Reply to the comments of reviewer #1 of "UK greenhouse gas measurements at two new tall towers for aiding emissions verification"

Ann R. Stavert¹, Simon O'Doherty², Kieran Stanley², Dickon Young², Alistair J. Manning³, Mark F. Lunt⁵, Christopher Rennick⁴ and Tim Arnold^{4,5}.

¹ Climate Science Centre, CSIRO Oceans & Atmosphere, Aspendale, VIC, 3195, Australia

² School of Chemistry, University of Bristol, Bristol, BS8 1TS, United Kingdom

³ Met Office, Exeter, Devon, EX1 3PB, United Kingdom

⁴National Physical Laboratory, Teddington, Middlesex, TW11 0LW, United Kingdom

⁵ School of GeoSciences, University of Edinburgh, Edinburgh, EH9 3FF, UK

General comments:

A full uncertainty analysis is evidently planned for a future paper, but more work is needed in the current manuscript to summarize the various contributions to the uncer- tainty of the datasets described here. This manuscript provides many details about systematic errors related to water vapor, but a concise summary of the impacts on the measurements is lacking. Ideally the dataset would include time and site dependent es- timates of this uncertainty component depending on the method of drying/water correc- tion employed. Other uncertainty components such as larger errors for measurements outside of the calibration range should also be reported for individual measurements.

Although some of the uncertainties discussed here are significant compared to the WMO internal reproducibility guidelines, these errors/uncertainties are likely very small relative to the so-called observation (model-data mismatch) errors assigned in the in- verse modeling. Some context about how the measurement uncertainty compares with model-representation errors would be helpful. It would also be helpful to see how the measurement uncertainty compares to signals of emissions. The WMO extended measurement compatibility goals should also be noted (+/- 0.2 ppm for CO2, +/- 5 ppb for CH4). The statistical analysis of the time series data needs major improvement. The Thoning fit is not a good choice for this dataset, as is clearly evident in Figure 5. Specific suggestions are provided below.

The inversion analysis as currently presented is not compelling due to only minor re- ported improvement in total uncertainty and apparent flaws in the inversion framework that cause uncertainties in some regions to increase with additional data. Since AMT is not suitable for a detailed discussion of the inverse modeling methodology perhaps bet- ter to omit. A simpler presentation of how the additional sites improve the sensitivity to surface fluxes could be substituted (i.e. a map of the total surface sensitivity/footprints estimated by the NAME model showing the impact of the additional sites). Discussion of the lab water vapor tests is very hard to follow and should be reorganized and significantly shortened. Specific suggestions are given below. It is not necessary to exhaustively present results from experiments that were inconclusive in the body of the paper. Although researchers who are struggling with similar issues might benefit from this information, it should be relegated to the Supplement or to an Appendix in order to simplify the main paper. Despite having direct experience with analyzing results from these types of water corrections, I found the presentation difficult to understand.

The authors would like to thank reviewer 1 for their helpful comments. We have endeavoured to incorporate all the suggestions that they have made. The locations of changes to the paper are given as Page Number, Line number of the new paper while the reviewers comments contain the locations in the original paper.

As requested by the reviewer a summary of the errors associated with the drying and water correction methods used at the Heathfield and Bilsdale sites is now included (see Table 5). Estimates of the error associated with making observations outside the range of the calibration suite has also been included (Page 14, lines 13 - 23). Discussion of these errors relatively to the WMO extended compatibility goals, model-data mismatch error and the magnitude of emissions signals have also been included (Page 28, lines 20 - 26). As advised by the reviewer we have significantly altered the statistical analysis, no longer

using the Thoning et al. approach and instead using box plots of linearly detrended data. References to specific changes are included in the replies to the specific comments below. Also, as suggested by the reviewers, due to the measurement focused scope of AMT we have removed details of the inversion modelling and its results. These will be discussed more fully in a later paper.

Following the advice of both reviewers we have removed a large section of the text relating to the first set of Nafion[®] drying tests and the accompanying figures. As such some technical corrections are no longer applicable, these have been noted as such. We have also simplified and reordered the discussion of the water correction work. The paper has been significantly restructured with the Nafion[®] and water correction work no longer a separate section. Instead they are incorporated in to the main experimental and discussion sections. Even with the inclusion of the extra information requested by the reviewers these changes have reduced the length main body of the paper by 6 pages and removed 4 figures.

Specific comments:

Page 1, line 27/28: "...this error is mostly calibrated out" is the 0.02 umol per mol error the remaining error after applying the calibration? Or is the nafion-related error « 0.02 ppm after calibration? In either case 0.02 ppm is nearly negligible and likely smaller than the total measurement error of the analytical system, which has not been adequately characterized. The 0.02ppm error is the error prior to carrying out the calibration. We expect the error following the calibration to be ≪ 0.01ppm. This has been clarified in the text. It is important to note that this error is only removed through calibration because the calibration cylinders are also passed through the Nafion® which may not be the case for all monitoring sites. As such we feel it is important to highlight and quantify this possible source of error. We would politely disagree with the reviewer, and suggest that a 0.02ppm bias is far from negligible - a bias of this magnitude represents 40% of the WMO Southern Hemisphere inter-laboratory compatibility guideline. As the reviewer correctly notes, other measurement errors associated with an analytical system are most likely to be larger, hence, in conjunction with this bias they could well lead to a significant error within a data set.

As outlined in the text, a separate paper currently in preparation will characterise the measurement error of the analytical system of these two new sites along with the other four sites within the UK GHG monitoring network.

Page 2, line 25: does AMT allow references to a paper in prep?

We have removed this reference.

Page 6, line 18: CRDS dwell times at each level are surprisingly long

The length of the CRDS sampling period at each height was designed to obtain as long as possible period of continuous data at each height during an hour. That is, for the site with three intakes 20 minutes was spent at each height during each hour. Such a sampling strategy is not unusual for greenhouse gases, for example the INFLUX experiment measured from some intakes for up to 40 minutes of every hour (Richardson et al., 2017).

Page 7: line 1, "This air is dried to <0.005%..." I think this refers to the counterpurge air but not totally clear at this stage if it might refer to the sample air. It would be useful to state what is level of drying that is accomplished with the nafion for the GC channel. We have clarified and added extra information to the text.

Page 6, line 21 - 23, "The counter purge is dried to < 0.005 % H₂O by the compressor (50 PLUS M, EKOM, Slovak republic) and a gas generator for total organic carbon systems (TOC-1250, Parker Balston, USA). Previous examinations of this drying method have found that samples are dried to < 0.0002 % H₂O (Young, 2007)."

Page 7: droplet test has weaknesses due to rapid changes of humidity that are inadequately resolved. Potential mismatchs/lags among co2/ch4/h2o channels. Also, I am not sure that Yver Kwok et al. 2015 is the best reference for this. I quickly checked and did not see any discussion of the droplet test in that paper. Maybe it would be better to cite the Rella 2013 AMT paper which describes several implementations of the droplet method. A citation for the Rella 2013 paper is currently lacking.

We thank the reviewer for highlighting this oversight. We have replaced the reference to Yver Kwok et al. 2015 with a reference to Rella 2013 (Page 7, line 20).

Page 7, line 15: A cylinder of air was not humidified. Instead, air from a cylinder was humidified.

Corrected

Page 8: data collected in the first 5 minutes following the injection was excluded. What is the maximum humidity sampled after these data have been excluded? (i.e. is the maximum h2o value included in the fits significantly lower than the 4.5% value mentioned on page 7, line 20?) It is frustrating & confusing that the water corrections are discussed in multiple sections of this paper (here on page 8 and again in section 3.3.1).

The text has been clarified to include

Page 8, line 8, "This 5-minute cut-off reduced the maximum H_2O value included in the fit to 4 % H_2O ."

Following the restructuring of the paper, the water correction work is now split into two components, an experimental (Sections 2.3.4, 2.3.5 and 2.3.6) which outlines how the water correction tests were conducted and results & discussion (Sections 3.3 and 3.4) which details the results of the experiments and their implications.

Page 8, line 22: What are the calibration gases spiked with? The below has been added to the text.

Page 15, lines 4 -10, "The cylinder spiking and filling techniques of the calibration cylinders varied. The Heathfield calibration suite and the second Bilsdale calibration suite were filled at GasLab MPI-BGC Jena and consisted of natural air spiked using a combination of pure CO₂ and a commercial mixture of 2.5 % CH₄ and 0.5 % CO in synthetic air. The "high" calibration cylinder of the first calibration suite used at the Bilsdale site was filled with peak-hour ambient air at EMPA, Dübendorf, Switzerland while the "low" and "mid" cylinders were based on Mace Head air, in the case of the "low" this was diluted with scrubbed natural air."

Page 9, line 1: It would be helpful to specify what is the maximum systematic error due to differences between sample and standard isotopic composition, since spiking standards can result in isotopic compositions that are outside the ambient range.

We cannot comment directly on our measurement error as we do not have measurements of the isotopic composition for our calibration gases. However, we refer readers to other papers who investigate this topic in detail and give an estimate based on the literature. Page 15, lines 13-19, "The effect of an isotopic mismatch between the calibration standards and the sample has been examined in detail by Flores et al. (2017), Griffith (2018) and Tans et al. (2017). With Griffith (2018) showing that, for a sample of 400 µmol mol⁻¹ CO₂ and 2000 nmol mol⁻¹ CH₄, the error will range between 0.001 – 0.155 µmol mol⁻¹ CO₂ and 0.1 – 0.7 nmol mol⁻¹ CH₄ depending on the magnitude of the sample to standard mismatch. As such, we expect a worst-case scenario estimate of the error associated with our measurements < 0.04 % for both CO₂ and CH₄."

Page 9, line 5: What is the uncertainty associated with the non-linearity correction? Dynamic dilution is a complex procedure and likely to have non-negligible uncertainties. Why not use a set of gravimetric standards instead?

These non-linearity corrections are typically exceptionally stable with Hammer et al. (2008) finding differences of < 0.2 nmol mol⁻¹ N₂O between non-linearity corrections determined 10 years apart and over the range of interest (326 to 340 nmol mol⁻¹ N₂O) highly linear. While, studies including Hall et al. (2011) and van der Laan et al. (2009) have found SF₆ to be very linear at mole fractions > 4 pmol mol⁻¹ and for the instrumental response function to be very stable over time. As such, we expect the uncertainty of the non-linearity correction to be small. Although not discussed in detail in this paper this uncertainty will be addressed in the associated uncertainty analysis paper as outlined in the introduction.

Page 2, lines 23 – 26, "A further paper, currently in preparation, will discuss the integration of these new sites with the existing UK Deriving Emissions linked to Climate Change (DECC) network (Stanley et al., 2018) funded by the UK Department of Business, Energy and Industrial Strategy (BEIS) and provide a full uncertainty analysis for data collected at all the DECC/GAUGE sites."

The UoB group, along with other groups (e.g. the Schauinsland GAW station) have a history of successfully using dynamic dilution for the quantification of the N₂O and SF₆ non-linearity. The laboratories at the BSD and HFD sites are physically small making it difficult to store or use a set of dedicated gravimetric N₂O and SF₆ standards at each site and shipping a set of cylinders between the sites would be logistically difficult considering the remote locations of and limited access to the sites. In contrast, the dynamic dilution method is able to generate multiple calibration points using just two cylinders and is logistically simple to deploy and ship between sites. As such, the decision was made to use this approach. We have also added the information below

Page 15, line 23 – Page 16, line 3, "This approach, dynamic dilution, has a history of use in similar field locations (Hammer et al., 2008) and is able to generate multiple calibration points using just two cylinders. This greatly reduces the number of cylinders needed, a key concern for space-limited locations like BSD and HFD. An assessment of the uncertainty associated with this non-linearity approach will be included in a future paper currently in preparation. However, previous studies (Hall et al., 2011;van der Laan et al., 2009;Hammer et al., 2008) have found the ECD detector response to be extremely stable over time and very linear for both SF₆ and N₂O in the mole fraction range typical of the HFD and BSD stations. As such, we expect the uncertainty of the nonlinearity correction to be very small."

Page 10, line 9: It is not clear how the long-term repeatability numbers here are being computed. The parenthetical description (xbar minus +/- one sigma) is not an adequate description. Are these numbers the mean standard deviation computed over all the tanks over a year? In any case, instead of "long-term repeatability" it would be better to report the "long-term reproducibility", since a metric of the compatibility of the measurements over periods of months to years is needed. These terms are defined in the Guide to the expression of uncertainty in measurement:

https://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf The difference between repeatability and reproducibility has to do with whether the conditions of the measurement are changed, and over timescales of months to years, standards are changing, ambient conditions, humidity levels are changing, etc. It seems unlikely that one could confidently interpret differences in CO2 in measurements made months or years apart at the level of 0.018 or 0.013 ppm.

We thank the reviewer for identifying the incorrect use of repeatability. We have now corrected this to reproducibility throughout the text.

A reproducibility of minute mean values at the level of 0.02 μ mol mol⁻¹ CO₂ over a period of years would indeed be unlikely however we are reporting the reproducibility of a 20-minute mean value which is much more stable. We have more clearly defined how these numbers have been calculated. See

Page 17, lines 3 - 4, "The long-term reproducibility of a 20-minute mean was estimated as the mean standard deviation of the daily 20-minute measurements of 4 standard cylinders at each site."

Page 10, line 21: How was the cycle time threshold of 8 sec determined? Are those data that are filtered based on cycle time obviously bad?

The 8-second cycle time threshold was determined after a close examination of the typical cycle times and their relationship to other key parameters. Longer cycle times were typically related to periods of more variable cavity pressure and outlet valve parameters. Measurements of cylinders with long cycle times were also typically significantly different, between 0.5 and 50 μ mol mol⁻¹ CO₂ from the mean cylinder value.

Page 11, section 2.2.2: The Thoning fits are most appropriate for remote sites, and the method often produces spurious results when gaps or large pollution events are present. These effects can be clearly seen in Figure 5. Figure 5 does not add any value to the paper and should be removed. More careful analysis is needed if seasonal cycles and trends are to be reported for these sites.

The Thoning fits and Figure 5 have been removed from the paper and replaced with a seasonal cycle analysis using box plots and simple trend estimate based on a linear fit to the data.

Page 18, Lines 2-6, "The long term trend in mole fraction at each site was estimated as the mean linear trend in the minute mean data over the period 2014-2017, inclusive. Seasonal and diurnal trends in the data were assessed using monthly and hour-of-day box plots of hourly means of detrended minute-mean data developed using the Python matplotlib.boxplot package. Here the long-term trend was removed by using a least-squares fit between a quadratic and the minute mean data."

Page 12, line 19: Are the HFD mean CO2 mole fraction of 407.5 ppm and the BSD mean of 404.7 the means over each entire record? If yes, then this quantity will seemingly be affected by the gaps and there is no use in reporting this quantity or interpreting the difference between the sites.

We thank the reviewer for noting this. References to mean mole fractions calculated in this manner have been removed.

Page 12, line 25: The text implies that there are some high CO2 events at BSD that do not have associated high CO. If that is the case then any such events are not likely to result from biomass burning and must have a different source (e.g. power plant plume).

We thank the reviewer for this observation and have included it in the text. Page 18, Lines 17 - 19, "While events that do not show corresponding high CO mole fractions, the majority of which occur in the higher two intakes, are likely to be driven by less local CO₂ sources, for example power plants."

Page 13, line 5: It seems unlikely that multi-day CO2 enhancements resulting from pollution over London or Europe would not be associated with elevated CO. Could the elevated CO2 result from advection of air from higher latitudes?

We thank the reviewer for this comment as it demonstrates that our text was not written clearly. We had meant to say that the difference in pollution patterns, short sporadic events vs. prolonged periods of increased background mole fraction, was not common to all gases between the two sites rather than not common between the gases at an individual site. In fact, coincident 2-3 day periods of increased CO, CH_4 , N_2O and SF_6 were observed at HFD. This has been clarified in the text.

Page 18, Lines 20 - 26, "HFD is located in southern England, just south of London (Figure 1). Here, high CO_2 events are typically longer — 2 to 3 days — and coincide with elevated CH_4 , CO, N_2O and SF_6 . Rather than appearing as peaks superimposed on a background value these periods have a positive shift in the entire diurnal cycle, suggesting a change in the background mole fraction. Air histories, based on the output of the Numerical Atmospheric dispersion Modelling Environment (NAME) Lagrangian dispersion model, outlined in Manning et al. (2011), for these periods of elevated CO_2 typically show the source of the air to be from over London or Europe."

Page 13, line 10: The typical diurnal variation of CO2 measured on tall towers is well understood, e.g., https://www.tandfonline.com/doi/pdf/10.3402/tellusb.v47i5.16070?needAccess=true We have clarified our text and included a reference to this paper. Page 19, Lines 1 -6, "Both sites show a clear relationship between CO_2 mole fraction and intake height with the lowest height generally having the most elevated mole fractions, followed by the higher heights (Figures 3a & 4a). This trend, also apparent for CH₄ and CO (Figures 3c & e and Figure 4c & e), is typical of tall tower measurements and is driven by proximity to surface sources (Bakwin et al., 1998). The gradient in CO_2 and CH₄ mole fraction is most apparent during the early hours of the morning (see Figure 3b & d and Figure 4b & d) when the boundary layer is the lowest."

Page 14, Line 15: See comment above. The authors are correct that a major issue with the Thoning fitting routine is the underlying FFT, which requires interpolation across gaps. But since the fits are obviously flawed, why not use alternative methods to investigate the seasonal cycle and trends? A simple analysis using monthly mean or median values would be much more robust and simple to implement and explain. And/or box and whisker or fiddle plots could be used to describe the seasonal cycles and trends. Any months with significant gaps could/should be removed.

All references to the Thoning fit have been removed and replaced with a discussion of seasonal cycles based on box plots.

Page 14, line 20: It is not surprising that CO is decreasing. This has been widely reported. It would be interesting to compare the trends for all of these gases with background values from Mace Head.

The -7 nmol mol⁻¹ yr⁻¹ trend observed in HFD and BSD CO is slightly larger than the -5 nmol mol⁻¹ yr⁻¹ trend apparent over the same period in the CO data collected at Mace Head. However, as HFD and BSD are on a different calibration scale to the MHD it is not possible to determine whether this is a real difference or a calibration artefact. As such, we decided not to include this in the text of the paper. However, as all three sites use common calibration scales for CO₂, CH₄, N₂O and SF₆ comparisons between the growth rates of these gases have been added to the text.

Page 20, Lines 1-7, "Of the 5 gases measured at HFD and BSD, CO is the only gas to show a decrease in mole fraction between 2013 and 2017, roughly -7 nmol mol⁻¹ yr⁻¹. In contrast, the CO₂ and CH₄ data increase by 2-3 μ mol mol⁻¹ yr⁻¹ and 5-9 nmol mol⁻¹ yr⁻¹ respectively, varying on the intake height. These agree well with the ~2 μ mol CO₂ mol⁻¹ yr⁻¹ and ~8 nmol CH₄ mol⁻¹ yr⁻¹ trends observed at Mace Head (MHD, 53.327 °N, -9.904 °E, Figure 1), a remote site within the UK DECC network located on the west coast of Ireland. The CO data collected at MHD is not on the NOAA x2014 CO calibration scale making direct comparisons between growth rates at the three sites meaningless."

Page 20, Lines 20-22, "The long-term trend, ~0.8 nmol N_2O mol⁻¹ yr⁻¹ (calculated using data from the 108 m and 100 m intakes at BSD and HFD over the period of coincident data collection, 2014 to mid 2016) also agrees well between the two sites and with MHD, also ~0.8 nmol N_2O mol⁻¹ yr⁻¹."

Page 21, Linese 17-18, "The long-term trend in the SF₆ mole fraction at BSD and HFD shows a gradual increase of 0.3 pmol mol⁻¹ yr⁻¹ again agreeing well with MHD which showed an identical growth rate."

Page 14, line 26: Again, box and whisker plots showing quantiles of the data would be helpful here to quantitatively describe any CO differences between sites. It is not useful to report the multi-year mean values without any uncertainties.

This has been corrected and a discussion based on box plots included. Page 20, Lines 8-10, "While the range of minute mean CO mole fractions was significantly larger at BSD, 63 to 9500 nmol mol⁻¹ than HFD, 60 to 4850 nmol mol⁻¹, the high CO values observed at BSD were relatively rare. This is reflected in the smaller quantile spread of the BSD data compared with the HFD data (Figure 7e & f)."

Page 15, N2O: There is a more recent Nevison et al paper that discusses N2O seasonality over the US (https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2017GB005759) We thank the reviewer for highlighting this new Nevison paper. However, as our paper, and the original Nevison reference, focuses on the UK while the new paper examines the USA

Page 16, Line 13: SF6 seasonality might be driven by atmospheric transport.

we feel the original paper is the more appropriate reference.

