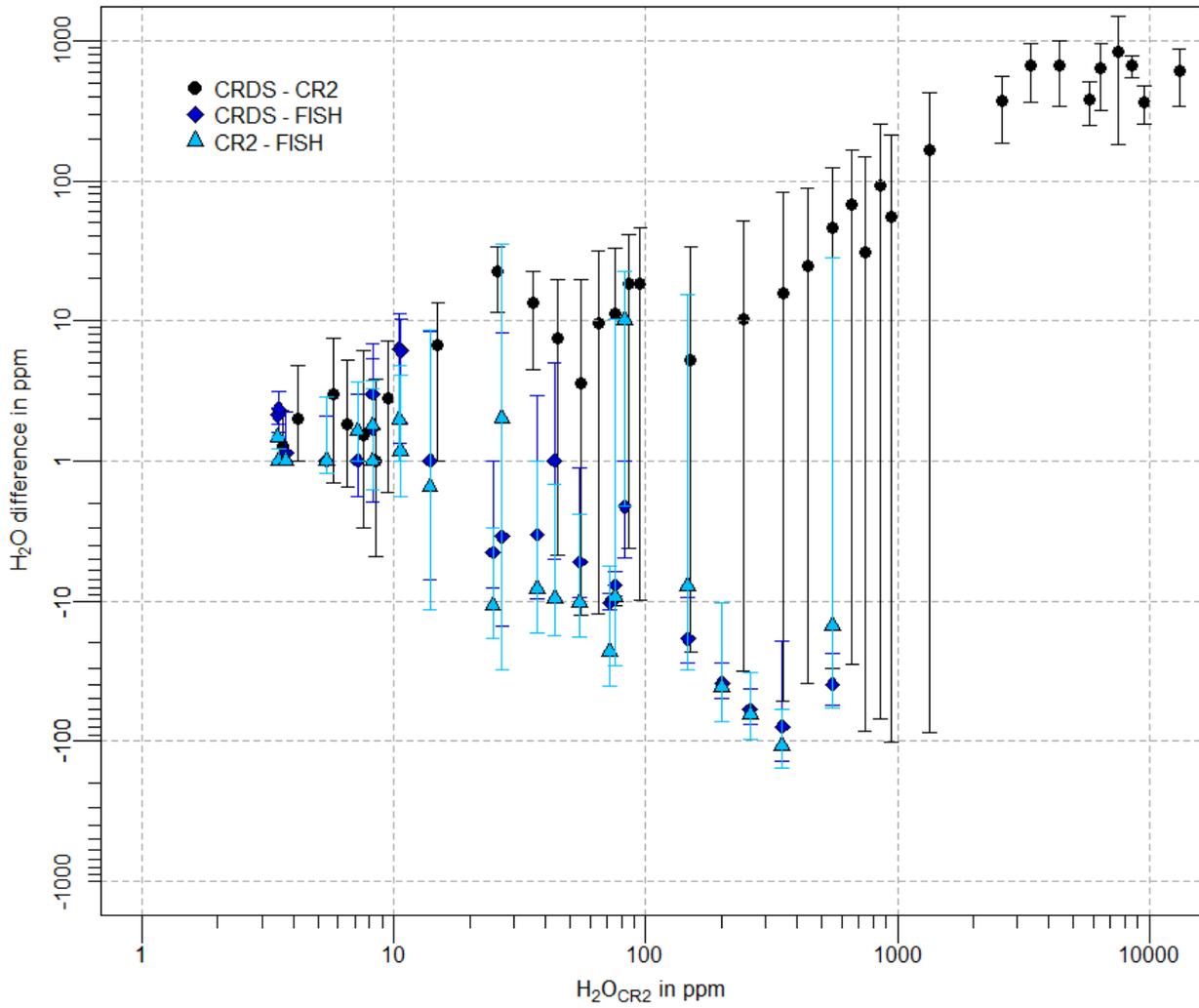


Figure 10: In-flight water vapor data **from** the CRDS analyzer (in black, 30 seconds mean in grey) and the reference instruments CR-2 (in dark blue) and FISH (in light blue) for the four flights on 26 May, 30 May, 31 May, and 1 June 2011. The corresponding atmospheric pressure levels are shown in green. The CRDS data are offset corrected and calibrated according to the 2013 dew point mirror comparison.