This increase in pollution events of SF6 during the winter months is seen across the combined DECC/GAUGE network (4 sites across the UK but not MHD) and is very intriguing. However, the air history maps for these events do not consistently indicate any one area in the UK or the continent suggesting that this is an emission rather than transport related increase. This has triggered further investigations in the inventory community but no conclusions have yet been made.

Page 21, Lines 21 -25, "This seasonal shift occurs across the wider DECC-GAUGE network and air history maps suggest that it is not associated with an obvious UK or continental region. As such, instead of an atmospheric transport driven shift we believe this to be a true change in emissions and hypothesise that this may be due to increased load on, and hence increased failure of, the electrical switchgear during the colder months."

Page 16, Line 23: The additional sites appear to have negligible impact on the inversions. The estimate of total emissions does not change significantly, and the uncertainties are only slightly reduced.

Section removed.

Page 17, discussion of Figure 6: Perhaps additional panels could be included to show the posterior flux distribution for the GAUGE and/or DECC cases. Figure 6a could be revised to show the magnitude of the redistribution relative to the mean. Otherwise it is some work for the reader to understand whether the redistribution is significant.

How can the addition of more data make the inversion estimate more uncertain in some regions? I think this can only be the case if the inversion framework underestimates the uncertainty with the 4-tower case. Some additional explanation is needed. Perhaps it would help to indicate which regions are significantly constrained using a footprint/sensitivity map. Following the comments of both reviewers the decision was made to remove the discussion of the effect of the two towers on UK flux estimates. This will be discussed in a separate paper.

Page 19, Line 27: Apparent typo. . . .lasted between 2-5 (minutes?) Section removed.

Page 26, discussion of figure S3: The average residuals given in Table 1 do not adequately describe the uncertainty indicated in these plots.

Table 1 now includes the interquartile range which better expresses the range of the uncertainties as indicated in plot S3.

Page 26, line 13: The reported large difference between the humidity of samples and standards is a bit concerning. Could this be mitigated by using a longer nation drier? And/or a chiller could be used to remove the bulk of the sample humidity upstream of the nation drier as is done in the NOAA tall tower systems.

Yes, the calibration gases could be further humidified using a longer Nafion[®] or the samples further dried using a chiller system. This suggestion has been incorporated into the conclusions and future work section (Page 28, Lines 15 - 16).

Page 26, line 15: "However it is likely to be a systematic offset of the order of -0.05 to -0.1 ppm. . ." based on what evidence? The residuals in S3? Or the Reum tests? Or something else?

This estimate was based on the residuals shown in Figure S3 (now Figure S5). This has been clarified in the text.

Page 23, Lines 17-20, "However, for BSD and HFD, assuming that the residuals of the droplet water tests are an accurate reflection of the likely error (Figure S3, we expect there to be a systematic offset of the order of -0.05 to -0.1 μ mol mol⁻¹ CO₂ and -1 to -2 nmol mol⁻¹ CH₄."

Page 26, last paragraph: Are these below ambient and above ambient results shown somewhere?

The coefficients for the above and below ambient tests are shown in Table 1 and plots of the residuals are included in the Supplement, Figure S6.

Page 27, line 10: It is surprising that the maximum humidity at these sites is <2%. This is a typo. It should have been 2.5%. This has been corrected (Page 25, Line 11).

Page 27, line 12: I understand that when the Nafion drier was installed then many of the air samples have H2O < 0.1%, but what are the implications? This paragraph is discussing Figure 10 a & b, and this particular figure does not seem to show any troubling implications for samples with H2O < 0.1%.

This paragraph was meant to convey that in light of the variability of the water correction over time, particularly at higher water contents, using a Nafion dryer to obtain a relatively low and stable sample water content was an advantage. This has been clarified.

Page 24, Lines 14 - 16, "In light of the temporal variability of the water correction over time, particularly at higher water contents, using a Nafion[®] dryer or alternative drying method to obtain a relatively low and stable sample water content would be an advantage."

Page 27, line 22: It seems very impractical to calibrate high-humidity sites weekly. Also, I am not convinced that the droplet test is accurate at very high humidity. The daily tests at U of Br in Figure 10e show extremely large variability at humidity > 2.5%.

As the reviewer has noted, performing weekly water calibrations would be a burdensome task and the droplet test can be unreliable at high humidities. This has been included in the text.

Page 24, Lines 24 – 27, "The impracticality of such a frequent testing regime along with the apparent unreliability of the droplet test at $H_2O > 2.5$ % (See Figure S3 g) mean that an

alternative method, possibly partial drying, or a higher level of uncertainty may need be applied to measurements made at higher water contents."

Page 28, line 6: Why was this test not performed at H2O < 0.7% since that is where the droplet test is unreliable due to rapidly changing H2O?

We used the lowest stable set point of the dewpoint generator, 2.5° C, which equated to a water content of 0.6-0.7 %. As such, we were unable to use the dew point generator to examine water contents < 0.7 %.

Page 28, line 21: The range of humidity for HFD and BSD here is different than given on Page 27 line 10 (max of 2.5% instead of 2.0%). Meanwhile the Introduction states that the DECC/GAUGE network observes samples with humidity of up to 3.5%. The values reported on Page 27, line 10 was a typo. This has been corrected. Other sites within the DECC/GAUGE network, specifically the Angus site in Scotland, experienced humidities > 2.5%.

Page 29, line 8: Clarify 0.5 to 3.5% in the Wet experiment versus < 0.31% for the others. Section removed

Page 29 & 30, discussion of Figures 12 & 13 and table 2. This discussion is extremely hard to follow. Since it seems that the droplet tests are highly uncertain below about <0.3%, then consider just omitting that discussion and sticking with the estimates from Reum. Or perhaps omit this discussion altogether. The tepid conclusion at the top of page 31 does not warrant the amount of discussion present in the current manuscript. If the experiments failed or were inconclusive then why include them? If you feel it is important for the community to understand the pitfalls of your attempt in order to avoid similar futile attempts, then perhaps relegate this discussion to the supplement.

As suggested by the reviewer we have removed this section and the accompanying figures and table.

Page 29, line 23: "Although small, these changes are an order of magnitude smaller. ..." This is puzzling. . .did you mean to write: ". . .these changes are an order of magnitude larger"? As suggested by the reviewer we have removed this section and the accompanying figures and table.

Page 31, line 25: The measured difference of 0.02 ppm is practically negligible and unlikely to impact any conceivable scientific analysis, except perhaps analysis of spatial gradients in the high southern hemisphere.

Yes, as the reviewer has stated these differences are small, particularly in the light of the magnitude of the signals observed in the UK regional network. However, small errors can compound and individually may be a cause of concern at "clean air" locations as such we feel it is important to highlight them.

Page 32, line 3: "These cylinders are very dry, H2O < 0.0001%..." But, above on page 26, it states that standards are significantly humidified by the nafion, and so the difference should be even smaller than the 0.005 ppm that you report.

The relationship between water content and CO₂ exchange discussed in this section and plotted on the X axis of Figure 9 is in reference to the water content of the sample (or standard) prior to it entering the Nafion[®]. Yes, the standards are humidified by the Nafion[®] but the CO₂ exchange occurs prior/during this humidification process.

Page 32, line 8: "This is not the case for the water correction bias, which varies with H2O (section 3.3.2)."

This sentence has been removed.

Page 32, line 22: Unfortunately, the inversion results as presented here indicate only marginal improvement in the flux estimates. However this is likely due to limitations of the inversion framework. In particular, it seems that the uncertainty estimated for the 4 tower network is too small.

At the suggestion of the reviewers references to the inversion study have been removed from the paper.

Page 33, line 33: I think this should read 0.015 umol/mol instead of 0.15. Yes, this was incorrect. We thank the reviewer for noting this. It has been corrected Page 28, Line 6.

Table 1: Mean residual is not an adequate diagnostic of the uncertainty, since the residuals vary strongly as a function of H2O. Perhaps include an example plot for one of these calibration instances in the Supplement.

The plots of all of these calibration instances are included in the supplement – see Figures S5 and S6. In an effort to encompass the variability of the residual Table 1 has also been altered to include not just the mean residual but the 25th and 75th percentiles.

Table 2: ISO format for dates is YYYY-MM-DD

The date format in Table 2 (now Table 5) has been corrected.

Figure 2: Include definition of TOC in the legend or the caption. It is mentioned in the text but should also be noted here.

This has been noted in the caption of Figure 2.

Figure 8 & 9: Essentially no discussion of these figures is provided. Either describe these figures or omit.

Figure 8 has been removed. Figure 9 has been moved to the supplement (now Figure S4) and referred to specifically Page 11, Lines 22, 23, 28, Page 12, Line 4 and Page 13, Line7.

Figure 11: Include the CO2 and CH4 values of the cylinders in the legend or the caption.

These values are now included in the caption of Figure 9 (Now Figure 11). "The (a) CO_2 and (b) CH_4 change in the Wet – Dry sample treatment difference with sample water content for cylinders UoB-04 (515.3 µmol mol⁻¹ CO_2 and 2585 nmol mol⁻¹ CH_4), H-296 (406.6 µmol mol⁻¹ CO_2 and 1947 nmol mol⁻¹ CH_4), UoB-06 (384.8 µmol mol⁻¹ CO_2 and 1975 nmol mol⁻¹ CH_4) and H-306 (372.5 µmol mol⁻¹ CO_2 and 1776 nmol mol⁻¹ CH_4). Error bars are the larger of either the standard deviation of the mean difference or the uncertainties of the two sample types added together in quadrature."

Table 2 & Table 3: What is the uncertainty of the assigned values for the calibration standards? The reproducibility of the NOAA scale for CO2 is estimated to be 0.03 ppm. But how well do the calibration centres propagate the scale?

The uncertainties provided by the calibration centres have been added to Table 3. The values reported in Table 4 were made at the AGAGE Mace Head laboratory or the University of Bristol laboratory. An uncertainty assessment for scale propagation in this manner is outside the scope of this paper but will be included in the future uncertainty paper.

References

Bakwin, P. S., Tans, P. P., Hurst, D. F., and Zhao, C.: Measurements of carbon dioxide on very tall towers: results of the NOAA/CMDL program, Tellus B, 50, 401-415, doi:10.1034/j.1600-0889.1998.t01-4-00001.x, 1998.

Flores, E., Viallon, J., Moussay, P., Griffith, D. W. T., and Wielgosz, R. I.: Calibration Strategies for FT-IR and Other Isotope Ratio Infrared Spectrometer Instruments for Accurate δ 13C and δ 18O Measurements of CO2 in Air, Analytical Chemistry, 89, 3648-3655,

10.1021/acs.analchem.6b05063, 2017.

Griffith, D. W. T.: Calibration of isotopologue-specific optical trace gas analysers: A practical guide, Atmos. Meas. Tech. Discuss., 2018, 1-19, 10.5194/amt-2018-187, 2018.

Hall, B. D., Dutton, G. S., Mondeel, D. J., Nance, J. D., Rigby, M., Butler, J. H., Moore, F. L., Hurst, D. F., and Elkins, J. W.: Improving measurements of SF6 for the study of atmospheric transport and emissions, Atmos. Meas. Tech., 4, 2441-2451, 10.5194/amt-4-2441-2011, 2011.

A gas chromatographic system for high-precision quasi-continuous atmospheric measurements of CO2, CH4, N2O, SF6, CO and H2:

https://www.researchgate.net/publication/255704060_A_gas_chromatographic_system_for_high-precision_quasi-

continuous_atmospheric_measurements_of_CO2_CH4_N2O_SF6_CO_and_H2, 2008. Manning, A. J., O'Doherty, S., Jones, A. R., Simmonds, P. G., and Derwent, R. G.: Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modeling approach, Journal of Geophysical Research: Atmospheres, 116, n/a-n/a, 10.1029/2010JD014763, 2011.

Richardson, S. J., Miles, N. L., Davis, K. J., Lauvaux, T., Martins, D. K., Turnbull, J. C., McKain, K., Sweeney, C., and Cambaliza, M. O. L.: Tower measurement network of in-situ CO2, CH4, and CO in support of the Indianapolis FLUX (INFLUX) Experiment, Elementa-Sci. Anthrop., 5, 14, 10.1525/elementa.140, 2017.

Stanley, K. M., Grant, A., O'Doherty, S., Young, D., Manning, A. J., Stavert, A. R., Spain, T. G., Salameh, P. K., Harth, C. M., Simmonds, P. G., Sturges, W. T., Oram, D. E., and Derwent, R. G.: Greenhouse gas measurements from a UK network of tall towers: technical description and first results, Atmos. Meas. Tech., 11, 1437-1458, 10.5194/amt-11-1437-2018, 2018. Tans, P. P., Crotwell, A. M., and Thoning, K. W.: Abundances of isotopologues and calibration of CO2 greenhouse gas measurements, Atmos. Meas. Tech., 10, 2669-2685, 10.5194/amt-10-2669-2017, 2017.

van der Laan, S., Neubert, R. E. M., and Meijer, H. A. J.: A single gas chromatograph for accurate atmospheric mixing ratio measurements of CO2, CH4, N2O, SF6 and CO, Atmos. Meas. Tech., 2, 549-559, 10.5194/amt-2-549-2009, 2009.

Young, D.: Development of a Pre-Concentration System for the Determination of Atmospheric Oxygenated Volatile Organic Compounds, Doctor of Philosphy, School of Chemistry, Faculty of Science, University of Bristol, Bristol, 2007.

Reply to the comments of reviewer #2 of "UK greenhouse gas measurements at two new tall towers for aiding emissions verification"

Ann R. Stavert¹, Simon O'Doherty², Kieran Stanley², Dickon Young², Alistair J. Manning³, Mark F. Lunt⁵, Christopher Rennick⁴ and Tim Arnold^{4,5}.

¹ Climate Science Centre, CSIRO Oceans & Atmosphere, Aspendale, VIC, 3195, Australia

² School of Chemistry, University of Bristol, Bristol, BS8 1TS, United Kingdom

³ Met Office, Exeter, Devon, EX1 3PB, United Kingdom

⁴National Physical Laboratory, Teddington, Middlesex, TW11 0LW, United Kingdom

⁵ School of GeoSciences, University of Edinburgh, Edinburgh, EH9 3FF, UK

General Comments

The paper reports on the installations of GHG measurement systems at two new tall tower sites, describes their setups and their contribution to uncertainty reduction for UK CH4 emission estimates. In addition, the paper describes extensive tests and evaluates the impact of two different sample air drying strategies applied for CRDS instruments in the lab and installed at these two sites. From these subjects, I consider the latter the most relevant and best described in the paper. The text's length, the writing style and the text structure make it often difficult to follow the work and reasoning of the authors. This regarding, substantial improvements are necessary (i.e. shorten and restructure the text). Also, two of the three instruments measure carbon monoxide (CO) – why was this opportunity not used to evaluate the effects of drying / water corrections on the measurements of CO as well? The experiments and findings related to air drying strategies represent an important contribution to this field of research and should thus be accepted for publication after major corrections and only if the comments within this review are adequately and fully addressed.

The authors would like to thank Reviewer 2 for their helpful comments. We have endeavoured to reply to and incorporate all the suggestions that they have made. As requested by the reviewer we have removed and simplified large sections of the text. As such some technical corrections are no longer applicable, these have been noted as such. Reviewer 2 commented on the lack of discussion of the impact of Nafion® drying and the CRDS water correction on CO mole fractions. As the Nafion® drying experiments were conducted using a Picarro CRDS G2301, which does not measure CO, the effect of Nafion® drying on CO mole fractions was not investigated. While instrument specific CO water corrections were calculated, the large minute-mean variability inherent in the G2401 CO measurements (> 4 nmol mol⁻¹) meant that the difference between the instrument specific and in-built correction was not statistically significant (Page 22, Lines 10-14). As such, these corrections were not presented in the body of the paper, however, further information can be found in Figure S5 of the supplementary. Even with the inclusion of the extra information requested by the reviewers these changes have reduced the length main body of the paper by 6 pages and removed 4 figures.

Specific comments and technical corrections

Page 1, Line 15 (1/15): "...were located...": I would rather go for "are located"

Removed.

1/20: "by between" -> "by up to 20 %" Section removed

1/29: "..default factory.." it is well established that the default factory water correction is not to be used for the purposes described – this fact should be stated more clearly here. Reference to default factory corrections has been removed from the abstract. Details of the issues associated with this correction are given in the introduction. Page 3, Lines 15-20, "Initially, it was hoped that the inbuilt water correction would remove the need for sample drying, inherent in most other methods (e.g. FTIR or NDIR) but subsequent studies questioned its stability over time and between instruments (Yver Kwok et al., 2015;Chen et al., 2010;Winderlich et al., 2010). In response to this, researchers have typically developed their own water corrections or have returned to sample drying in order to minimise the effect (Welp et al., 2013;Winderlich et al., 2010;Schibig et al., 2015)."

2/17: write out the names of the gas molecules at first use. Completed as requested Page 25, Line 15.

3/3: need "While,"? Section removed.

3/10: delete "recognised" Deleted.

3/16: A dramatic reduction in the cost of these instruments could be debated. I suggest you keep only the next sentence, starting in line 17. Sentence deleted.

4/4: Please explain why dry zero air was chosen for the counter flow purge gas (instead of using the reflux method with sample air from the instrument's outlet). This is relevant because it is on one side true that the drying with dry (zero) air is more effective, but on the other side trace gas species gradients can influence the transition dynamics of these trace species across the Nafion membrane and thus the composition of the dried sample air as well.

Page 3, Line 22 – 25, "For ease of servicing, the CRDS instrumentation at GAUGE and UK DECC Network sites was initially deployed using an identical drying method to that of the co-located GC instrumentation. This method relied on drying the sample with a Nafion[®] water permeable membrane in combination with dry zero air as a counter purge gas."

If I am not much mistaken, Nafion must be written as Nafion[®]. We have replaced all references to "Nafion" with "Nafion[®]".

5/18: correct to: "m a.g.l." Corrected throughout text.

6/2: it would be advantageous, if you stated at least once the serial numbers of the Picarro instruments used in your setups.

This information has been added to the text.

Page 5, Line 20, "Both sites are equipped with a CRDS (G2401 Picarro Inc., USA, CFKADS2094 and CFKADS2075 deployed at Bilsdale and Heathfield, respectively) making high frequency (0.4 Hz) CO_2 , CH_4 , CO and H_2O measurements."

7/7: Please summarize here briefly the reason for discontinuing the Nafion drying. State the main arguments/problems that are thus avoided or minimized. If the reader jumps to Section 3.3.3 to find out about it s/he will have a hard time finding the explanation. Sentence in 7/12 is similarly ambiguous.

These have been clarified.

Page 6, Line 25 to Page 7, Line 6. "In an attempt to minimise the water correction required for dry mole fraction CRDS measurements, CRDS samples were initially dried using a Nafion[®] in an identical manner to those of the GC-ECD. This resulted in air samples with water mole fractions between 0.05 and 0.2 % H₂O depending on the original moisture content of the air. However, due to concerns that the mole fraction gradient between the sample and the Nafion[®] counter purge might lead to CO₂ transport across the Nafion[®] membrane this drying approach was discontinued. The CRDS Nafion[®] drying systems were removed on the 17th of June 2015 & 30th of September 2015 at BSD and HFD, respectively and undried air analysed and the data post corrected with an instrument specific water correction."

And

Page 7, Lines 18 - 19, "Motivated by the possibility of CO₂ transport across the Nafion[®] membrane, the decision was made to measure wet samples and correct using an instrument specific water correction."

7/26: Note: Picarro analyzers are not calibrated for H2O measurements and often show different positive or negative values close to zero at third place behind the comma. For the sake of completeness of information, it would be informative to know what the zero was for water on all three Picarro instruments.

The below information has been added to the text.

Page 8, Lines 23-27, "As Picarro analysers are not calibrated for H_2O measurements when measuring dry air they often show different positive or negative values close to zero. For the instruments studied in this paper these "zero-water" values were 0.00001, -0.0003 and -0.002 for the Bilsdale, Heathfield and University of Bristol laboratory instruments respectively. These values were determined using measurements of cylinders of dry air where the first 120 minutes were ignored and the "zero-water" value calculated as the mean H_2O of the subsequent data (> 60 min)."

8/2: please add a short justification/explanation on how the criteria were decided on (particularly the thresholds for 1 min mean standard deviations).

The below information has been added to the text.

Page 7, Lines 4-9, "Data collected in the first five minutes immediately following the injection, the typical line equilibration period, were excluded from the fit. This avoids using data adversely effected by the effect of rapid changes in H_2O content on the cell pressure sensor, as identified by Reum et al. (2018) and the erroneous post-injection CO_2

enhancement identified by (Rella et al., 2013). Again, due to cell pressure sensor concerns, data points with minute-mean H_2O standard deviations > 0.5 % H_2O were excluded. This 5-minute cut-off reduced the maximum H_2O value included in the fit to 4 % H_2O ."

8/11: I am not convinced by the nomenclature for gases; what do you mean with "close- toambient standard"? From the continuation of the text, I guess this is a "working tank" that helps you to account/correct for short-term drift? You should then use "standard gas" for what you now call "calibration gas".

As the reviewer has noted the standard cylinder is used to account for short-term drift and as such could also be called a "working tank" (Page 14, Line 7). The text has been updated to include a reference to this term. However, we would prefer to leave the labelling of the calibration cylinders unchanged. These cylinders are calibrated to WMO calibration scales and used to propagate these scales to the atmospheric observations (i.e. calibrate) as such we feel "calibration cylinders" is the most appropriate terminology. The use of the nomenclature "standard cylinder" and "calibration cylinders" in this way is also consistent with other related publications (e.g. Stanley et al. (2018)) and as such we feel it is important to maintain it.

8/17: "greatly increases the error" – what does this explicitly mean? How much? Better reformulate, drop "greatly" and, if possible, add facts-based estimates. Added to the text.

Page 14, Lines 13 - 23, "Assigning mole fractions to values outside the range of the calibration suite will increase the error. The magnitude of this error will depend on the magnitude of the mole fraction difference between the lowest calibration cylinder and the sample. This error has been estimated using measurements made at the Heathfield site of cylinders of known CO mole fractions, 6 and 57 nmol mol⁻¹ CO below the lowest calibration cylinder. These show a percentage error of 2.41 and 3.09 %, respectively. A similar assessment of the error associated with samples above the highest calibration standard were made using cylinders 87 and 686 nmol mol⁻¹ CO above the highest calibration standard. These correspond to percentage errors of 2.98 and 2.56%, respectively. As all the minute-mean CO measurements below of the calibration range are within 57 nmol mol⁻¹ of the lowest calibration cylinder and the vast majority of minute-mean CO measurements above the calibration cylinder (99%) we expect that this error would typically be < 3%."

8/21: no "target gas" seems to be used – for the appropriate QC, this is required. There is a comment on this in Chapter 4. – but it would be helpful if the authors comment on how the absence of a target gas is influencing the current performance of the system (i.e. what is the impact on the uncertainties, etc.).

Target tanks, as defined by the World Meteorological Organisation (WMO, 2016), are an extremely useful quality control tool and "function as a warning that there might be a problem" but they are "not to be used to define a second, optional, path of traceability". As such, although they can be used to assess the uncertainty of the measurements, their use (or not) does not contribute to the uncertainty of any measurements. With this in mind, the authors believe that, although the use of a target tank would be a useful addition to the system, the lack thereof does not negatively impact the uncertainties of the measurements it merely impacts our ability to characterise these uncertainties.

This has been clarified in the text.

Page 29, Lines 3 - 8, "Future improvements to the Bilsdale and Heathfield records include the addition of target tanks at the sites. Although the use of target tanks do not directly influence measurement uncertainty they allow independent long-term monitoring of instrument performance and are a useful tool for assessing measurement uncertainty. The development of a full uncertainty analysis incorporating such target tank measurements, along with an assessment of the calibration strategy and any instrumental, water correction and sampling errors is also planned."

9/8: In Table 3, the pre- and post-site calibration values should be given (where applicable). Where available the pre- and post-site calibration values have been added to Table 3.

10/2: The statement "... with the mean absolute precision increasing (i.e. becoming less precise)" seems wrong to me.

We can understand why this sounds odd to the reviewer and have clarified the text. Page 16, Lines 23 – 24, "Both sites showed a small trend with the mean absolute precision worsening with increasing CO₂ and CH₄ mole fraction."

10/22: the numbers for pressure and temperature are switched – correct. Page 17, Line 17, Corrected.

12/15: 2.3 Results and discussion; 3.3 has the same name, which is confusing. A paper should have one Results/Discussion section.

The paper has been restructured to remove the second Results/Discussion section.

13/19: What does "...green space..." refer to?

Green space refers to protected areas of undeveloped landscape this can include areas of native vegetation along with parks and reserves, however, for clarity the term has been replaced with "native vegetation" (Page 19, Lines 9-10).

14/8-19: I can only second the comments by Anonymous Referee #1 on this subject. As detailed in our reply to reviewer #1 the FFT based curve fitting method has been removed from the paper and diurnal and seasonal trends are now examined using box plots.

14/23: I am reserved about calling a 3.5 years long data set as sufficient for calculating longterm trend curves.

Section removed

15/2: put spaces between number and unit (e.g., 108 m). Corrected throughout the text.

17/25: you can safely drop "extremely low" without losing any information. What were the values for CO?

Removed "extremely low". Carbon monoxide was <12 nmol mol⁻¹, this information has been added to the text (Page 7, Line 11).

18:21: this would be an excellent place to state the maximum humidity of air in the cylinders.

This section has been removed and instead the typical humidity of the dried cylinders has been included in an earlier part of the paper.

Page 7, Line 21, "...air from a cylinder of dry (<0.002 % H₂O) natural air..."

19/3: typo in the flow range

Corrected Page 11, Line 6, "...had a flow range of 0.1–0.5 L min⁻¹..."

19/14: analogously to I.D.; 1 " O.D. Corrected Page 10, Line 25, "...coil of ¼ " diameter (I.D. 0.12 ")..."

19/17: please specify "polymeric plastic tubing" better and explain why it was necessary to include it (particularly as this introduces further possible "active" surfaces that can influence the performance of the system).

We understand that this does introduce possible active surfaces, however, due to the fittings of the DPG this was unavoidable. The text has been updated to address this and to specify the exact tubing type.

Page 11, Lines 9-12, "Other than the cryogenic water trap and two short sections (< 10 cm) of ¼" (O.D.) "Bev-a-line" plastic tubing immediately prior to and post the DPG, 1/16" stainless steel tubing was used throughout the system. Due to the air output and input connections of the DPG the use of the plastic tubing was unavoidable."

19/27: typo "2-5" Section removed.

20/7: please be more explicit than "were monitored closely and regularly logged." The text has been updated to Page 12, lines 9-11, "Flow rates, cylinder pressure, chamber temperature and H2O trap temperature were manually logged after each valve position

change and when the water trap was inserted into the silicone oil bath."

20/27: a flowmeter rather "monitors' than "ensures" Section removed.

21/2: it would be helpful for the reader, if you added how this value translates to % H2O This equates to approximately 2.9% H₂O. This value has been added to the text Page 11, Line 19.

21/7: please specify "stabilise".

Page 11, Lines 18- 19, "A stable water content was defined as one where the standard deviation of the minute mean values was < $0.003 \ \% H_2O$ for a 15-min period."

21/16: I am not sure what you mean here – if we are talking about H2O < 0.0001 %, we go far below a dew point of 0 deg.C – please explain. Sentence removed.

21/19: please clarify the frequency 1.5 Hz/data recording strategy here (as the Picarros record with 0.4 Hz).

Picarro's operate by taking an individual measurement of each gas of interest, in the case of a G2401 instrument CO2, CH4, H2O and CO, in series. For our instruments, these measurements are typically 0.7 sec (approx. 1.5 Hz) apart. Instead one, perhaps more correctly, could consider the time it takes to complete a full cycle through all the gases of interest. For our instruments, this typically takes 2.5 sec (approx. 0.4 Hz). We have updated the text to reflect this.

22/7: you should number the equations

Equations numbered

24/21: use "target gas"

The authors would prefer not to use the phrase "target gas" as this is a specific term reserved for a long term surveillance tank which is completely independent of the calibration and drift correction of the instrument (WMO, 2016). The tank referred to here is used to drift correct the measurements and cannot be classed as a target tank. Instead, as these cylinders are measured periodically over multiple years the raw uncorrected measurements can be used to estimate typical instrumental drift. This has been clarified in the text.

Page 13, Line 26 - "For the UoB CRDS, raw long-term measurements of standard cylinders have shown the typical instrumental drift to be < $0.001 \ \mu mol \ mol^{-1} \ day^{-1} \ CO_2$ and < $0.03 \ nmol \ mol^{-1} \ day^{-1} \ CH_4$."

25/14: wrong use of c.f. (cf) – use vs. Corrected Page 22, Line 18.

27/6: "WMO internal reproducibility"; also, decide what you want to use: internal reproducibility guidelines/requirements/bounds and use throughout the text consistently. Please also explain somewhere (best at first use) why you (correctly) aim for Internal reproducibility and not for the Recommended compatibility.

Corrected throughout to the text to "WMO internal reproducibility guidelines". Text added to clarify the choice of internal reproducibility guidelines.

Page 4, Lines 17 – 22, "The importance of these errors are assessed in comparison to the WMO internal reproducibility guidelines (WMO, 2016) which incorporate not only the instrumental precision but uncertainties related to other components sample collection and measurement including drying. These internal reproducibility guidelines are typically half the WMO recommended compatibility goals which, unlike the reproducibility guidelines, are driven by the need for compatibility between datasets."

30/8: can you please explain in a bit more detail how you did this – and how you were able to estimate Nafion-related errors, since, to my knowledge, Reum et al. (2017) did not use one?

Section removed

31/23: would you be able to discuss the role/importance of the trace gas species gradients between sample and counter flow gas? (see 4/4)

Added to the text.

Page 26, Lines 6 – 11, "Nafion[®] membranes, when combined with a dry counter purge gas stream, can be used to effectively dry air samples. This drying process is driven by the moisture gradient between the "wet" sample and the dry counter purge. In a similar manner, as long as the membrane is permeable to the gas, a sample to counter purge gradient in any other trace gas species will also drive exchange. In an effort to quantify the magnitude of CO_2 and CH_4 exchange a series of experiments measuring the composition of the Nafion[®] counter purge gas were conducted."

32/3: see 21/16

Clarified

Page 27, Lines 11 - 12 "These cylinders are very dry, $H_2O < 0.0001$ %, equivalent to the driest conditions studied in the DPG experiments"

37: North and scale are missing.

North and scale have been added. Figure 1.

38: all abbreviations used should be explained in the figure's caption

Caption now reads "Figure 2: A generalised schematic showing the initial Bilsdale and Heathfield site setup of the cavity ringdown spectrometer (CRDS) and the Gas Chromatograph – Electron Capture Dectector (GC-ECD) including the gas generator (TOC) and back pressure regulator (BP). Note that Bilsdale has three inlets, while Heathfield has only two as shown here. The Nafion[®] drying system located downstream of the CRDS multiport valve was removed at both sites in 2015. Black arrows and lines show the direction of sample, standard and calibration gas flow. Grey dashed lines and arrows show the flow path of the Nafion[®] counter purge gas."

39: I presume "Hour of day" means local time? Better replace with "Hour (LT)". Was summer / winter time used as well?

Data were not adjusted for summer/winter time daylight savings. All data were recorded and plotted on UTC as such, "Hour of day" has been replaced with "Hour (UTC)". Figure 6.

41: see 14/8-19 comment regarding spikes

This figure has been replaced with two using box plots (Figures 6 and 7) rather than a FFT based curve fitting method to examine diurnal and seasonal trends. This has removed the spikes.

Figure S3: equations not legible

Equation text font size has been increased on Figures S5 and S6.

References

Chen, H., Winderlich, J., Gerbig, C., Hoefer, 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 (CO₂ and CH₄) using the cavity ring-down spectroscopy (CRDS) technique, Atmos. Meas. Tech., 3, 375-386, 10.5194/amt-3-375-2010, 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, Atmospheric Measurement Techniques, 6, 837-860, 10.5194/amt-6-837-2013, 2013. Reum, F., Gerbig, C., Lavric, J. V., Rella, C. W., and Göckede, M.: Correcting atmospheric CO₂ and CH₄ mole fractions obtained with Picarro analyzers for sensitivity of cavity pressure to water vapor, Atmos. Meas. Tech. Discuss., 2018, 1-29, 10.5194/amt-2018-242, 2018. Schibig, M. F., Steinbacher, M., Buchmann, B., van der Laan-Luijkx, I. T., van der Laan, S., Ranjan, S., and Leuenberger, M. C.: Comparison of continuous in situ CO2 observations at Jungfraujoch using two different measurement techniques, Atmos. Meas. Tech., 8, 57-68, 10.5194/amt-8-57-2015, 2015.

Stanley, K. M., Grant, A., O'Doherty, S., Young, D., Manning, A. J., Stavert, A. R., Spain, T. G., Salameh, P. K., Harth, C. M., Simmonds, P. G., Sturges, W. T., Oram, D. E., and Derwent, R. G.: Greenhouse gas measurements from a UK network of tall towers: technical description and first results, Atmos. Meas. Tech., 11, 1437-1458, 10.5194/amt-11-1437-2018, 2018. Welp, L. R., Keeling, R. F., Weiss, R. F., Paplawsky, W., and Heckman, S.: Design and performance of a Nafion dryer for continuous operation at CO₂ and CH₄ air monitoring sites, Atmos. Meas. Tech., 6, 1217-1226, 10.5194/amt-6-1217-2013, 2013.

Winderlich, J., Chen, H., Gerbig, C., Seifert, T., Kolle, O., Lavrič, J. V., Kaiser, C., Höfer, A., and Heimann, M.: Continuous low-maintenance CO₂/CH₄/H₂O measurements at the Zotino Tall Tower Observatory (ZOTTO) in Central Siberia, Atmos. Meas. Tech., 3, 1113-1128, 10.5194/amt-3-1113-2010, 2010.

WMO: 18th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques (GGMT-2015), Geneva, Switzerland, 2016.

Yver Kwok, C., Laurent, O., Guemri, A., Philippon, C., Wastine, B., Rella, C. W., Vuillemin, C., Truong, F., Delmotte, M., Kazan, V., Darding, M., Lebègue, B., Kaiser, C., Xueref-Rémy, I., and Ramonet, M.: Comprehensive laboratory and field testing of cavity ring-down spectroscopy analyzers measuring H_2O , CO_2 , CH_4 and CO, Atmos. Meas. Tech., 8, 3867-3892, 10.5194/amt-8-3867-2015, 2015.

UK greenhouse gas measurements at two new tall towers for aiding emissions verification

Ann R. Stavert¹, Simon O'Doherty², Kieran Stanley², Dickon Young², Alistair J. Manning³, Mark F. Lunt⁵, Christopher Rennick⁴ and Tim Arnold^{4,5}.

¹ Climate Science Centre, CSIRO Oceans & Atmosphere, Aspendale, VIC, 3195, Australia ² School of Chemistry, University of Bristol, Bristol, BS8 1TS, United Kingdom ³ Met Office, Exeter, Devon, EX1 3PB, United Kingdom

⁴ National Physical Laboratory, Teddington, Middlesex, TW11 0LW, United Kingdom

⁵ School of GeoSciences, University of Edinburgh, Edinburgh, EH9 3FF, UK

10

Correspondence to: Ann R. Stavert (ann.stavert@csiro.au)

Abstract. Under the UK focused Greenhouse gAs and Uk and Global Emissions (GAUGE) project, two new tall tower greenhouse gas (GHG) observation sites were established in the 2013/2014 Northern

- 15 Hemispheric winter. These sites located at existing telecommunications towers utilised a combination of cavity ring-down spectroscopy (CRDS) and gas chromatography (GC) to measure key GHGs (CO₂, CH₄, CO, N₂O and SF₆). Measurements were made at multiple intake heights on each tower. <u>CO₂ and</u> CH₄ dry mole fractions were calculated from either CRDS measurements of wet air which were post corrected with an instrument specific empirical correction or samples dried to between 0.05 and 0.3 %
- 20 H₂O using a Nafion[®] dryer, with a smaller correction applied for the residual H₂O. The impact of these two drying strategies was examined. Drying with a Nafion[®] drier was not found to have a significant effect on the observed CH₄ mole fraction; however, Nafion[®] drying did cause a 0.02 µmol mol⁻¹ CO₂ bias. This bias was stable <u>for sample CO₂ mole fractions between 373 and 514 µmol mol⁻¹ and for sample H₂O up to 3.5 %. As the calibration and standard gases are treated in the same manner, <u>the 0.02 µmol</u></u>
- 25 <u>mol⁻¹ CO₂ bias</u> is mostly calibrated out with the residual error ($\ll 0.01 \mu mol mol⁻¹ CO₂$) well below the World Meteorological Organization's (WMO) reproducibility requirements. Of more concern was the error associated with <u>the</u> empirical instrument specific water correction algorithms. These corrections are relatively stable and reproducible for samples with H₂O between 0.2 and 2.5 %, CO₂ between 345 and

6	Put u u
U	Deleted: were
•(Deleted: two
Ó	Deleted: ,
í	Deleted: Heathfield (HFD) and Bilsdale (BSD),
-6	Deleted. The inclusion of the two additional tower stations within

Deleted: The mclusion of the two additional tower stations within the existing UK Deriving Emissions linked to Climate Change (DECC) network of four stations was found to reduce the uncertainty of CH_4 UK emission estimates by between 10-20 %.

Deleted: with

Deleted: this error

Deleted: both default factory and

449 μ mol mol⁻¹ and CH₄ between 1743 and 2145 nmol mol⁻¹. However, the residual errors in these corrections increase to > 0.05 μ mol mol⁻¹ for CO₂ and > 1 nmol mol⁻¹ for CH₄ (greater than the WMO internal reproducibility guidelines) at higher humidities and for samples with very low (< 0.5 %) water content. These errors also scale with the absolute magnitude of the CO₂ and CH₄ mole fraction. As such,

water corrections calculated in this manner are not suitable for samples with low (< 0.5 %) or high (> 2.5 %) water contents and either alternative correction methods should be used or partial drying or humidification considered prior to sample analysis.

Deleted: were

eletea: found

1 Introduction

5

- 10 The adverse effects of anthropogenically driven increases of greenhouse gas concentrations on global temperatures and climate have been well established (IPCC, 2013). Governmental efforts to curb these emissions include the UK 2008 Climate Change Act, which requires the UK to decrease its GHG emissions by 80 % of 1990 levels by 2050 (Parliment of the United Kingdom, 2008 Chapter 27). This in turn motivated the Greenhouse gAs Uk and Global Emissons (GAUGE) project, which aimed to better
- 15 quantify the UK <u>carbon dioxide (CO₂), methane (CH₄) and nitrous oxide (N₂O) emissions. These new emission estimates would then be used to assess the impact of emission abatement and reduction strategies. Key to the GAUGE project was combining new and existing GHG data streams, including high-density regional observation studies, tall tower sites, moving platforms (ferry and aircraft) and satellite observations, with innovative modelling approaches.</u>
- 20 This paper describes the establishment of two new UK GHG tall tower (TT) sites funded under the GAUGE project. Here we provide an analysis of the observations made at the sites and investigate the error associated with empirical instrument specific water correction algorithms and the Nafion®-based sample drying approach used at these TT sites. A further paper, currently in preparation, will discuss the integration of these new sites with the existing UK Deriving Emissions linked to Climate Change (DECC)
- 25 network (Stanley et al., 2018) funded by the UK Department of Business, Energy and Industrial Strategy (BEIS) and provide a full uncertainty analysis for data collected at all the DECC/GAUGE sites. <u>A second</u> companion paper<u>also</u> in preparation, will discuss the integration and inter-calibration of all the CO₂,

Deleted: 1

Deleted:

D	eleted: is divided into two sections. Firstly, it
D	eleted: ,
D	eleted: es
- (D	eleted:
D oł tr	eleted: sites and outlines additional insights that these new bservations provide when quantifying UK GHG emissions and ends. Secondly, this paper
D	eleted: s
	eleted: (Stanley, Stavert et al.
D	eleted:)
- (D	eleted: While, a
D	eleted: (Stavert, Stanley et al.,
. (n	eleted:)

 CH_4 , CO, N_2O and SF_6 data streams including near surface, tall tower, ferry and aircraft measurements along with an analysis of the impact of identified site biases on UK GHG emission estimates. Like the UK DECC network, the new sites, Bilsdale (BSD) and Heathfield (HFD), are equipped with a combination of cavity ring-down spectrometer (CRDS) and gas chromatograph (GC) instrumentation

5 (Stanley et al., 2018). These instruments, along with the associated calibration gases (linked to, <u>WMO</u> calibration scales) and automated sampling systems are located at the base of telecommunication towers within the UK. Further details of the sites and instruments used along with a description of the data collected to date are provided in the subsequent sections.