5



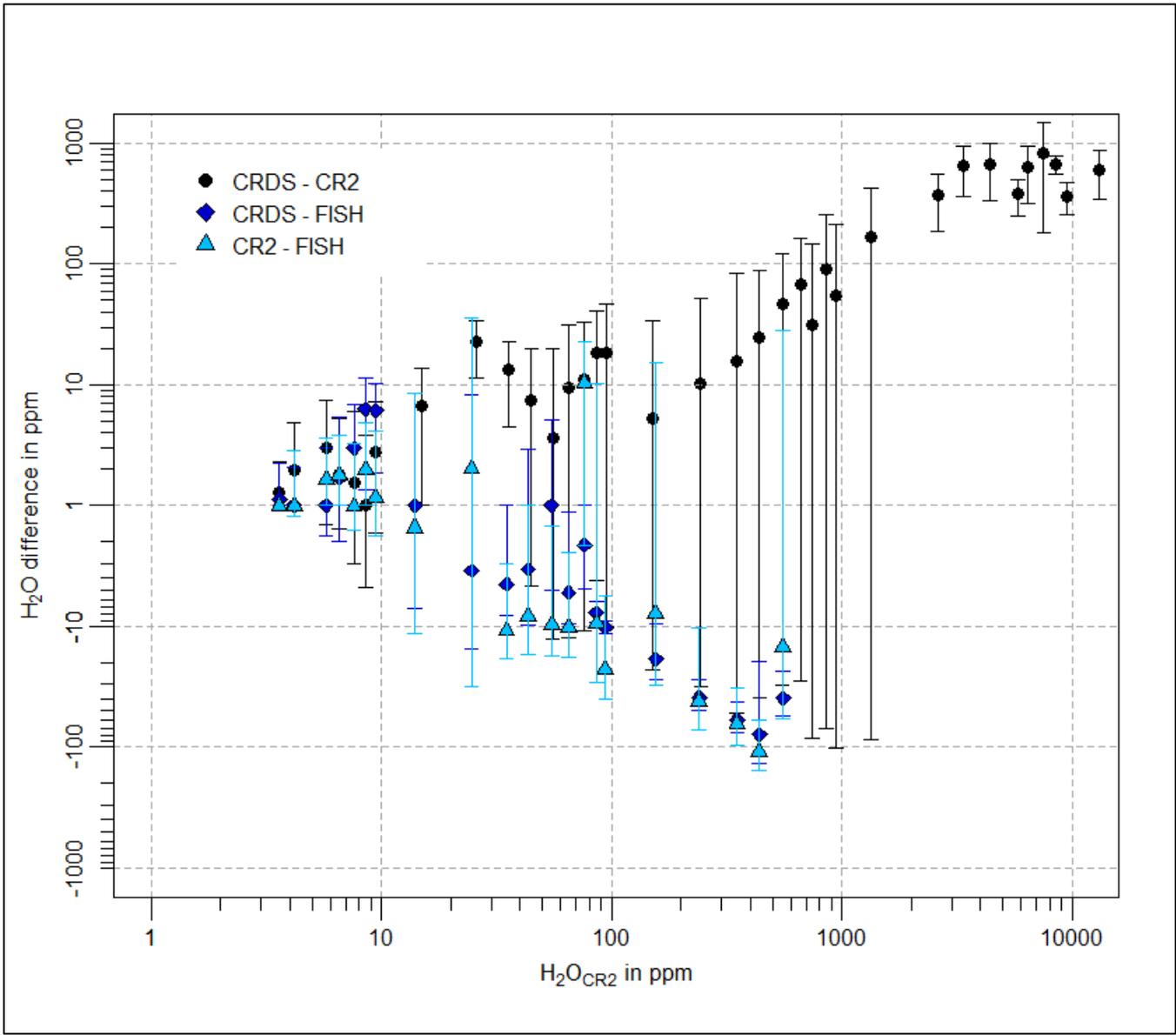


Figure 11: **Absolute differences** of the 30 seconds mean CRDS and CR-2 in-flight data (black points), CRDS and FISH data (dark blue diamonds), and CR-2 and FISH data (light blue triangles) averaged over intervals of 1 ppm, 10 ppm, 100 ppm, 1000 ppm, and 10000 ppm water vapor in the corresponding water vapor ranges of 0-10 ppm, 10-100 ppm, 100-1000 ppm, 1000-10000 ppm, and >10000 ppm, respectively, against the CR-2 water vapor flight data. Error bars indicate the standard deviations of the average differences. **The water vapor measurements of CR-2 are chosen as x-axis, because they cover all flights in contrast to the FISH data. For plotting reasons all differences <1 ppm are set to 1 ppm.**