The precision, stability, relative autonomy and robustness of CRDS instrumentation has led to a rapid

- 10 increase in their deployment in global, continental and regional GHG monitoring networks including the GAUGE network, the European Integrated Carbon Observing System (ICOS) (Yver Kwok et al., 2015) and the Indianapolis Flux Experiment (INFLUX) (Turnbull et al., 2015). These instruments also claim the advantage of being able to measure un-dried ("wet") air samples which are then post corrected to "dry" values using an inbuilt algorithm (Rella, 2010).
- 15 Initially, it was hoped that the inbuilt water correction would remove the need for sample drying, inherent in most other methods (e.g. FTIR or NDIR) but subsequent studies questioned its stability over time and between instruments (Yver Kwok et al., 2015;Chen et al., 2010;Winderlich et al., 2010). In response to this, researchers have typically developed their own water corrections or have returned to sample drying in order to minimise the effect (Welp et al., 2013;Winderlich et al., 2010;Schibig et al., 2015;Rella et al.,
- 20 2013). As such the examination of any errors or biases induced by drying and water correction methods is essential for fully quantifying the uncertainty of CRDS measurements.

For ease of servicing, the CRDS instrumentation at GAUGE and UK DECC Network sites was initially deployed using an identical drying method to that of the co-located GC instrumentation. This method relied on drying the sample with a Nafion® water permeable membrane in combination with dry zero air

25 as a counter purge gas. Here, due to the moisture gradient between the sample and the counter purge, the water passed from the wet sample through the membrane to the dry counter purge. Drying in this manner has a history of successful application for the measurements of halocarbons (Foulger and Simmonds, 1979), N₂O (Prinn et al., 1990) and SF₆ (Fraser et al., 2004). However, studies have found that CO₂ and

3

Deleted: recognised

Deleted: World Meteorological Organization

Deleted: in the first part of this paper. As an example of the impact of these two new sites in terms of UK GHG emissions estimates, inversion based estimates of UK CH4 emissions and their associated uncertainties, with and without the new TT sites, are also included.

Deleted: Recent advances in tuneable diode lasers used within CRDS instrumentation has led to a dramatic reduction in the cost of these instruments. This, along with their Deleted:

Deleted: H₂O Deleted: have

Deleted: The Deleted: deployed CH_4 can also pass across a dry Nafion membrane (Chiou and Paul, 1988) and that this transport increases with the water saturation of the membrane (Naudy et al., 2014). As the transport process is driven by a partial pressure difference between the sample and counter purge gas it is possible that changes in the sample CO_2 and CH_4 mole fraction relative to the counter purge gas, along with the <u>water (H₂O)</u> content of the sample, may alter the magnitude of any cross-membrane leakage.

- A study by Welp et al. (2013) examined this issue and concluded that the leakage was small and well within the WMO <u>compatibility</u> guidelines. However, the drying approach used by Welp et al. (2013) is not directly comparable to that of the GAUGE sites as they used dry sample gas as the counter purge rather than zero air. That study also only examined two water contents (0 % or 2 % H₂O) and conducted
- 10 only dry (0 % H₂O) experiments on samples with CO₂ and CH₄ mole fractions above ambient concentrations. Considering the importance of water in gas transport across the membrane (Chiou and Paul, 1988) and the range of water contents observed in undried air samples measured within the DECC/GAUGE network (up to 3.5 % H₂O) further investigation of this issue was required.

As such, this paper also aims to quantify the magnitude of Nafion[®] CO₂ and CH₄ transport using the 15 drying method used at the DECC/GAUGE TT sites along with errors associated with instrument specific water corrections. It also examines how these might change within the range of H₂O, CO₂ and CH₄ mole fractions typically observed at these sites. The importance of these errors are assessed in comparison to the WMO internal reproducibility guidelines (WMO, 2016) which incorporate not only the instrumental precision but uncertainties related to other components sample collection and measurement including

20 drying. These internal reproducibility guidelines are typically half the WMO recommended compatibility goals which, unlike the reproducibility guidelines, are driven by the need for compatibility between datasets.

2 Experimental

5

2.1 Site descriptions

- 25 Two new tall tower sites, Heathfield (HFD; 50.977 °N, 0.231 °E) and Bilsdale (BSD; 54.359 °N, -1.150 °E) were established at existing telecommunication towers in December 2013 and January 2014,
 - 4

Deleted: comparability

Deleted: the second half of

Deleted:

respectively. The general set up of these sites is similar to that described for the DECC sites in Stanley et al. (2018) and the locations of these two new sites relative to these sites described in Stanley et al. (2018) are shown in Figure 1.

Heathfield is located in rural East Sussex, 20 km from the coast. The closest large conurbation (Royal

5 Tunbridge Wells) is located 17 km NNE from the tower. The area surrounding the tower is > 90 % woodland and agricultural <u>areas</u> with some residential (0.7 %) and light industrial areas (0.3 %) (East Sussex in figures, 2006). Notable local industry includes a large horticultural nursery located only 200 m north of the tower.

Bilsdale is a remote moorland plateau site within the North York Moors National Park. It is 25 km NNW of Middlesbrough (the closest large urban area) and 30 km from the coast. The tower is situated in a predominantly rural area, including moorland, woodland, forest and farmland (North York Moors National Park Authority, 2012;Chris Blandford Associates, 2011).

Inverted stainless steel intake cups were mounted at 42, 108 and 248 m.a.g.l. (metres above ground level) on the BSD tower and 50 and 100 m.a.g.l. at HFD. Air was pulled through the intake cups via ½ " Synflex

15 Dekabon metal/plastic composite tubing (EATON, USA) and a 40 μm filter (SS-8TF-40, Swagelok, UK) using a line pump (DBM20-801 linear pump, GAST Manufacturing, USA) operating at > 15 L min⁻¹. The instruments located at the sites sub-sampled from the tower intakes via a T-piece prior to the line pump. Further details can be found in Stanley et al. (2018).

2.2 Instrumentation

- 20 Both sites are equipped with a CRDS (G2401 Picarro Inc., USA, CFKADS2094 and CFKADS2075 deployed at Bilsdale and Heathfield, respectively) making high frequency (0.4 Hz) CO₂, CH₄, CO and H₂O measurements. A GC coupled to a micro-electron capture detector (GC-ECD, Agilent GC-7890) is used to measure N₂O and SF₆ every 10 mins. For further instrumental details, including flow diagrams and column details, see Stanley et al. (2018).
- 25 The sample lines, calibration and standard gas cylinders are linked to two multiport valves (EUTA-CSD10MWEPH, VICI Valco AG International, Switzerland), one for the CRDS and a second for the GC-

5

Deleted: green space

Deleted:

ECD, the output of each valve is connected to the intakes of the instruments. Filters (7µm, SS-4F-7, Swagelok, UK) are located on the intake lines prior to the valve while a 2µm filter (SS-4F-2, Swagelok, UK) is located between the valve and the CRDS. The GC-ECD flow path, instrumentation and part numbers are described in detail in Stanley et al. (2018). However, in brief, air entering the GC-ECD

5 system is first dried (Section 2.3.1) before flushing an 8 mL sample loop. The contents of the loop are transferred onto a combination of pre-, main and post chromatographic columns using P-5 carrier gas (a mixture of 5 % CH₄ in 95 % Ar; Air Products, UK).

The automated switching of valves and control of GC-ECD temperatures and flows, as well as logging the data and a range of other key parameters (flows, pressures, temperatures) is achieved using custom

10 Linux based software (GCWerks, www.gcwerks.com). The CRDS instrument makes measurements at each intake height, switching between heights every 20 mins at BSD and 30 mins at HFD. While the GC-ECD measures only a single intake, initially the 108 m.a.g.l. intake at BSD (switched to the 248 m.a.g.l. intake on 17th March 2017) and the 100 m.a.g.l. intake at HFD. Other than the tower sample lines, all tubing within the system is 1/16 ", 1/8 " or ¹/₄ " (O.D.) stainless steel (Supelco, Sigma-Aldrich, UK). A

15 generalised diagram of the original sampling scheme for the two sites is shown in Figure 2.

2.3 Sample Drying and CRDS water correction

2.3.1 GC-ECD

All samples measured on the GC-ECD (air, standards and calibration) are dried using a Nafion[®] permeation drier (MD-050-72S-1, Permapure, USA) prior to analysis. The counter purge gas for the drier

20 is generated from compressed room air. <u>The counter purge</u> is dried to <<u>0.005 % H₂O by the compressor</u> (50 PLUS M, EKOM, Slovak Republic) and a <u>gas_generator designed for total organic carbon instruments</u> (TOC-1250, Parker Balston, USA). <u>Previous examinations of this drying method have found that samples are dried to < 0.0002 % H₂O (Young, 2007).</u>

2.3.2 CRDS

25 In an attempt to minimise the water correction required for dry mole fraction CRDS measurements, CRDS samples were initially dried using a Nafion[®] in an identical manner to those of the GC-ECD. This resulted

6

Deleted:	
Deleted:	
Deleted:	

Deleted: This air	
Deleted: zero air	
Deleted:	

Formatted: Subscript

in air samples with water mole fractions between 0.05 and 0.2 % H₂O depending on the original moisture content of the air. However, <u>due to concerns that the mole fraction gradient between the sample and the Nafion® counter purge might lead to CO₂ transport across the Nafion® membrane this <u>drying approach</u> was <u>discontinued</u>. The CRDS Nafion® drying systems were removed <u>on the 17th of June 2015 & 30th of</u></u>

5 September 2015 at BSD and HFD, respectively<u>and</u> undried air analysed and the data post corrected with an instrument specific water correction.

2.3.3 Composition of the counter purge dry air stream

As the drying technique implemented in this study uses a Nafion® drier which relies on a dry counter purge air stream measurements of the HFD, BSD and University of Bristol (UoB) laboratory counter 10 purge were made using the HFD, BSD and UoB CRDS instruments, respectively. All counter purge streams showed mole fractions of CO₂ < 0.3 µmol mol⁻¹, CH₄ < 2 nmol mol⁻¹, CO < 12 nmol mol⁻¹ and H₂O < 0.01 % (Figure <u>\$1</u>). All these zero air streams have CO₂ and CH₄ mole fractions far lower than the 2015 mean global concentrations, 400.99 µmol mol⁻¹ CO₂ and 1840 nmol mol⁻¹ CH₄ (Dlugokencky and Tans, 2015;Dlugokencky, 2015). While the CO mole fraction is significantly lower than the minimum

15 <u>CO mole fractions typically observed at the HFD and BSD sites</u>, $\sim 60 \text{ nmol mol}^{-1}$ <u>As such there is a clear</u> and sizable partial pressure difference across the Nafion® membrane for <u>all three</u> species.

2.3.4 Calculating instrument specific water corrections,

Motivated by the possibility of CO₂ transport across the Nafion[®] membrane, the decision was made to measure wet samples and correct using an instrument specific water correction. These corrections were

- 20 determined in the field by conducting a droplet test, similar to those described in Rella et al. (2013). In this test, air from a cylinder of dry (< $0.002 \ \% H_2O$) natural air was humidified and the change in CO₂ and CH₄ mole fraction with water content examined. In brief, a 1.5 m length of 3/8 " Synflex Dekabon metal/plastic composite tubing (EATON, USA) was introduced between the standard cylinder outlet and the CRDS intake. Distilled water (0.7 mL) was injected through a septum located on a T-piece fixed on
- 25 the "cylinder end" of the Dekabon tubing (See Figure <u>\$2</u> for flow diagram). This water evaporated into the sample stream, with the H₂O mole fraction typically peaking at up to 4.5 % (dependent on room

Del	eted		meti	hoo
-----	------	--	------	-----

Deleted: found to be problematic. As such, the

Deleted: S2

Formatted: Not Superscript/ Subscript	
Deleted: , creating	
Deleted: both	
Deleted: at site	
Formatted: Heading 3	
Deleted: perceived weaknesses in	
Deleted: use	
Deleted: driers in this application	
Deleted: used in Yver Kwok et al. (2015).	
Formatted: Subscript	

Deleted: S1

temperature) before decreasing to pre-injection concentrations. The effect of this changing H_2O concentration on the raw (without the inbuilt H_2O correction) CO_2 and CH_4 concentrations was then observed. The experiment was repeated in at least triplicate annually.

Data collected in the first five minutes immediately following the injection, the typical line equilibration

- 5 period, were excluded from the fit. This avoids using data adversely effected by the effect of rapid changes in H₂O content on the cell pressure sensor, as identified by Reum et al. (2018) and the erroneous postinjection CO₂ enhancement identified by (Rella et al., 2013). Again, due to cell pressure sensor concerns, data points with minute-mean H₂O standard deviations > 0.5 % H₂O were excluded. This 5-minute cutoff reduced the maximum H₂O value included in the fit to 4 % H₂O.
- 10 A water correction was then determined from a fit between the "wet"/mean "dry" ratio and the H₂O of the droplet test data and the equation given by Rella (2010). Here we defined "dry" data as any data with $H_2O < 0,003$ %, as measured by the CRDS, and the remaining data as "wet". We use minute mean uncorrected CRDS CO₂ and CH₄ data for this analysis, that is, minute averaged data from the "co2_wet" and "ch4 wet" columns of the raw Picarro data files along with data from the "h2o" column. This H₂O
- 15 data, unlike the "h2o_reported" data has been corrected for spectral self-broadening as detailed in Rella (2010). <u>A similar analysis was conducted for CO however this used the "co" data, which has water vapour</u> and line interference corrections applied to it. The raw co values (i.e. "co_wet") are not provided in the CRDS output files.

The fit was conducted using orthogonal distance regression weighted by both the minute mean standard

20 deviation of the H₂O and gas of interest (CO₂ or CH₄). The resulting correction parameters are shown in Table 1. These corrections were then applied to minute mean observational data through the GCWerks software completely bypassing the built-in CO₂ and CH₄ water corrections.

As Picarro analysers are not calibrated for H₂O measurements when measuring dry air they often show different positive or negative values close to zero. These "zero-water" values were 0.00001, -0.0003 and

25 <u>-0.002 for the Bilsdale, Heathfield and University of Bristol laboratory instruments respectively. These values were determined using measurements of cylinders of dry air where the first 120 minutes were ignored and the "zero-water" value calculated as the mean H₂O of the subsequent data (> 60 min).</u>

· · · · · · · · · · · · · · · · · · ·	_
Formatted: Subscript	
Deleted: 001	
Formatted: Not Highlight	
Deleted: Data collected in the first 5 minutes immediately	
following the injection and data with minute mean standard	
deviations > 0.5 % H ₂ O, 0.05 μ mol mol ⁻¹ CO ₂ or 0.4 nmol mol ⁻¹	¹ CI
were excluded from the fit.	
×	
Deleted:	

Formatted: Tab stone: 11 75 cm Loft

2.3.5 Temporal stability and mole fraction dependence of instrument specific water corrections

The typical temporal stability and mole fraction dependence of the CRDS water correction was examined using a laboratory based CRDS (G2301, Picarro Inc., USA; CO₂, CH₄ and H₂O series). Here the water correction was determined using the droplet experiment, as described in Section_2.3.4. The mid-term and short-term stabilities were examined by repeating the experiment approximately weekly over a three-month period and daily for a 5-day period using a cylinder of dried ambient mole fraction air. A set of instrument specific water corrections was also determined in triplicate, using dried sub- and above ambient air mole fraction cylinders. As this instrument was not able to measure CO the effect of CO mole fraction on the CRDS instrument specific water correction is not addressed in this paper.

10 2.3.6 Assessing the CRDS water correction

5

The CRDS water correction was also assessed through a series of simple Dew Point Generator (DPG; Licor LI-610 Portable Dew Point Generator, USA) experiments. Here four cylinders of dry air with varying CO₂ and CH₄ mole fractions (Table 2) were humidified to a range of dew points between 2.5 and $30 \degree$ C (0.6 to 3.5 % H₂O) and measured, with and without cryogenic drying, using a Picarro G2301 CRDS

- 15 at the University of Bristol (UoB) laboratory. Here the difference between the reported mole fractions (with instrument specific water correction applied) for the dried and wet air streams was interpreted as the error in the CRDS water correction. Exact details of the experimental set up can be found in the supplement Section \$1 and Figure \$3.
- 2.3.7 Quantifying CO₂ and CH₄ cross membrane transport using measurements of the counter 20 purge.gas

Experimental details

An experiment was designed observe gas exchange across the Nafion® membrane by measuring the counter purge gas before (CP_{in}) and after (CP_{out}) the Nafion® while varying the water and CO₂ and CH₄ content of the sample gas stream. This experiment followed a series of inconclusive experiments which

25 aimed to quantify gas exchange across the Nafion® membrane by directly measuring changes in the sample stream.

Formatted: Not Highlight

Formatted: Not Highlight

Deleted: .One of the outputs of V1 went via the Nafion dry and Nafion wet sample streams.3.2.4 Experiment 3 — Quantifying CO₂ and CH₄ cross membrane transport using measurements of the counter purge gas[¶] Finally, a third experiment was conducted. It aimed to

Deleted:

Deleted: (CP_{out}). For this experiment, two changes were made to the DPG sample flow path described previously (Figure 7). Firstly, the water trap was moved from its original location and placed immediately prior to the CRDS intake. This ensured that any difference in CP_m and CP_{out} was driven solely by transport across the Nafion membrane rather than by an artefact of differing H₂O mole fractions. Secondly.)



In this experiment, a system (Figure 3) was constructed allowing the controlled humidification, using a DPG of two high-pressure cylinders one of dry near ambient and one above ambient CO_2 and CH_4 mole fraction (Table 2; UoB-15 and UoB-16). These humidified air samples were measured using the UoB laboratory Picarro CRDS. Cylinder delivery pressure was controlled using a single stage high purity

- 5 stainless steel Parker Veriflo regulator (95930S4PV3304, Parker Balston, USA) or a TESCOM regulator (64-2640KA411, Tescom Europe). The DPG was used to humidify the cylinder air to a range of dew points between 5 and 25 °C, equating to water contents of between 0.786 ± 0.001 and 2.883 ± 0.003 % H₂O. The DPG was also bypassed allowing the dry cylinder air (< 0.0001 % H₂O, dew point < -70 °C) to be measured directly.</p>
- 10 The full system flow path is shown in Figure 3, but in summary the output of the cylinder regulator was plumbed to the input of the DPG. A T-piece (T1) connected prior to the DPG input vented any excess gas via a flow meter (F1) ensuring that the DPG input remained at close to ambient atmospheric pressure throughout the experiment. The output of the DPG passed through a second T-piece (T2) with the overflow outlet also connected to a flow meter (F2) to monitor that the CRDS input pressure remained
- 15 near ambient. Typical output flows were 0.1 and 0.3 L min⁻¹ for F1 and F2 respectively. Following this the humidified cylinder air was further split using a T-piece (T3), with half the flow passing through the Nafion® before reaching a 4-port 2-position valve, V1 (EUDA-2C6UWEPH, VICI Valco AG International, Switzerland, actually a 6-port valve configured as a 4-port valve). The other half bypassed the Nafion® and connected directly to V1. The first output of V1 connected to a multiport valve
- 20 (EUTA-CSD10MWEPH, VICI Valco AG International, Switzerland) valve while the second connected to a pump (PICARRO Vacuum pump S/N PB2K966-A) set to a flow rate matching that of the CRDS (0.3 L/min) to ensure uniform flow through both branches of the system. The V1 was controlled manually using a VALCO electronic controller and universal actuator while the multiport valve was controlled by the GCWerks software. The output of the multiport valve was connected to the CRDS via a cryogenic
- 25 water trap. The cryogenic water trap consisted of a coil of ¹/₄." diameter (I.D. <u>0.12</u>.") stainless steel tubing immersed in a Dewar of silicone oil (Thermo Haake SIL 100, Thermo Fisher Scientific, USA). The silicone oil was cooled using an immersion probe (CC-65, NESLAB) to less than -50 °C.

Deleted: 7

-(Deleted: Dew Point Generator (
•(Deleted: ; Licor LI-610 Portable Dew Point Generator, USA)
•(Deleted: 4
•(Deleted: a Picarro G2301 CRDS at the University of Bristol (
•(Deleted:) laboratory

Deleted: 7

Deleted: 3.36mm

Counter purge air, both before (CP_{in}) and after (CP_{out}) the Nafion® were also sampled using the multiport valve. To do this a T-piece (T4) was placed on the counter purge tubing prior to the Nafion® connecting to the multiport valve while a second T-piece (T5) located after the Nafion® was again connected to the multiport valve. Two flowmeters, F3 and F4, were used to monitor the counter purge flow. Flowmeter F3

5 was placed on the outflow of the Nafion® counter purge prior to T5 while a second F4 was connected to one output branch of T5. These flowmeters had a flow range of 0.1-0.5 L min⁻¹ (FR2A12BVBN-CP, Cole-Palmer, USA). When not sampling the counter purge F3 and F4 had flow rates of 0.4 L min⁻¹, when sampling CP_{out} the F4 flowrate dropped to 0.2 L min⁻¹.

Other than the cryogenic water trap and two short sections (< 10 cm) of 1/4_" (O.D.) "Bev-a-line" plastic 10 tubing immediately prior to and post the DPG, 1/16 " stainless steel tubing was used throughout the system. Due to the air output and input connections of the DPG the use of the plastic tubing was unavoidable.

The experiment was conducted in a temperature-controlled laboratory at 19 °C, and thus, at temperatures lower than a number of the dew points used within the experiment. Hence, in order to avoid condensation

- 15 forming on the walls of the tubing, all components of the cylinder air flow path between the DPG and the multiport valve, excluding the water trap, and the pump were contained within a chamber heated to > 32 °C. Tubing between the heated chamber and the input of the CRDS was also heated with heating tape to > 32 °C while the internal temperature of the CRDS was > 32 °C throughout the experiment. The multiport valve was heated to > 25 °C hence only dew points below 25 °C (2.9 % H₂O) were used in this experiment.
- 20 As the reliability of CRDS water correction was also under investigation it was important to isolate the effect of the Nafion® from that of the CRDS water correction. To do this the experiment was conducted in three stages (see Figure S4). Firstly, the H₂O content of the DPG humidified sample stream was allowed to stabilise (Figure S4 purple). A stable water content was defined as one where the standard deviation of the minute mean values was <0.003 % H₂O for a 15-min period, During this period the H₂O trap remained
- 25 out of the Dewar of silicone oil and the CRDS measured an undried, Nafion® bypassed sample, while the secondary pump maintained the flow of DPG sample through the Nafion®. After this criterion was reached the second stage was commenced, Here the H₂O trap was inserted into the silicone oil and the water content monitored until 10 minutes of dry air (defined as < 0.002 % H₂O) was obtained (Figure S4

Deleted: u

Deleted: 5

Moved (insertion) [8]

the standard deviation of the minute mean values was $< 0.003 \ \% H_2O$	Moved up	[8]: A stable water content was defined as one where
	the standard	deviation of the minute mean values was < 0.003 % H2O
for a 15-min period.	for a 15-min	period.

Deleted: H₂O trap was removed from the Dewar of silicone oil and the CRDS measured an undried, Nafion® bypassed sample, while the secondary pump maintained the flow of DPG sample through the Nafion®

Deleted: Secondly

Deleted: cylinder

<u>grey</u>). Together these two stages took typically 2 to 3 hours to complete — allowing the Nafion® time to equilibrate while ensuring that the H₂O trap was drying the sample and the DPG had reached the required set point. The multiport valve was then used to switch between the CP_{in} or CP_{out} flows, measuring each for repeated 20-minute blocks (n > 3) at each dew point (see Figure, S4, red and blue). The experiment

5 was also repeated with the DPG excluded and the cylinder of dried air measured directly, a water content of < 0.0001 % equating to a dew point of < -70 °C.

It is important to note that the DPG was not independently calibrated but the H₂O concentration was measured directly by the CRDS during the initial part of the experiment. These values were used as the reference H₂O concentration in all calculations and plots. Flow rates, cylinder pressure, chamber

10 temperature and H₂O trap temperature were manually logged after each valve position change and when the water trap was inserted into the silicone oil bath.

Data processing

All CO₂ and CH₄ data were corrected using the instrument specific water correction (Section 2,3.4). Minute mean values of all data were calculated from the raw 0.4 Hz data and exported from the GCWerks

15 software. Data processing was completed using code written using the Anaconda distribution of the Python programming language (Python Software Foundation, 2017;van Rossum, 1995) and a variety of standard packages including NumPy1.11.1 (Walt et al., 2011), SciPy 0.18.1 (Jones et al., 2001) and Matplotlib 2.0.2 (Hunter, 2007).

The counter purge measurements made during the humidification experiments represent a combination 20 of effects.

$$CP_{in} = True_{CP} \tag{1}$$

$$CP_{out} = True_{CP} - N_{X\%} \tag{2}$$

Where,

*True*_{cp} the true mole fraction of the counter purge gas

25 $N_{X\%}$ is the effect of the Nafion® at $X_2\%$ H₂O in the sample stream X% is the water content of the sample gas before the Nafion®

(Deleted:
(Deleted: X
(Formatted: Not Highlight
(Deleted: f
(Formatted: Not Highlight

Deleted: 1.3

Hence the difference between the mean of CP_{in} and the mean of CP_{out} represents both any transport of CO₂ (or CH₄) through the Nafion® membrane and the effect of the water correction.

In order to remove any valve switching or line equilibration effects the first 5 mins of data of each sample period was discarded and the mean of the final 15-minute period of each sample type at each dew point

- 5 was calculated. The uncertainty of this mean was determined as the 95_% confidence interval based on the larger of either the standard deviation of the minute means or average of the standard deviations of the minute means. Examples of the raw data collected during the experiment are given in Figure <u>\$4</u>. As the experiment was subject to a small temporal drift the mean CP_{in} values were linearly interpolated and the *CP_{out}* - *CP_{in}* difference calculated as the difference between the CP_{out} and time adjusted CP_{in} values
- 10 and the uncertainty estimated as the combined uncertainty of the CP_{in} and CP_{out} values.

Key experimental assumptions

These experiments assume that any changes in the CO_2 or CH_4 mole fraction are driven solely by the Nafion® drying processes. Other possible sources of error or bias included, adsorption and desorption effects within the regulator and walls of the tubing, gas solubility within the condenser of the dew point

15 generator and instrumental drift.

Regulator and tubing adsorption and desorption effects has been previously examined by Zellweger and Steinbacher (2017, personal communication). They found that for Parker Veriflo type regulators, as used in this experiment, the effects can be quite large, up to 0.5 μ mol mol⁻¹ CO₂ or 2 nmol mol⁻¹ CH₄. But that these effects were only evident at flow rates < 250 ml min⁻¹ and after significant periods of stagnation (15

20 hours). Considering the high flow rates (> 1 L min⁻¹) and long flushing times (2 to 3 hours) used in our experiment it is highly unlikely that regulator effects would make a significant impact on the results. As discussed earlier, a lengthy equilibration period was used at the start of each DPG run and following any change in DPG set point. This was to account for the dissolution of sample gas, in particular CO₂, in

the DPG water chamber. After this initial equilibrium period there were no rapid changes in the CO₂ mole 25 fraction with only a slow drift, apparent in the data. CRDS instrumental drift is also typically very small

and slow. For the UoB CRDS instrument, long-term measurements of target style standard cylinders have shown the drift to be $< 0.001 \ \mu mol \ mol^{-1} \ day^{-1} \ CO_2$ and $< 0.03 \ nmol \ mol^{-1} \ day^{-1} \ CH_4$. These drift rates are at least an order of magnitude smaller than the mole fraction differences observed in this study.

13

Formatted: Not Highlight

Deleted: 8

Deleted: in Section 3.2.2

Although small, any time dependent drifts were accounted for by temporally interpolating between each block of data. Also key to the design of this experiment is the examination of differences between two very similar mole fractions rather than absolute mole fractions. As such, any systematic errors that might drive a systematic offset cancel out and any mole fraction depended biases are minimised.

5 2.3.8 Calibration and traceability

Calibration procedures for both the CRDS and GC-ECD are as described in detail in Stanley et al. (2018). In brief, CRDS measurements are calibrated <u>using a close-to-ambient standard ("working tank"</u>) and a set of three calibration cylinders, which span the typical ambient range (Table <u>3</u>). Only a small number of elevated observations, < 0.4 % of the CO₂ and < 1.5 % of the CH₄ minute mean observations, were

10 outside the range of the calibration cylinders. However, a much higher proportion of the CO observations were outside the range of the calibration suites used at site, 28 % at BSD and 43 % at HFD, with the majority of these data points (> 98 %) below the lowest calibration cylinder.

Assigning mole fractions to values outside the range of the calibration suite <u>will increase</u> the error. The magnitude of this error will depend on the magnitude of the mole fraction difference between the closest

- 15 calibration cylinder and the sample. This error has been estimated using measurements made at the Heathfield site of cylinders of known CO mole fractions, 6 and 57 nmol mol⁻¹ CO below the lowest calibration cylinder. These show a percentage error of 2.41 and 3.09 %, respectively. A similar assessment of the error associated with samples above the highest calibration standard were made using cylinders 87 and 686 nmol mol⁻¹ CO above the highest calibration standard. These correspond to percentage errors of
- 20 2.98 and 2.56 %, respectively. As all the minute-mean CO measurements below of the calibration range are within 57 nmol mol⁻¹ of the lowest calibration cylinder and the vast majority of minute-mean CO measurements above the calibration range are within 686 nmol mol⁻¹ of the highest calibration cylinder (99 %) we expect that this error would typically be < 3 %.</p>

Daily measurements of the ambient standard are used to account for any linear drift, while monthly 25 measurements of the calibration suite are used to characterise the nonlinear instrumental response. This calibration procedure is controlled by the GCWerks software and allows near real-time examination of calibrated data, Deleted: against both

Deleted: 2

Deleted:

Deleted: greatly increases

Deleted: estimates.	and should be reflected in increased uncertainty
Formatte	d: Superscript
Formatte	d: Superscript
Formatte	d: Not Highlight
Formatte	d: Not Highlight
Formatte	d: Not Highlight
Formatte	d: Not Highlight
Formatte	d: Superscript
Formatte	d: Not Highlight
Formatte	d: Superscript
Formatte	d: Not Highlight
Formatte	d: Highlight
<u>All</u> CRDS standards and calibration gases are composed of natural air, <u>some</u> spiked or diluted with scrubbed natural air (TOC gas generator, Model No. 78-40-220, Parker Balston, USA) to achieve the required concentrations of CO₂, CH₄ and CO. <u>All standard cylinders were filled at Mace Head with well-</u> mixed Northern Hemisphere air. The cylinder spiking and filling techniques of the calibration cylinders

- 5 varied. The Heathfield calibration suite and the second Bilsdale calibration suite were filled at GasLab MPI-BGC Jena and consisted of natural air spiked using a combination of pure CO₂ and a commercial mixture of 2.5 % CH_d and 0.5 % CO in synthetic air. The "high" calibration cylinder of the first calibration suite used at the Bilsdale site was filled with peak-hour ambient air at EMPA, Dübendorf, Switzerland while the "low" and "mid" cylinders were based on Mace Head air, in the case of the "low" this was
- 10 diluted with scrubbed natural air. Using natural air based calibration and standard gases removes any j pressure broadening effects inherent in the use of non-matrix matched artificial standards (Nara et al., 2012). As the CRDS is an isotopologue-specific method filling the cylinders in such a manner ensures j that the isotopic composition was as close to those of the sampled air as possible. The effect of an isotopic j mismatch between the calibration standards and the sample has been examined in detail by Flores et al. /
- 15 (2017) Griffith (2018) and Tans et al. (2017) With Griffith (2018) showing that, for a sample of 400 μ mol mol⁻¹ CO₂ and 2000 nmol mol⁻¹ CH₄, the error will range between 0.001 0.155 μ mol mol⁻¹ CO₂ and 0.1 0.7 nmol mol⁻¹ CH₄ depending on the magnitude of the sample to standard mismatch. As such, we expect a worst-case scenario estimate of the error associated with our measurements < 0.04 % for both CO₂ and CH₄.
- 20 GC-ECD measurements are made relative to a natural air standard of known N₂O and SF₆ concentration. This standard is measured hourly and used to linearly correct the samples (Table <u>4</u>). The instrumental nonlinearity response was characterised prior to deployment by dynamically diluting a high concentration standard with zero air and was repeated in the field at the BSD site on 30th September 2015. This approach, dynamic dilution, has a history of use in similar field locations (Hammer et al., 2008) and is able to
- 25 generate multiple calibration points using just two cylinders. This greatly reduces the number of cylinders needed, a key concern for space-limited locations like BSD and HFD. An assessment of the uncertainty associated with this non-linearity approach will be included in a future paper currently in preparation. However, previous studies (Hall et al., 2011;van der Laan et al., 2009;Hammer et al., 2008) have found

Deleted: re all

Formatted: Subscript
Formatted: Subscript
Deleted: This
Deleted: isotopologue
Deleted: the calibration and
Deleted: standard
Deleted: were filled at Mace Head with well-mixed Northern Hemisphere air in an effort to
Formatted: Not Highlight
Deleted: e ratios of the standard and calibration gases were
Formatted: Not Highlight
Deleted: (seeChen et al. (2010);Lee et al. (2006);Nara et al. (2012)
Formatted: Not Highlight
Formatted: Not Superscript/ Subscript
Formatted: Subscript
Formatted: Not Superscript/ Subscript
Formatted: Subscript
Formatted: Not Superscript/ Subscript
Formatted: Subscript
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Superscript/ Subscript
Formatted: Subscript
Formatted: Subscript
Deleted: for further discussion).
Formatted: Highlight
Deleted: 3
Formatted: Font color: Auto

the ECD detector response to be extremely stable over time and very linear for both $SF_{\underline{6}}$ and $N_{\underline{2}}O$ in the mole fraction range typical of the HFD and BSD stations. As such, we expect the uncertainty of the nonlinearity correction to be very small.

GC-ECD and CRDS standards and calibration cylinders were, where possible, calibrated both before and

- 5 after deployment at the sites. If these two measurements agreed then a mean mole fraction was used, otherwise a linearly drift corrected mole fraction was used. The CRDS cylinders were calibrated through WMO linked calibration centres (either WCC-EMPA, Dübendorf, Switzerland or GasLab MPI-BGC, MPI, Jena, Germany). This ties the ambient measurements to the WMO CO₂ x2007 (Zhao and Tans, 2006), CH₄ x2004A (Dlugokencky et al., 2005) and CO x2014A (Novelli et al., 1991) scales. The
- 10 calibration of the GC-ECD standards was conducted at either the AGAGE Mace Head laboratory or the University of Bristol laboratory and are reported here on the recently released SIO-16 N₂O scale and the SIO SF₆ scale. Most cylinders were or will be calibrated before and after deployment and the mean of the two values used. Some cylinders, due to logistical constrains were only calibrated once (Table <u>4</u>).

2.3.9 Instrument short-term precision and long-term repeatability

- 15 The short-term (1-minute) precision of the CRDS data was determined as the mean of the standard deviations of the 1-minute mean data. This was calculated <u>from</u> measurements of the standard cylinder and the calibration suite allowing the relationship between CO₂, CH₄ and CO mole fraction and short-term precision to be examined. <u>This analysis included 18 cylinders covering a wide range of mole fractions (Table 3).</u>
- 20 The mean absolute short-term precision for all cylinders was consistent between the two sites across all three gases. At BSD the short-term precision was 0.024 µmol mol⁻¹ CO₂, 0.18 nmol mol⁻¹ CH₄ and 4.2 nmol mol⁻¹ CO while at HFD it was 0.021 µmol mol⁻¹ CO₂, 0.22 nmol mol⁻¹ CH₄ and 6 nmol mol⁻¹ CO₂ Both sites showed a small trend with the mean absolute precision worsening with increasing CO₂ and CH₄ mole fraction. However, this was not observed in the relative precision which remained unchanged
- 25 at ~ 0.005 % for CO₂ and ~ 0.01 % for CH₄. This was not the case for CO where the relative precision improved with increasing mole fraction from ~ 4 % at CO < 100 nmol mol⁻¹ to < 1.5 % at CO > 250 nmol mol⁻¹. We suspect that this tendency is inherent in the spectroscopic approach as the CO peak measured

J	Formatted:	Subscript
Î	Formatted:	Subscript

X

Deleted: 215

Deleted: for

Deleted: n

Deleted: 3

Moved (insertion) [7] Deleted: 2 Deleted: Long-term repeatability of 20 minute means were calculated by examining the standard deviation of the daily standard cylinder measurements. Moved up [7]: This analysis included 18 cylinders covering a wide range of mole fractions (Table 2). Deleted: ± 0.004 Deleted: ± 0.04 Deleted: ± 0.04 Deleted: ± 0.04 Deleted: ± 0.04 Deleted: ± 1 Deleted: ±1 Deleted: ±1 Deleted: (x ± 1σ) Deleted: increasing (i.e. becoming less precise)

by the Picarro CRDS is much smaller than those of the CO₂ and CH₄ (Chen et al., 2013) and hence more susceptible to noise in the baseline particularly at low mole fractions.

The long-term reproducibility of a 20-minute mean was estimated as the mean standard deviation of the daily 20-minute measurements of 4 standard cylinders at each site. Like short-term precision, mean long-

term repeatability (calculated over a period of approximately a year) is consistent between the two sites, 5 0.018 and 0.013 µmol mol⁻¹ CO₂, 0.20 and 0.20 µmol mol⁻¹ CH₄ and 1.1 and 1.7 µmol mol⁻¹ CO at BSD and HFD respectively.

Repeatability of individual injections on the GC instruments were calculated as the standard deviation of the hourly standard injection. These were found to be < 0.3 nmol mol⁻¹ and < 0.05 pmol mol⁻¹, for N₂O and SF₆ respectively, and did not differ between the two sites.

10

2.4 Data analysis

2.4.1 Data quality control

A three-stage data flagging and quality control system was used for the HFD and BSD data. Initially, automated flags based on the stability of key parameters including cell temperature & pressure and

- 15 instrument cycle time (the time taken to collect and process each measurement) were applied. Here, data with a cycle time > 8 seconds were filtered out along with any data with cell temperature outside the range 45 ± 0.02 °C or cell pressure outside 140 ± 0.1 Torr. Secondly, a daily manual examination of the GC chromatograms and key GC/CRDS parameter values of each site were made. Data points were flagged if instrument parameters varied beyond thresholds determined to reduce their accuracy and a
- 20 reason for the removal was logged. Finally, all sites were reviewed simultaneously and the mixing ratio of the same gas from each site are overlaid to look for differences between sites. Any significant differences between the background values at each site were investigated by examining key instrumental parameters, calibration pathways and 4-hourly air mass history maps to ensure that these differences represent true signals rather than instrumental or calibration driven artefacts. The hourly air
- 25 mass history maps were produced using the Numerical Atmospheric dispersion Modelling Environment (NAME) Lagrangian dispersion model (Manning et al., 2011).

Formatted: Font color: Text 1

Deleted: ± 0.009	
Deleted: ± 0.001	
Deleted: ± 0.04	
Deleted: ± 0.01	
Deleted: ± 0.3	
Deleted: ± 0.3	
Deleted: $(\overline{x} \pm 1\sigma)$	

Deleted: - quality controlstatistical processing

Deleted: pressure	
Deleted: Torr	
Deleted: temperatures	
Deleted: °C.	



2.4.2 Statistical processing, baseline fitting and seasonal cycles

The long term trend in mole fraction at each site was estimated as the mean linear trend in the minute mean data over the period 2014-2017, inclusive, Seasonal and diurnal trends in the data were assessed using monthly and hour-of-day box plots of hourly means of detrended minute-mean data developed using

5 <u>the Python matplotlib.boxplot package. Here the long-term trend was removed by using a least-squares</u> <u>fit between a quadratic and the minute mean data.</u>

3 **Results and discussion**

3.1 CO₂, CH₄ and CO key features

The minute mean CO₂ observations range between a low of 379.50 to a high of 497.48 $\mu mol\ mol^{-1}\ CO_2$

- 10 at Heathfield and 379.77 to 587.17 μmol mol⁻¹ CO₂ at Bilsdale. High CO₂ mole fractions observed at BSD are generally higher than those of the HFD site (Figures <u>4a & 5a</u>). The high mole fraction events observed at BSD are generally sporadic lasting only a couple of hours and appear as a brief pulse relative to the normal diurnal cycle; a pattern indicative of a nearby point source. Considering BSD is remote from large conurbations, measured signals are expected to be dominated by biogenic sources. In
- 15 this instance, we suspect high mole fraction events at BSD are due to local heather (*Calluna vulgaris*) burning. These CO₂ events also typically coincide with periods of elevated CH₄ and CO, again suggesting a biomass burning source. While events that do not show corresponding high CO mole fractions, the majority of which occur in the higher two intakes, are likely to be driven by more remote CO₂ sources, for example power plants.
- 20 JHFD is located in southern England, just south of London (Figure 1). Here, high CO₂ events are typically longer 2 to 3 days and coincide with elevated CH₄, CO, N₂O and SF₆. Rather than appearing as peaks superimposed on a background value these periods have a positive shift in the entire diurnal cycle, suggesting a change in the background mole fraction. Air histories, based on the output of the Numerical Atmospheric dispersion Modelling Environment (NAME) Lagrangian dispersion model, outlined in
- 25 Manning et al. (2011), for these periods of elevated CO₂ typically show the source of the air to be from over London or Europe.