5

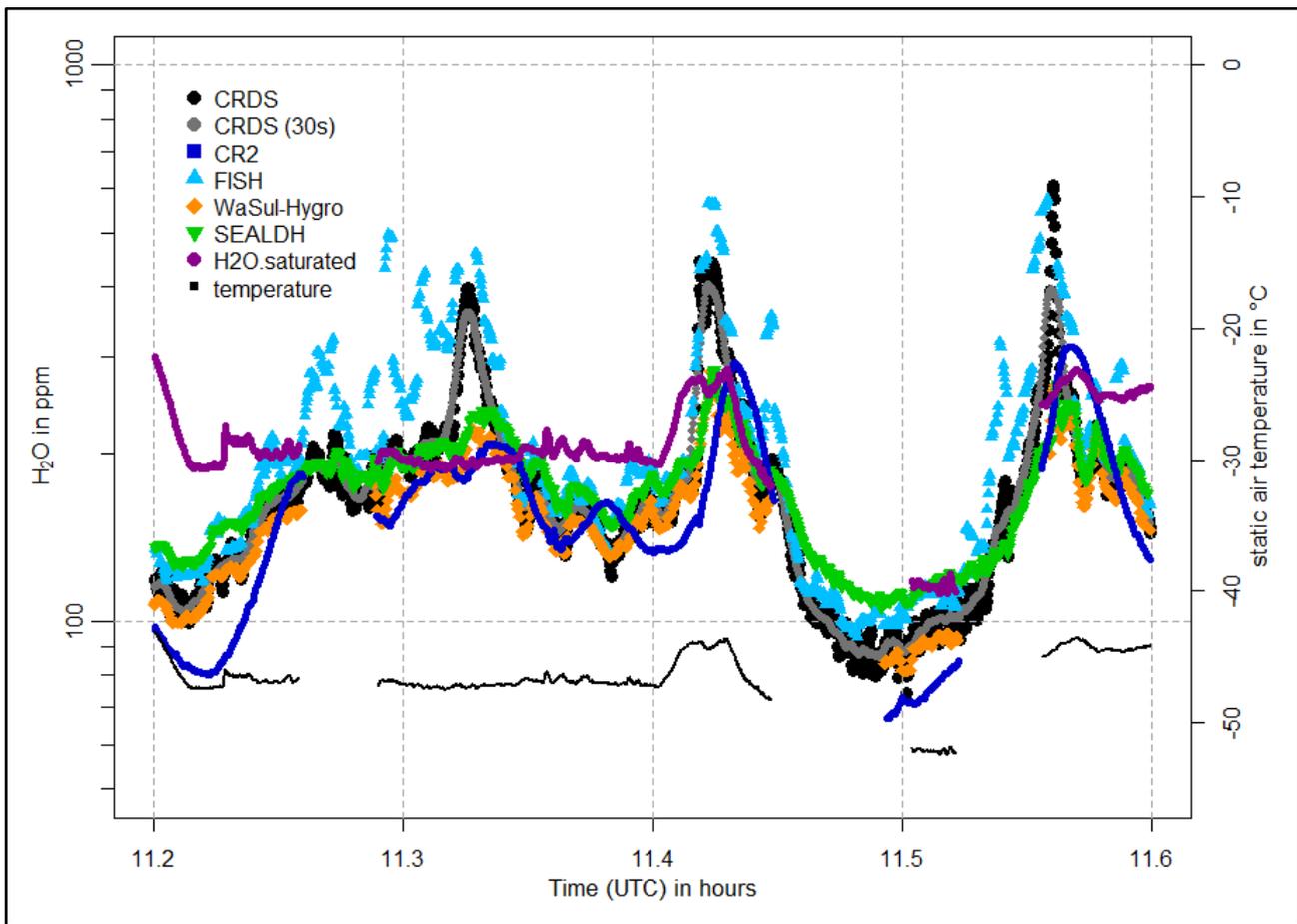


Figure 12: Water vapor mole fractions offfrom the CRDS analyzer (~~offset-corrected and calibrated according to the 2013 dew point mirror comparison~~) (black points, 30 seconds mean as grey points), the CR-2 (dark blue squares) and FISH (light blue triangles) instruments, as well as the WaSul-Hygro (orange diamonds) and SEALDH-I (green triangles) analyzers, during the flight on 1 June 2011, in the presence of clouds. Also shown are the water vapor mole fraction corresponding to saturation (violet points) and the static air temperature (black line).