Formatted: Not Highlight

Deleted: Annual seasonal cycles were determined as the difference between the statistically smoothed curve and the long-term trend. Diurnal cycles were calculated using detrended hourly mean data (i.e. hourly mean minus long term trend). These data were split into 24 hour-of-the-day bins for each season and the mean, 5th and 95th percentile of each bin determined. The mean value of all the hourly means at each sample height were then added to each bin mean for plotting.

Formatted: Not Highlight

Deleted: Statistical baselines and long term trends for each gas at each height were calculated using the CCGFILT Python package (Thoning 2015) This is based on the curve fitting technique originally used by Thoning et al. (1989). In this package the data is initially fitted to a function combining simple polynomial (the longterm trend) and harmonic functions (annual seasonality). Secondly, the statistical baseline (smoothed curve) is defined using the shortterm deviations from this function. These fit residuals are transformed into the frequency domain using a Fast Fourier Transform (FFT), filtered using a low-pass filter with a short term cut off of 80 days and a long term cut off of 667 days. The inverse FFT is then used to transform the data back into the time domain. To reduce the computational complexity of the fitting long term trends and the statistical baselines were determined using hourly CO2, CH4 and CO means, while the N2O and SF6 data were fit at their native 10-minute frequency.

Formatted: Not Highlight

Formatted: Not Highlight

Deleted: <#>2.2.3 Modelling

<#>The observations can be combined with the output from atmospheric transport models, prior information and uncertainty assessments of each of the components in an inverse modelling framework to estimate regional emissions of each gas. Examples of different inverse methods that have been [1]

Formatted: Heading 1

Deleted: 2.3.1	
Deleted: While the HFD mean CO ₂ mole fraction, 407.: ¹ , is slightly higher than the mean mole fraction of the BS	⁵ μmol mol ⁻ D site ₃ [2]
Deleted: 3a	
Deleted: 4a	
Deleted:	
Formatted: Subscript	
Deleted: In contrast,	
Formatted: Subscript	
Formatted: Subscript	
Formatted: Subscript	
Deleted: and	
Deleted: drive	
Deleted: This difference in pollution event patterns is n in the CH ₄ CO N ₂ O or SE data	ot apparent

Both sites show a clear <u>relationship between CO_2 mole fraction and intake height with the lowest height</u> generally having the most elevated mole fractions, followed by the higher heights (Figures <u>4a & 5a</u>). This trend, also apparent for CH₄ and CO (Figures 4b & c and Figure 5b & c), is typical of tall tower measurements and is driven by proximity to surface sources (Bakwin et al., 1998). The gradient in CO₂

5 and CH4 mole fraction is most apparent in the warmer seasons and during the early hours of the morning (Figure 6a, b, c & d) when the boundary layer is the lowest.

The timings <u>and magnitude</u> of the HFD and BSD seasonal cycles are similar, with CO₂ mole fractions highest in the colder months and lowest during the Northern Hemisphere <u>summer</u> (Figure <u>7a & b</u>). Although both sites are located in areas <u>consisting of predominantly agricultural space or native</u>

10 vegetation the HFD site is more urbanised. This appears to be reflected in more elevated CO₂ events in winter relative to the BSD site (Figure 6a & b). The HFD CO seasonal cycle also shows an increased prevalence of high events relative to BSD while the summer is consistent between the two sites (Figure 6e & f). Suggesting that the HFD site is more sensitive to fossil fuel emissions.

As with the seasonal cycle, the shape of the CO₂ diurnal cycle is similar at both sites, with mole fractions

15 peaking near sunrise and the lowest CO₂ mole fractions observed in the late afternoon (Figures <u>6a & b</u>). Again, the amplitude of these cycles varies between the sites with HFD, the more anthropogenically influenced site showing <u>a</u> higher maximum in the early morning <u>during</u>.

Although there is a very large range in the minute mean CH₄ observations, 1841 to 3065 nmol mol⁻¹ at BSD and 1843 to 3877 nmol mol⁻¹ at HFD, > 99.99 % of measurements, are less than 2400 nmol mol⁻¹

- 20 CH4, with only 6 events in the combined record exceeding this threshold. These events have been clipped from the data shown in Figures <u>4b</u> and <u>5b</u> for ease of viewing. Like CO₂, the CH4 observations show seasonal cycles with the mole fractions the highest in the winter months and the lowest in midsummer (Figure <u>7c & d</u>). <u>A small</u> CH4 diurnal cycle peaks in the morning usually 1 to 2 hours after sunrise (this is after the CO₂ maximum) and then dips in the mid-afternoon (Figures <u>6c & d</u>). The CH4 diurnal cycle
- 25 is also more pronounced and smoother in the HFD data and evident throughout the year, whereas the BSD cycle is only strongly apparent in the summer months. This could be linked to differences in the relative magnitude of key local sources/sinks of CH₄ between the two sites.

Deleted: pattern in mean Deleted: with Deleted: 3a Deleted: 4a Deleted: is also apparent for CH4 (and to a lesser degree CO), which is also measured at multiple intakes and is most likely driven by the closer proximity to surface sources (Figures 3c & e and Figure 4c & e) Deleted: Deleted: (see Deleted: 3b Deleted: & d Deleted: and Figure 4b Deleted: At these times Deleted: and CO2 and CH4 released from the surface sources during the night build up leading to relatively higher mole fractions at lower elevations. This difference disappears later in the day as the sun warms the atmosphere and thermal mixing causes the boundary laver to rise Formatted: Tab stops: 15 cm, Left Deleted: summer; however, the magnitude of the summer draw down appears approximately 25 % larger at BSD than HFD Deleted: 5a Deleted: of significant agriculture and green space Formatted: Subscript Deleted: a significantly reduced summer CO2 uptake and slightly larger Deleted: increase Deleted: enhanced Deleted: winter maximum Deleted: minimum Deleted: 3b Deleted: and Deleted: 4 Deleted: an approximately 3 µmol mol-1 CO2 Deleted: 3d Deleted: 4c Deleted: clear Deleted: 5c Deleted: Interestingly, there is also a gradient in the amplitude of 31 Deleted: Deleted: 3d Deleted: and 4

Of the 5 gases measured at HFD and BSD, CO is the only gas to show a decrease in mole fraction between 2013 and 2017, roughly -7 nmol mol⁻¹ yr⁻¹. In contrast, the CO₂ and CH₄ data increase by 2-3, µmol mol⁻¹ 1 yr⁻¹ and 5-9, nmol mol⁻¹ yr⁻¹ respectively, varying on the intake height. These agree well with the ~2 µmol CO₂ mol⁻¹ yr⁻¹ and ~8 nmol CH₄ mol⁻¹ yr⁻¹ trends observed at Mace Head (MHD, 53.327 °N, -9.904)

5 °E, Figure 1), a remote site within the UK DECC network located on the west coast of Ireland. The CO data collected at MHD is not on the NOAA x2014 CO calibration scale making direct comparisons between growth rates meaningless.

While the range of minute mean CO mole fractions was significantly larger at BSD, 63 to 9500 nmol mol⁻¹, than HFD, 60 to 4850 nmol mol⁻¹, the high CO values observed at BSD were relatively rare. This is

10 reflected in the smaller quantile spread of the BSD data compared with the HFD data (Figure 7e & f).

3.2 N₂O and SF₆ key features

The range of N_2O mole fractions observed from the two intakes of comparable height, 108 m at BSD and 100 m at HFD, were very similar, 326.6 to 340.0 and 326.4 to 338.5 nmol mol⁻¹ for BSD and HFD, respectively (Figures <u>4d</u> and <u>5d</u>). The N₂O data from the higher (248 m) intake at BSD, has parrower

15 range, <u>especially in the cooler months of the year than the lower 108m data (Figure 6g). As described earlier the smaller range in the 248 m data is typical of tall tower measurements and driven by increased mixing with increasing altitude, which reduces the influence of local point sources.</u>

The N₂O mole fraction seasonal cycle of both sites shows an unusual pattern with two maxima per year, one in early spring and a second in autumn (Figure 7g & h). Both the timings and amplitudes of these

20 cycles, are similar at both sites. The long-term trend, ~0.8 nmol N₂O mol⁻¹yr⁻¹ (calculated using data from the 108 m and 100 m intakes at BSD and HFD over the period of coincident data collection, 2014 to mid 2016) also agrees well between the two sites and with MHD, also ~0.8 nmol N₂O mol⁻¹yr⁻¹.
 A previous study. Nevison et al. (2011) examined the monthly mean N₂O seasonality of baseline mole

fraction data at Mace Head (MHD, 53.327 °N, -9.904 °E, Figure 1), a remote site within the UK DECC

25 network located on the west coast of Ireland. They found that although biogeochemical cycles predict a single thermally driven summer time maximum in N₂O flux (and hence mole fraction) (Bouwman and Taylor, 1996), they actually observed a late summer minimum, with a single N₂O concentration peak in

Deleted: A known weakness of the statistical fitting procedure is that, due to the FFT used, any gaps in the record must be filled using interpolation before analysis. This can lead to spuriously high (or low) values near gaps within the fitted seasonal cycle if the ends of the gaps occur during a significantly elevated (or reduced) event. These "end effects" are very evident in the HFD 50 m intake record for CH4 (and CO and SF6) in late 2017 leading to the large spikes in the seasonal cycle evident in Figure 5c, e and 1.1 Of the 5 gases measured at HFD and BSD, CO is the only gas to show a decrease in the long-termn trend of the statistically defined baselineole fraction between 2013 and 2017, roughly-6 mmol mol ¹¹ yr ⁻¹ (Figure 5f)
Deleted: approximately 3µmol mol ⁻¹ yr ⁻¹ and 7 [5]
Formatted: Not Highlight
Formatted: Not Highlight
Deleted:
Formatted [6]
Deleted: (Figure 5b and d)
Formatted[7]
Deleted: . The HFD long-term trend curves in CO2, CO and CH4

Deleted: The HFD long-term trend curves in CO₂, CO and CH₄ are generally positively offster lealive to BSD - most likely driven by the prevalence of polluted air at the HFD site. The ... hile the range of minute mean CO mole fractions was significantly larger at a...St BS..., D averaged 269 mm0 mol³ but varied between ... 3 and ... [5]

Formatted: Not Superscript/ Subscript

Deleted: . The mean HFD CO minute mean value was slightly higher, 277 nmol mol⁻¹, however the...FD, CO range was significantly narrower, ...0 to 4850 nmol mol⁻¹, t. [9]

Deleted: 2.3.2

Formatted: Heading 2

Deleted: Mean ...₂O mole fractions observed from the two intakes of comparable height, 108 m at BSD and 100 m at HFD, were found to be identical (329.5 nmol mol⁻¹) and the range of N₂O observations were also ...ery similar, 326.6 to 340.0 and 326.4 to 338.5 nmol mol⁻¹ for BSD and HFD, respectively (Figures 3g ...d and 4g...d). The N₂O data from the higher (248 m) intake at BSD, has has not only a higher mean value, 331.1 nmol mol⁻¹ but a ...arrower range, 329.3 to 335.0 nmol mol⁻¹ ... [10]

(Formatted: Not Highlight

Deleted: The 248m data analysed in this paper was from the 17th March 2017 (compared with late 2013 for the other heights) till the 19th June 2017. Hence this data is likely to be higher in N₂O relative1 **Deleted:** 5g

Eormatted

V	ronnatteu	[12]		
Ì	Deleted:are similar at both sites. and the long-term trend,			
ļί	~0.8 (Figure 5h)	[13]		
l	Formatted	[14]		
Ì	Deleted: are lso agrees well between the two sites and very			
V	similar between the two sites	[15]		
Ì	Deleted:			

spring. This was attributed to the winter intrusion of N₂O depleted stratospheric air and its delayed mixing into the lower troposphere. In contrast, a UK focused inversion study Ganesan et al. (2015), found that N₂O flux seasonality is driven not just by seasonal changes in temperature but by agricultural fertilizer application and post-rainfall emissions. They predict the largest net N₂O fluxes will occur between May

- 5 and August while agricultural fluxes will peak during spring for eastern England and summer time for central England. However, the exact timings of these fluxes can vary year-to-year as they depend not only on the scheduling of agricultural fertilizer application but on rainfall and temperature. Like MHD, BSD and HFD are expected to experience a decrease in N₂O driven by stratospheric intrusion, which would account for the springtime maximum and summer minimum. However, both BSD and HFD are located
- 10 much closer to significant agricultural sources of N₂O than MHD. Hence, it is likely that they are much more influenced by agricultural N₂O fluxes. As such, it is possible that although a summer time maximum in N₂O flux is completely offset by stratospheric intrusion, this summer time maximum may be so large that the residual autumn tail of this event appears as a second maximum at BSD and HFD.

Clear diurnal cycles in N₂O were observed at the HFD for the spring, summer and autumn months with the maximum N₂O mole fraction occurring 2 hours after sunrise and the minimum in the mid-afternoon

(Figure <u>6h</u>). These cycles were not <u>as apparent at BSD (Figure <u>6g</u>). The long-term trend in the SF₆ mole fraction at BSD and HFD shows a gradual increase of 0.3 pmol mol⁻¹ yr⁻¹ <u>again agreeing well with MHD which showed an identical growth rate</u>. Although the predominant sources of SF₆ are electrical switchgear, which is not expected to have significant seasonality, there was</u>

- 20 a small seasonal cycle observed (Figure <u>7i & j</u>). This cycle is more apparent in the <u>108 m BSD data and appears as a slight (0.1 to 0.15 pmol mol⁻¹) enhancement in SF₆ in the winter months. <u>This seasonal shift occurs across the wider DECC-GAUGE network and air history maps suggest that it is not associated with an obvious UK or continental region. As such, instead of an atmospheric transport driven shift we believe this to be a true change in emissions and hypothesise that this may be due to increased load on,</u></u>
- 25 and hence increased failure of, the electrical switchgear during the colder months. SF₆ mole fractions averaged 8.9 pmol mol⁻¹ at both BSD and HFD. While HFD, located closer to large conurbations than BSD, typically saw higher SF₆ pollution events. This was reflected in its larger range of 8.1 to 34.2 pmol mol⁻¹ compared with 8.1 to 22.9 pmol mol⁻¹ at BSD (Figures 4e and 5e).

Deleted: 3h)			
()			
Deleted: (Figu	re 5j)		
Deleted: 5i			
Deleted:			
Deleted: We			

Deleted: 3i

3.3 Site specific water corrections

The annually determined instrument specific water corrections are typically very similar at each site, often within the 95 % confidence interval of the triplicate runs (Table 1), suggesting that the corrections are fairly stable between years and instruments. The residuals of the instrument specific water corrections are

- 5 generally quite small, with 25th and 75th quartiles of -0.03 and 0.05, μ mol mol⁻¹ CO₂, and -0.4 and 0.3 pmol mol⁻¹ CH₄ (Table 1). The mean absolute residuals are, on average, smaller than those of the inbuilt correction and are notably smaller at higher H₂O content (see Figure S5). For example, the mean absolute residuals for 2015 data from HFD with H₂O > 2 % are 0.04 and 0.09 μ mol mol⁻¹ CO₂ and 0.4 and 1.2 nmol mol⁻¹ CH₄ for the new and inbuilt correction, respectively.
- 10 While instrument specific CO water corrections were calculated, the large minute-mean variability inherent in the G2401 CO measurements (> 4 nmol mol⁻¹) meant that the difference between data corrected using the instrument specific and in-built correction was not statistically significant. As such, these corrections were not presented in the body of the paper, however, further information can be found in Figure S5 of the supplementary.
- 15 Plots of the residuals typically show a common pattern, with the residual of zero at 0 % H₂O, before dipping below zero and then returning to zero at H₂O between 0.2 and 0.5 % (Figure <u>\$5</u>). Unlike other tests, the depth and width of this dip is more pronounced for BSD 2017. However, the BSD 2017 data both spans a wider range of H₂O contents than the earlier <u>BSD</u> tests (0 to 3.5 % <u>ys</u>. 0 to 2.2 %) and has far fewer data points in the 0.1 to 1 % H₂O range (0.9 % of all data points <u>ys</u>. 34 % and 27 % for BSD
- 20 2015 and 2016, respectively). The BSD 2017 0.1 to 1.0 % minute mean data also have an average standard deviation an order of magnitude larger than those of 2015 and 2016 (Figure <u>\$5a</u>, b & c). Refitting the BSD 2017 correction using only data H₂O < 2.2 % decreases the depth of the deviation by 0.05 µmol mol⁻¹ CO₂ and 0.3 nmol mol⁻¹ CH₄ as well as decreasing its width slightly but the deviation remains. This suggests that the presence of the dip is robust but the change in <u>its shape between 2017 and 2016 may</u>
- 25 well be a fitting artefact.

<u>Reum et al. (2018)</u> previously identified this pattern in water correction residuals and linked it to a pressure sensitivity at low water vapour mole fractions. They proposed an alternative fitting function incorporating the "pressure bend" although they do not recommend using this fit for data collected during

1	Moved a	own [1]: <#> cross membrane transport for ambient
_	Deleted:	<#>Experiment 1 —
	Moved d	own [6]: <#> At least 15 minutes of data were
	Deleted:	<#>3.2.2
1	Deleted:	<#>Quantifying CO2 and CH4 cross membrane[16]
$\langle \rangle$	Deleted:	<pre><#>with varying sample H2O mole fraction [17]</pre>
	Deleted:	<#>of
	Deleted:	<#>Nation
	Deleted	A areas membrane transport for embiant CO2 and of
	Deleted.	<pre>closs memorane transport for amolent CO2_mu[18]</pre>
	Maurad d	····· [19]
	Moved d	own [z]: <#>
	Deleted:	<#> [20]
	Deleted:	<#>.
	Moved d	own [3]: <#> ¶
	Moved d	own [4]: <#> Flow meters F1 and F2 had a range of
	Moved d	own [5]: <#> to a second identical valve, V2, while
	Deleted:	<#> [21]
	Deleted:	<#>.
	Deleted:	<pre><#> to a second identical valve, V2, while the second []</pre>
	Deleted:	<#> Flow meters F1 and F2 had a range of 0.1-1.L[22]
	Deleted:	<#>One of the outputs of V1 went via the Nafion
	Deleted:	<#> dry and Nafion wet sample streams.
	Deleted:	<#> At least 15 minutes of data were collected after that
	Deleted:	[24]
	Deleted:	#Sobserve gas exchange across the Nation membranes
	Deleted:	(#>) For this experiment, two changes were made that
	Deleted.	(#>). For this experiment, two enanges were made [27]
	Deleted.	
	Deleted:	*#>was placed between v2 and the CRDS intake, [29]
	Deleted:	<#> connected to the
	Deleted:	<#>multiport valve allowing the sampling
	Deleted:	<#> of the Nation
	Deleted:	<#>
	Deleted:	<#>counter purge
	Deleted:	<#>gas
	Deleted:	<#> prior to
	Deleted:	<#>(CPin) and after
	Deleted:	<#>the Nafion
	Deleted:	<#> (CPout). A fourth flowmeter (F4) was placed after)
	Deleted:	<#> to
	Deleted:	<#>heat
	Deleted:	<#>the
	Deleted:	<#> new
	Deleted:	<#> multiport valve to > 25 °C only dew points
	Deleted:	<#>of
	Deleted:	<#>25 °C
	Deleted:	<#>or less
	Deleted:	<pre><#> were used in this experiment.¶</pre>
	Formatted	
	Deleted:	====================================
	Deleted	<pre><#> experiment was conducted in three stages Firstbal</pre>
	Deleted	=
	Deleted:	stating
	Delatad:	
	Deleted:	-m-stage
	Deleted:	we use rize use removed from the Dewar of [34]
	Deleted:	\# <u>.</u>

the droplet test due to the paucity of stable data typically obtained between 0.02 and 0.5 % H₂O and the effect of rapidly changing H₂O on the cell pressure sensor. Implementing a more controlled water test at the sites would allow the use of the new fitting function. But due to the complexity of such a test this would be logistically difficult at remote field sites.

- 5 It is also important to note that the magnitude of the dip observed by <u>Reum et al. (2018)</u> in their controlled water tests, ~ 0.04 μmol mol⁻¹ CO₂ and <u>1</u> nmol mol⁻¹ CH₄, are roughly half those observed for the HFD, BDS and UoB droplet tests. As such the increased residuals observed for our water corrections between 0.02 and 0.5 % H₂O are likely to be primarily driven by the rapidly changing H₂O content inherent in the droplet test rather than represent a true error in the water correction.
- 10 The poor performance of the CRDS pressure sensor at low H₂O mole fractions, 0.02 to 0.5 % H₂O, is not expected to be a large source of error for undried samples as the majority of these, 92 % of the BSD and HFD data, contain > 0.5 % H₂O. But this is likely to be a large source of error for Nafion[®] dried samples where 95 % of HFD and 92 % of BSD are < 0.5 %. Calibration gases, although partially humidified to < 0.015 % H₂O, when flowed through a damp Nafion[®] are still far drier than Nafion[®] dried air samples
- 15 which averaged 0.2 % H₂O. As such this effect will not be accounted for as part of the calibration process. It is difficult to quantify this error, as it will vary with sample water content and the sensitivity of the individual instrument's pressure sensor to low H₂O mole fractions. However, for BSD and HFD₂ assuming that the residuals of the droplet water tests are an accurate reflection of the likely error (Figure S5), we expect there to be a systematic offset of the order of -0.05 to -0.1 µmol mol⁻¹ CO₂ and -1 to -2
- 20 nmol mol⁻¹ CH₄.

The sample mole fraction dependence of the CRDS water correction was examined by conducting water droplet tests using dry cylinders of above and below ambient mole fractions (Section <u>2.3.5</u>). Specific above and below ambient water corrections were calculated based on these data sets (Table1 and Figure <u>S6</u>). Jf the water correction was independent of sample mole fraction then the residuals should be identical

25 for both correction types. Although the above and below ambient residual plots are similar they do differ slightly with the residual of the above mole fraction sample becoming more positive at higher H₂O mole fractions while the below ambient mole fraction residuals become more negative. This is reflected in the

Deleted: Fitting the HFD and BSD data using this function reduced
the dip in the residual but did not significantly alter the mean
residual. However, due to the limitations of the droplet water test as
outlined by Reum et al. (2017) we have decided not to implement
this.

Deleted: Reum et al. (2017)

Deleted: 0.5

Deleted: 4 fold smaller than

Deleted: it is likely

-(Delete	ed: 3	1

1	Deleted: The difference (residuals) between these two types of corrected data and the true dry mole fraction of each cylinder was then determined.			
Deleted: However, a				
Deleted: CO2 and CH4 mean r				
Deleted: does not change significantly				
Ľ	Deleted: ,			
Ý	Deleted: shape of the CO2 residual changes slightly with the			

		difference in mean residuals and the shift in the interquartile ranges as seen for both CO2 and CH4 in	(Formatted: Subs	cript
		Table 1,	······(Formatted: Subs	cript
		The change in the difference between dry mole fractions calculated using the earliest instrument specific	(Deleted:	
		water correction and subsequent water corrections for each instrument with water concentration is shown			
	5	in Figure §a & b. For a typical air sample (1.5 % H ₂ O, 400 μ mol mol ⁻¹ CO ₂ and 2000 nmol mol ⁻¹ CH ₄)	(Deleted: 10	
		shifting between the annual water corrections drives CO_2 and CH_4 changes of < 0.05 μ mol mol ⁻¹ and < 1			
		nmol mol ⁻¹ . However, this difference does change with water content and can increase outside the WMO			
		reproducibility bounds at higher (> 2.5 %) H ₂ O contents. For example, the difference between CO ₂ dry			
		mole fractions calculated using the Bilsdale 2015 and 2017 H_2O correction increases to 0.12 µmol mol ⁻¹			
	10	at 2.5 $\%$ H ₂ O. It's also important to note that these differences will scale with CO ₂ and CH ₄ mole fraction.			
		Nevertheless, at the range of ambient water contents observed at BSD and HFD (0.1 to 2.5 %) these			
		differences remain below the WMO comparability guidelines (WMO, 2016) for CO ₂ and CH ₄ mole	(Deleted: (WMO, 2	2016)
		fractions $<750~\mu mol~mol^{-1}$ and $<4000~nmol~mol^{-1},$ respectively, as observed in BSD and HFD air			
		samples. In light of the temporal variability of the water correction over time, particularly at higher water			
	15	contents, using a Nafion® dryer or alternative drying method to obtain a relatively low and stable sample			
		water content would be an advantage,		Deleted: This is n	ot the
		A comparison of the individual daily and weekly tests, Figures &c & d and 10e & f, conducted using the		example 38 % at BS samples had H ₂ O <	SD an 0.1 %
		UoB instrument, show the daily tests to be far more similar than the weekly tests. That is, the variability		Deleted: 10	-
		over the 3-month period of the weekly test is much larger than that of the 5-day period of the daily test.			
	20	However, the variability of the weekly tests is similar to those of the annual tests, Figure & and b,		Deleted: 10	
		suggesting that, within the bounds of the data typically observed at the BSD and HFD sites, the use of			
		annually derived instrument specific water corrections are sufficient. This may not be the case at sites			
		with higher levels of humidity and CO_2 and CH_4 mole fractions where water corrections may need to be			
		determined more frequently, perhaps even weekly. The impracticality of such a frequent testing regime			
	25	along with the apparent unreliability of the droplet test at $H_2O > 2.5$ % (for example Figure S5g) mean			
		that an alternative method, possibly partial drying, or a higher level of uncertainty may need be applied	< (Deleted: Alternati	ively,
n			Contract 1	· · · · · · · · · · · · · · · · · · ·	

to measurements made at higher water contents.

Deleted:	10	

eleted: This is not the case for samples dried with a Nafion drier. re the CRDS is measuring samples with low water contents, for ample 38 % at BSD and 42 % at HFD of the Nafion dried air mples had $H_2O < 0.1$ %. leted: 10

_			
Dal	-t	•	10
De	eleu		IU.

leted: Alternatively, Deleted: could

3.4 Quantifying the CRDS water correction error using the dew point generator

The change in the CRDS water correction with sample H_2O content was characterised using the difference between the *Wet* and *Dry* DPG runs. This error typically had a shallow negative parabolic trend for both CO₂ and CH₄ (Figure 2) and was similar to the shape seen in the residual of the CRDS water corrections

5 (Figure <u>\$5 and \$6</u>) with the error negative at H₂O mole fractions near 0.5 %, becoming more positive between 1 and 2 % H₂O before dropping at higher H₂O contents.
 Although the UoB CRDS was not deployed in the field we expect the results of the DPG tests to be typical

of most Picarro G2401 CO₂/CH₄ CRDS instrumentation. The DPG tests show that for ambient and below ambient mole fraction samples the CH₄ error remained within the WMO internal reproducibility 10 guidelines (WMO, 2016) at all water contents examined, that is 0.6 to 3.5 % H₂O, while the CO₂ error

- increased outside the guidelines for $H_2O > 2.5$ %. CO_2 errors increased rapidly outside this range reaching 0.3 µmol mol⁻¹ at 3.5 % H₂O. These results are broadly consistent with those of the droplet test residuals. Unlike the ambient and below ambient samples, the CRDS water correction error of the above ambient sample, UoB-04, exceeded the WMO internal reproducibility guidelines for both CO_2 and CH_4 at most
- 15 H₂O mole fractions. For the H₂O range of the BSD and HFD sites the error peaked at 0.1 µmol mol⁻¹ for CO₂ near 1.75 % H₂O and at 2 nmol mol⁻¹ CH₄ near 2.25 % H₂O. As discussed earlier in Section 3.3, the absolute error in the CRDS water correction will scale with the absolute mole fraction of the sample due to the structure of the correction. The UoB CRDS correction was also optimised using a cylinder of significantly lower mole fraction (397.38 µmol mol⁻¹ CO₂ and 1918.73 nmol mol⁻¹ CH₄ compared with
- 20 515.4 μmol mol⁻¹ and 2579.5 nmol mol⁻¹). This shift in error/residual was also observed in the H₂O droplet tests using higher mole fraction cylinders although it appears larger for the DPG tests, most likely due to the higher mole fractions used within these tests (515.4 and 2579.5 compared with 449.55 μmol mol⁻¹ CO₂ and 2148 nmol mol⁻¹ CH₄, respectively).

 The full range H₂O mole fractions observed at the HFD and BSD sites, 0.05 to 2.5 % H₂O, a were not

 25
 examined in these tests which due to limitations inherent in the experimental set up were restricted to a

 H_2O range of 0.6 - 3.5 %. However, it is possible to conclude that for observations of ambient and below ambient CO₂ and CH₄ mole fractions with $H_2O > 0.6$ % the water driven error in the CRDS water correction is not likely to be a major source of uncertainty. Even at other DECC sites that are subject to

1	1
	-
~	•

Deleted: $(Wet - Dry = E_{X\%})$
Deleted: 11
Formatted: Not Highlight
Deleted: S3

Deleted: 7

Deleted: 1

Deleted: . A

(Deleted: Considering the typical
λ	Formatted: Subscript
λ	Formatted: Subscript
1	Formatted: Subscript
(Formatted: Subscript
~~(Deleted: ny
(Deleted: for these sites

higher humidity, for example the Angus site (Stanley et al., 2018) periods of high (> 2.5 % H₂O) water content are rare, < 0.03 % of the data record. In contrast, as elevated CO₂ and CH₄ mole fractions are regularly observed at both the HFD and BSD sites, the increase in CRDS error with mole fraction is a source of concern and must be quantified as part of a full uncertainty analysis.

5 3.5 Quantifying Nafion[®] cross membrane transport

Nafion® membranes, when combined with a dry counter purge gas stream, can be used to effectively dry air samples. This drying process is driven by the moisture gradient between the "wet" sample and the dry counter purge. In a similar manner, as long as the membrane is permeable to the gas, a sample to counter purge gradient in any other trace gas species will also drive exchange. In an effort to quantify the

- 10 magnitude of CO₂ and CH₄ exchange a series of experiments measuring the composition of the Nafion® counter purge gas were conducted. During these experiments all measurement of the Nafion® counter purge (CP_{in} and CP_{out}) were cryogenically dried to < 0.002 % H₂O prior to CRDS analysis. <u>Hence</u> the need to use an empirical CRDS water correction and any error associated with the correction was removed and differences between the CP_{in} and CP_{out} samples can be solely attributed to transport across the
- 15 Nafion membrane (N_{X%}). The results of these experiments are shown in Figure 10. The counter purge experiments conducted with both the ambient (UoB-15) and above ambient (UoB-16) mole fraction cylinders show identical changes in CO₂ and CH₄ mole fractions, respectively. The wet sample N_{X%} difference is consistently positive for CO₂ with the CP_{out} mole fraction an average of 0.021 $\pm 0.002 \ \mu$ mol mol⁻¹ ($\bar{x} \pm 95 \%$ conf. int., n > 19) higher than CP_{in}, reflecting a loss from the sample to the
- 20 counter purge across the Nafion[®] membrane (Figure 10a). Although small, this value is an order of magnitude larger than the average standard deviation of the 15 min block means (0.002 µmol mol⁻¹ CO₂) making it well within the typical measurement precision. This difference decreases slightly with decreasing sample water content but it is never zero. Even with a dry sample, the CP_{out} CP_{in} difference (Nx%), 0.015 ± 0.003 µmol mol⁻¹ CO₂, is still positive. This is in line with previous studies, which have

25

found that, although water substantially increases membrane perme	ability, even dry membranes are
permeable to CO2 (Ma and Skou, 2007; Chiou and Paul, 1988). As	s earlier studies have found that

Deleted:

Deleted: the	
Deleted: hence	
Deleted: it	
Deleted: . As such,	
Deleted: solely	
Deleted: 4	

Deleted:

Deleted: negative

membranes can take more than a week to fully dry out (Chiou and Paul, 1988), it is also highly likely that the relatively brief length of this study (4 to 5 hours) was too short to remove all H₂O from the membrane. The CH₄ CP_{in} and CP_{out} mole fraction difference for both dry and wet samples is also slightly positive, 0.03 ± 0.01 and 0.04 ± 0.02 nmol mol⁻¹ CH₄, respectively (Figure 10c). This value is very close to the

5 measurement precision, with the average CH₄ standard deviation of the 15-min block means of the order of 0.02 nmol mol⁻¹ CH₄.

The ~ 0.02 μ mol mol⁻¹ loss of CO₂ across the Nafion[®] membrane from the sample stream to the counter purge observed here, although small, is of the order of the WMO internal reproducibility <u>guidelines</u>, 0.05 μ mol mol⁻¹ in the northern hemisphere and 0.025 μ mol mol⁻¹ in the southern hemisphere (WMO, 2016),

- 10 and must be acknowledged. However, the calibration gases are also passed through the Nafion. These cylinders are very dry, $H_2O < 0.0001$ %, equivalent to the driest conditions studied in the DPG experiments (Figure 10a and b) and as such would be expected to show similar CO₂ loss across the Nafion. membrane, ~ 0.015 µmol mol⁻¹. Hence, as the bias is constant with sample CO₂ and H₂O mole fractions and as a bias would be present in both the calibration gases (~ 0.015 µmol mol⁻¹) and samples
- 15 (~ 0.02 µmol mol⁻¹) the majority of the bias will be calibrated out, with only a very small (≤ 0.005 µmol mol⁻¹) constant bias, of the order of the instrumental precision, remaining.
 In contrast, the mean CH₄ Nafion bias, 0.04 ± 0.02 nmol mol⁻¹, is at least an order of magnitude smaller than the WMO internal reproducibility guidelines (WMO, 2016) and is extremely close to the typical measurement precision suggesting that it is not a bias of concern.

20 4 Conclusions and future work

The newly established Bilsdale and Heathfield tall tower measurement stations provide important new data sets of GHG observations. These high-precision continuous in situ measurements show clear long term increases in baseline CO₂, CH₄, N₂O and SF₆ mole fraction and capture the seasonal and diurnal cycles of these key gases. <u>It is expected</u> that these observations, when combined with regional inversion

25 modelling, will significantly improve our ability to quantify UK greenhouse gas emissions — both reducing the uncertainty and improving the spatial and temporal resolution. Future work using this data

Deleted: t
Deleted: requirements
Palatad: (WMO 2016)
Deleted. (wM0, 2010)
Deleted:
Deleted: the 0 °C data points
Deleted: used in
Deleted: Experiment 3

Deleted: However, unlike the CO2 difference,

Deleted: This is not the case for the water correction bias, which varies with H₂O (Section 3.3.2).

Deleted: (WMO, 2016)

Deleted:

Deleted: 4

Deleted: 4

Considering how poorly the results from Experiment 1 and 2 characterise the N_{xNs} value it is interesting to see that the CO₂ N_{xNs} value obtained here is within the bounds of those estimates while the CH₄ estimates are not. It is not clear whether this is a fortuitous quirk or because the CRDS water correction error is two-fold larger for CH₄ than CO₂ relative to the typical atmospheric mole fractions

Deleted: As demonstrated for CH4, w Deleted: e

is focusing on better estimates of UK GHG emissions with a particular emphasis on the UK carbon budget.

An examination of the Nafion R drying method found it to have a small inherent CO₂ bias of 0.02 µmol mol⁻¹; however, this bias did not vary significantly with sample water content > 0.7 % H₂O or CO₂ mole

5 fraction. Even samples as dry as the calibration gases were affected by this Nafion[®] bias, although to a smaller degree — ~ 0.015 µmol mol⁻¹ for H₂O < 0.0001 % — as residual moisture remained in the membrane. Thus, as calibration gases are dried in an identical manner to the samples, this bias is mostly calibrated out with only a very small (≤ 0.005 µmol mol⁻¹) constant residual bias of the order of the instrumental precision. As such, the Nafion[®] drier itself when, used in this manner, does not contribute 10 a significant bias to the resulting CO₂ observations.

In contrast, the errors associated with the CRDS water correction for samples with low water contents (< 0.5 %), like those dried using a Nafion[®] drier, can be significant and difficult to adequately quantify using the current in field techniques. Hence, even though Nafion[®] driers are not themselves an inherent source of bias, for the CRDS instrumentation examined in this study the incomplete drying of the sample

- 15 is a significant source of error. Altering the drying method to better match the moisture content of the calibration gases to the sample may minimise this error. It is also important to note that this may not be the case for other CRDS instrumentation or optical techniques that use alternative cell pressure sensors. Similarly, samples with high water contents > 2.5 % H₂O or CO₂ or CH₄ mole fractions significantly above ambient mole fractions are also subject to larger water correction errors.
- 20 Estimates of these types of errors for Bilsdale and Heathfield have been given in Table 5 and range between 0.2 and 0.05 μmol mol⁻¹ CO₂ and 3 and 1 nmol mol⁻¹ CH₄. While these errors are significant relative to the WMO internal reproducibility goals they are typically smaller than the extended WMO measurement compatibility goals (± 0.2 μmol mol⁻¹ CO₂ and ± 5 nmol mol⁻¹ CH₄). It is also important to note that they are orders of magnitude smaller than baseline excursions observed at the sites (see Figures)
- 25 <u>4 & 5). They are also a factor of 10 smaller than the CH_d model-data mismatch within the UK DECC network as estimated by Ganesan et al. (2015) at ~ 20 nmol mol⁻¹.
 While, drift in the instrumental water correction typically small it is important that it is identified and accounted for through regular water tests. The necessary frequency of these water tests will depend on
 </u>

Deleted: 8

Deleted: This

Formatted: Subscript	
Formatted: Subscript	
Formatted: Subscript	

Formatted: Not Superscript/ Subscript

Formatted: Superscript	
Formatted: Subscript	
Formatted: Subscript	
Formatted: Superscript	

Deleted:

Deleted: Undried

Deleted: air samples with water contents typically between 0.5 and $2.5 \ \% H_2O$ are not subject to this error. However, care must be taken to ensure that any

the stability of the individual instrument and the typical CO₂, CH₄ and H₂O mole fractions at the given location and should be determined on a case by case basis.

Future improvements to the Bilsdale and Heathfield records include the addition of target tanks at the sites. Although the use of target tanks do not directly influence measurement uncertainty they allow

5 independent long-term monitoring of instrument performance and are a useful tool for assessing measurement uncertainty. The development of a full uncertainty analysis incorporating such target tank measurements, along with an assessment of the calibration strategy and any instrumental, water correction and sampling errors is also planned.

Deleted: Future improvements to the Bilsdale and Heathfield records include the addition of target tanks at the sites to allow independent long-term monitoring of instrument performance and the development of a full uncertainty analysis incorporating calibration, instrumental, water correction and sampling errors.

5 Acknowledgements

- 10 This study was funded under the NERC Greenhouse Gas Emissions and Feedbacks Program as part of the Greenhouse gAs UK and Global Emissions (GAUGE) area grant number NE/K002449/1NERC. This grant also covered the establishment and early running costs of the stations. Operating costs of the Bilsdale site after 17th September 2016 were funded by the UK Department of Business, Energy and Industrial Strategy (formerly the Department of Energy and Climate Change) through contract TRN1028/06/2015.
- 15 The National Physical Laboratory (NPL) took responsibility for the Heathfield site on the 30th September 2017 by the UK's Department for Business, Energy and Industrial Strategy as part of the National Measurement System Programme.

The authors would also like to acknowledge the support of Dr Carole Helfter and Dr Neil Mullinger from the NERC Centre for Ecology and Hydrology (CEH), Edinburgh, Scotland who helped to maintain and

20 run the Bilsdale site. Lastly the authors would like to thank Dr Joseph Pitt from the University of Manchester for the use of the dew point generator used during a series of preliminary studies.

29

Deleted: 5.

7 References

Bakwin, P. S., Tans, P. P., Hurst, D. F., and Zhao, C.: Measurements of carbon dioxide on very tall towers: results of the NOAA/CMDL program, Tellus B, 50, 401-415, doi:10.1034/j.1600-0889.1998.t01-4-00001.x, 1998.

- 5 Bouwman, A. F., and Taylor, J. A.: Testing high-resolution nitrous oxide emission estimates against observations using an atmospheric transport model, Global Biogeochemical Cycles, 10, 307-318, 10.1029/96gb00191, 1996.
- Chen, H., Winderlich, J., Gerbig, C., Hoefer, 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 (CO₂ and CH₄) using the cavity ring-down spectroscopy (CRDS) technique, Atmos. Meas. Tech., 3, 375-386, 10.5194/amt-3-375-2010,
- 10 2010. Chen, H., Karion, A., Rella, C. W., Winderlich, J., Gerbig, C., Filges, A., Newberger, T., Sweeney, C., and Tans, P. P.: Accurate
- measurements of carbon monoxide in humid air using the cavity ring-down spectroscopy (CRDS) technique, Atmos. Meas. Tech., 6, 1031-1040, 10.5194/amt-6-1031-2013, 2013.
- Chiou, J. S., and Paul, D. R.: Gas permeation in a dry Nafion membrane, Industrial & Engineering Chemistry Research, 27, 2161-2164, 10.1021/ie00083a034, 1988.
 - Chris Blandford Associates: North Yorkshire and York Landscape Characterisation Project, London, 2011.
- Dlugokencky, E. J., Myers, R. C., Lang, P. M., Masarie, K. A., Crotwell, A. M., Thoning, K. W., Hall, B. D., Elkins, J. W., and Steele, L. P.: Conversion of NOAA atmospheric dry air CH4 mole fractions to a gravimetrically prepared standard scale, Journal of Geophysical Research: Atmospheres, 110, n/a-n/a, 10.1029/2005JD006035, 2005.
- 20 Flores, E., Viallon, J., Moussay, P., Griffith, D. W. T., and Wielgosz, R. I.: Calibration Strategies for FT-IR and Other Isotope Ratio Infrared Spectrometer Instruments for Accurate δ13C and δ18O Measurements of CO2 in Air, Analytical Chemistry, 89, 3648-3655, 10.1021/acs.analchem.6b05063, 2017.

Foulger, B. E., and Simmonds, P. G.: Drier for field use in the determination of trace atmospheric gases, Analytical Chemistry, 51, 1089-1090, 10.1021/ac50043a074, 1979.

- 25 Fraser, P., Porter, L. W., Baly, S. B., Krummel, P. B., Dunse, B. L., Steele, L. P., Derek, N., Langenfelds, R. L., Levin, I., Oram, D. E., Elkins, J. W., Vollmer, M. K., and Weiss, R. F.: Sulfer hexafluoride at Cape Grim: Long term trends and regional emissions, 18-23, 2004. Ganesan, A. L., Manning, A. J., Grant, A., Young, D., Oram, D. E., Sturges, W. T., Moncrieff, J. B., and O'Doherty, S.: Quantifying methane and nitrous oxide emissions from the UK and Ireland using a national-scale monitoring network, Atmos. Chem. Phys., 15, 6393-6406, 10.5194/acp-15-6393-2015, 2015.
- 30 Griffith, D. W. T.: Calibration of isotopologue-specific optical trace gas analysers: A practical guide, Atmos. Meas. Tech. Discuss., 2018, 1-19, 10.5194/amt-2018-187, 2018.
- Hall, B. D., Dutton, G. S., Mondeel, D. J., Nance, J. D., Rigby, M., Butler, J. H., Moore, F. L., Hurst, D. F., and Elkins, J. W.: Improving measurements of SF6 for the study of atmospheric transport and emissions, Atmos. Meas. Tech., 4, 2441-2451, 10.5194/amt-4-2441-2011, 2011.
- 35 A gas chromatographic system for high-precision quasi-continuous atmospheric measurements of CO2, CH4, N2O, SF6, CO and H2: https://www.researchgate.net/publication/255704060 A gas chromatographic system for high-precision quasicontinuous atmospheric measurements of CO2 CH4 N2O SF6 CO and H2, 2008. Hunter, J. D.: Matplotlib: A 2D Graphics Environment, Computing in Science & Engineering, 9, 90-95, 10.1109/MCSE.2007.55, 2007.

IPCC: Summary for Policymakers, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth

40 Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1–30, 2013.

Ma, S., and Skou, E.: CO₂ permeability in Nafion® EW1100 at elevated temperature, Solid State Ionics, 178, 615-619, dx.doi.org/10.1016/j.ssi.2007.01.030, 2007.

45 Manning, A. J., O'Doherty, S., Jones, A. R., Simmonds, P. G., and Derwent, R. G.: Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modeling approach, Journal of Geophysical Research: Atmospheres, 116, n/a-n/a, 10.1029/2010JD014763, 2011.

Nara, H., Tanimoto, H., Tohjima, Y., Mukai, H., Nojiri, Y., Katsumata, K., and Rella, C. W.: Effect of air composition (N₂, O₂, Ar, and H₂O) on CO₂ and CH₄ measurement by wavelength-scanned cavity ring-down spectroscopy: calibration and measurement strategy, Atmos. 50 Meas. Tech., 5, 2689-2701, 10.5194/amt-5-2689-2012, 2012.

Naudy, S., Collette, F., Thominette, F., Gebel, G., and Espuche, E.: Influence of hygrothermal aging on the gas and water transport properties of Nafion® membranes, Journal of Membrane Science, 451, 293-304, dx.doi.org/10.1016/j.memsci.2013.10.013, 2014.

30

Nevison, C. D., Dlugokencky, E., Dutton, G., Elkins, J. W., Fraser, P., Hall, B., Krummel, P. B., Langenfelds, R. L., O'Doherty, S., Prinn, R. G., Steele, L. P., and Weiss, R. F.: Exploring causes of interannual variability in the seasonal cycles of tropospheric nitrous oxide, Atmos. Chem. Phys., 11, 3713-3730, 10.5194/acp-11-3713-2011, 2011.

North York Moors National Park Authority: North York Moors National Park Management Plan: A wider view, Helmsley, 2012.

- 5 Novelli, P. C., Elkins, J. W., and Steele, L. P.: The development and evaluation of a gravimetric reference scale for measurements of atmospheric carbon monoxide, Journal of Geophysical Research: Atmospheres, 96, 13109-13121, doi:10.1029/91JD01108, 1991. Prinn, R., Cunnold, D., Rasmussen, R., Simmonds, P., Alyea, F., Crawford, A., Fraser, P., and Rosen, R.: Atmospheric emissions and trends
- of nitrous oxide deduced from 10 years of ALE-GAGE data, Journal of Geophysical Research: Atmospheres, 95, 18369-18385, 10.1029/JD095iD11p18369, 1990.
- 10 Rella, C.: Accurate Greenhouse Gas Measurements in Humid Gas Streams Using the Picarro G1301 Carbon Dioxide / Methane / Water Vapor Gas Analyzer 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, Atmospheric Measurement Techniques, 6, 837-860, 10.5194/amt-6-837-2013, 2013.

- 15 Reum, F., Gerbig, C., Lavric, J. V., Rella, C. W., and Göckede, M.: Correcting atmospheric CO₂ and CH₄ mole fractions obtained with Picarro analyzers for sensitivity of cavity pressure to water vapor, Atmos. Meas. Tech. Discuss., 2018, 1-29, 10.5194/amt-2018-242, 2018. Schibig, M. F., Steinbacher, M., Buchmann, B., van der Laan-Luijkx, I. T., van der Laan, S., Ranjan, S., and Leuenberger, M. C.: Comparison of continuous in situ CO2 observations at Jungfraujoch using two different measurement techniques, Atmos. Meas. Tech., 8, 57-68, 10.5194/amt-8-57-2015, 2015.
- 20 Stanley, K. M., Grant, A., O'Doherty, S., Young, D., Manning, A. J., Stavert, A. R., Spain, T. G., Salameh, P. K., Harth, C. M., Simmonds, P. G., Sturges, W. T., Oram, D. E., and Derwent, R. G.: Greenhouse gas measurements from a UK network of tall towers: technical description and first results, Atmos. Meas. Tech., 11, 1437-1458, 10.5194/amt-11-1437-2018, 2018. Tans. P. P. Crotwell, A. M., and Thoning, K. W.: Abundances of isotopologues and calibration of CO2 greenhouse gas measurements.

Atmos. Meas. Tech., 10, 2669-2685, 10.5194/amt-10-2669-2017, 2017.

25 Turnbull, J. C., Sweeney, C., Karion, A., Newberger, T., Lehman, S. J., Tans, P. P., Davis, K. J., Lauvaux, T., Miles, N. L., Richardson, S. J., Cambaliza, M. O., Shepson, P. B., Gurney, K., Patarasuk, R., and Razlivanov, L: Toward quantification and source sector identification of fossil fuel CO₂ emissions from an urban area: Results from the INFLUX experiment, Journal of Geophysical Research: Atmospheres, 120, 292-312, 10.1002/2014JD022555, 2015.

van der Laan, S., Neubert, R. E. M., and Meijer, H. A. J.: A single gas chromatograph for accurate atmospheric mixing ratio measurements
 of CO2, CH4, N2O, SF6 and CO, Atmos. Meas. Tech., 2, 549-559, 10.5194/amt-2-549-2009, 2009.
 van Rossum, G.: Python tutorial, Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, 1995.
 Walt, S. v. d., Colbert, S. C., and Varoquaux, G.: The NumPy Array: A Structure for Efficient Numerical Computation, Computing in Science & Engineering, 13, 22-30, 10.1109/MCSE.2011.37, 2011.

Welp, L. R., Keeling, R. F., Weiss, R. F., Paplawsky, W., and Heckman, S.: Design and performance of a Nafion dryer for continuous operation at CO₂ and CH₄ air monitoring sites, Atmos. Meas. Tech., 6, 1217-1226, 10.5194/amt-6-1217-2013, 2013.

Winderlich, J., Chen, H., Gerbig, C., Seifert, T., Kolle, O., Lavrič, J. V., Kaiser, C., Höfer, A., and Heimann, M.: Continuous lowmaintenance CO₂CH₄/H₂O measurements at the Zotino Tall Tower Observatory (ZOTTO) in Central Siberia, Atmos. Meas. Tech., 3, 1113-1128, 10.5194/amt-3-1113-2010, 2010.

WMO: 18th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques (GGMT-2015), Geneva, Switzerland, 2016.

- Young, D.: Development of a Pre-Concentration System for the Determination of Atmospheric Oxygenated Volatile Organic Compounds, Doctor of Philosphy, School of Chemistry, Faculty of Science, University of Bristol, Bristol, 2007. Yver Kwok, C., Laurent, O., Guemri, A., Philippon, C., Wastine, B., Rella, C. W., Vuillemin, C., Truong, F., Delmotte, M., Kazan, V., Darding, M., Lebègue, B., Kaiser, C., Xueref-Rémy, I., and Ramonet, M.: Comprehensive laboratory and field testing of cavity ring-down
- 45 spectroscopy analyzers measuring H₂O, CO₂, CH₄ and CO, Atmos. Meas. Tech., 8, 3867-3892, 10.5194/amt-8-3867-2015, 2015. Zhao, C. L., and Tans, P. P.: Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air, Journal of Geophysical Research-Atmospheres, 111, D08s09 10.1029/2005/d006003, 2006.

50





Figure 1: Locations of the GAUGE Bilsdale (BSD) and Heathfield (HFD) sites, shown in black and the UK DECC Mace Head (MHD), Ridge Hill (RGL), Tacolneston (TAC) and Angus (TTA) sites, shown in grey.



Figure 2: A generalised schematic showing the initial Bilsdale and Heathfield site setup of the cavity ringdown spectrometer (CRDS) and the Gas Chromatograph – Electron Capture Dectector (GC-ECD) including the dry gas generator (TOC) and back pressure regulator (BP). Note that Bilsdale has three inlets, while Heathfield has only two as shown here. The Nafion[®] drying system located downstream of the <u>CRDS</u> multiport valve was removed at both sites in 2015. Black arrows and lines show the direction of sample, standard and calibration gas flow. Grey dashed lines and arrows show the flow path of the Nafion[®] counter purge gas.

5

Deleted: CRDS

Formatted: Normal



Figure 3: A schematic of the humidification system used in the Nafion® counter purge experiment. Here the TOC is the dry gas generator. The Black arrows and lines show the direction of sample gas flow. Grey dashed lines and arrows show the flow path of the Nafion® counter purge gas. Heated zones are shown in yellow.

Formatted: Normal



Deleted:





Deleted:



Formatted: Caption

Figure 6: Box and whisker plots (25^{th} to 75^{th} percentiles) of the the diurnal cycle by season of detrended hourly mean values for (a) & (b) CO₂, (c) & (d) CH₄, (e) & (f) CO, (g) & (h) N₂O and (i) & (j) SF₆ of the Bilsdale 42 m (blue), 108 m (green) & 248 m (purple) and Heathfield 50 m (red) & 100 m (yellow) intake heights.

5











 Figure 9: The (a) CO₂ and (b) CH₄ change in the Wet – Dry sample treatment difference with sample water content for cylinders

 UoB-04 (515.3 µmol mol⁻¹ CO₂ and 2585 nmol mol⁻¹ CH₄), H-296 (406.6 µmol mol⁻¹ CO₂ and 1947 nmol mol⁻¹ CH₄), UoB-06 (384.8

 ymol mol⁻¹ CO₂ and 1975 nmol mol⁻¹ CH₄) and H-306 (372.5 µmol mol⁻¹ CO₂ and 1776 nmol mol⁻¹ CH₄). Error bars are the larger of either the standard deviation of the mean difference or the uncertainties of the two sample types added together in quadrature.

(Formatted: Subscript
\sim	Formatted: Superscript
X	Formatted: Subscript

Formatted: Normal



1

Figure 10: Change in the counter purge in (CP_{in}) and out (CP_{out}) (a) CO₂ and (c) CH₄ mole fraction with sample water content for ambient (UoB-15) and above ambient (UoB-16) mole fraction cylinders. Change in (b) CP_{in} and (d) CP_{out} water content with changing sample water content. Note that the gas stream was cryogenically drived before analysis. Error bars are larger of either the standard deviation of the mean difference or the uncertainties of the two sample types added together in quadrature. The dotted lines in (a) and (c) are the respective WMO internal reproducibility guidelines.

Deleted: 14

Table 1 – Instrument specific water corrections for the Bilsdale (BSD), Heathfield (HFD) and University of Bristol (UoB) CRDS instruments. The parameters shown are the mean \pm the 95% confidence interval of tests repeated in triplicate. Water corrections labelled High and Low were determined using an above ambient and below ambient mole fraction cylinder, respectively, while the rest were determined using an ambient mole fraction cylinder. The mean residual along with the interquartile range of the residuals are included.

Deleted: is the mean of the absolute values of all residuals $(1\sigma^2)$

				A		В			Mean residual (25 th -75 th quartile) CO ₂ µmol mol ⁻¹ CH ₄ nmol mol ⁻¹	n
		2015	-0.0157	±	0.0001	0.00018	±	0.00008	0.0003 (-0.01 - 0.01)	4
	BSD	2016	-0.01578	±	0.00004	0.00022	±	0.00002	-0.002 (-0.01 - 0.01)	3
		2017	-0.01556	±	0.00005	0.00008	±	0.00002	-0.001 (-0.01 - 0.02)	5
60	HFD	2015	-0.01558	±	0.00008	0.00010	±	0.00004	-0.002 (-0.02 - 0.02)	3
02		2016	-0.0154	±	0.0001	0.00004	±	0.00003	0.004 (-0.003 - 0.02)	1*
		2015	-0.0156	±	0.0003	0.0001	±	0.0001	-0.00002 (-0.03 - 0.03)	3
		2016	-0.01577	±	0.00007	0.00020	±	0.00004	-0.02 (-0.03 - 0.02)	13
	UoB	2017	-0.01558	±	0.00008	0.00012	±	0.00004	-0.006 (-0.02 - 0.007)	3
		2016 High	-0.0160	±	0.0003	0.0003	±	0.0001	0.007 (0.007 - 0.05)	3
		2016 Low	-0.01606	±	0.00005	0.00030	±	0.00002	-0.02 (-0.02 - 0.02)	3
		2015	-0.0138	±	0.0002	0.0005	±	0.0001	-0.02 (-0.2 - 0.1)	4
	BSD	2016	-0.0139	±	0.0002	0.0006	±	0.0001	-0.04 (-0.2 - 0.1)	3
		2017	-0.01309	±	0.00009	0.00014	±	0.00002	-0.04 (-0.2 - 0.1)	5
	HFD	2015	-0.01273	±	0.00004	0.00013	±	0.00004	-0.03 (-0.2 - 0.2)	3
CH ₄		2016	-0.0119	±	0.0005	-0.0002	±	0.0002	-0.09 (-0.4 - 0.3)	1*
		2015	-0.0137	±	0.0003	0.0002	±	0.0001	-0.06 (-0.2 - 0.1)	3
		2016	-0.0139	±	0.0001	0.00025	±	0.00005	0.002 (-0.2 - 0.2)	13
	UoB	2017	-0.0139	±	0.0001	0.00027	±	0.00006	-0.04 (-0.2 - 0.1)	3
		2016 High	-0.01393	±	0.00005	0.0004	±	0.0001	0.1 (-0.02 - 0.3)	3
		2016 Low	-0.01402	±	0.00005	0.00028	±	0.00008	-0.02 (-0.1 - 0.1)	3

*The fitted parameter and $1\sigma^2$ of a single test due to a leak in the septum

42

Table 2 – The cylinders used during the dew point generator CRDS water correction, Nafion® counter purge and UoB instrument specific water tests. Most measurements were made in-house and only corrected for linear drift against a standard calibrated at WCC-EMPA, Dübendorf, Switzerland and hence are simply indicative of the expected mole fractions. While those marked * were calibrated at GasLab MPI-BGC, Jena, Germany and linked to the WMO x2007 CO_2 and x2004A CH₄ scales.

The set former	Galladaa	CO ₂	CH ₄	
I est type	Cylinder	µmol mol-1	nmol mol ⁻¹	
	H-306	372.5	1776	
Dew point generator CRDS water correction	UoB-06	384.8	1975	
	H-296	406.6	1947	
	UoB-04	515.3	2585	
N. C	UoB-15	399.3	1928	
Nation® counter purge	UoB-16	430.7	2015	
	USN20104095*	346.91 ± 0.06	1742.9 ± 0.3	
UoB instrument specific water correction	H-283	379.1	1815	
	USN20104068*	449.49 ± 0.05	2145.0 ± 0.4	

Table 3 – CRDS calibration and standard cylinder mole fractions and usage start dates for the Heathfield (HFD) and Bilsdale (BDS) sites. Where available the mole fractions measured prior to and after deployment are given. Reported mole fractions from the WCC-EMPA, Dübendorf, Switzerland are given as mean ± uncertainty. *Mole fraction measurement from GasLab MPI-BGC, Jena, Germany are given as mean ± 1 σ .

Cylinder		CO2 WMO x2007 µmol mol ⁻¹		CH₄ WMO x2 nmol m	004A ol ⁻¹	CO WMO x2014 nmol mol ⁻¹		Start date – End date	
			Prior	Post	Prior	Post	Prior	Post	
		Low	-	379.2 ± 0.2	-	1807 ± 3	-	124 ± 2	2014 - 30 - 1 - 2015 - 04 - 24
	Calibration Suite #1	Ambient	-	394.7 ± 0.2	-	1889 ± 4	-	131 ±2	2014-02-20 - 2015-11-07
		High	-	456.5 ± 0.2	-	2074 ± 4	-	583 ± 6	2014-01-30 - 2015-04-24
	Calibration	Low	379.51 ± 0.06	-	1812.5 ± 0.02	-	74.6 ± 0.3	-	2016-01-20 - Current
BSD	Suite #2*	Ambient	418.63 ± 0.06	-	2090.0 ± 0.03	-	246.1 ± 0.4	-	2015-10-02 - Current
		High	471.17 ± 0.06	-	2404.8 ± 0.04	-	469.2 ± 0.5	-	2015-10-02 - Current
	Standard	H-239	-	395.2 ± 0.2	-	1900 ± 4	-	118 ± 3	2014-01-30 - 2014-09-23
		H-252	402.3 ± 0.2	402.3 ± 0.2	1906 ± 2	1906 ± 4	138 ± 2	144 ± 3	2014-09-23 - 2015-07-22
		H-251	402.2 ± 0.2	402.3 ± 0.2	1906 ± 2	1906 ± 4	138 ± 2	145 ± 3	2015-07-22 - 2016-05-06
		USN- 20141394*	399.31 ± 0.05	-	1939.3 ± 0.02	-	123.7 ± 0.3	-	2016-05-06 - Current
	Calibration Suite*	Low	369.24 ± 0.06	-	1845.9 ± 0.3	-	128.8 ± 0.3	-	2013-12-16 - Current
		Ambient	420.24 ± 0.06	-	1993.8 ± 0.3	-	321.7 ± 0.5	-	2013-12-16 - Current
		High1	441.26 ± 0.06	-	2211.0 ± 0.4	-	224.23 ± 0.4	-	2013-12-16 - 2017-01-27
HFD		High2	477.59 ± 0.06	-	2282.1 ± 0.4	-	104.65 ± 0.3	-	2017-02-24 – Current
	Standard	H-240	-	394.3 ± 0.2	-	1882 ± 4	-	121 ± 3	2013-12-16 - 2014-12-17
		H-255	402.1 ± 0.2	402.1 ± 0.2	1908 ± 2	1908 ± 4	135 ± 2	141 ± 3	2014-12-17 - 2015-10-21
		H-254	402.1 ± 0.2	402.2 ± 0.2	1908 ± 2	1908 ± 4	135 ± 2	142 ± 3	2015-10-21 - 2016-09-21
		H-285	393.6 ± 0.2	-	1928 ± 4	-	105 ± 2	-	2016-09-21 - Current

Table 4 – GC-ECD standard cylinder mole fractions and usage start dates

		N ₂ O	SF ₆		
Site	Cylinder	SIO-16	SIO-SF6	Start date	
		nmol mol ⁻¹	pmol mol ⁻¹		
HFD	H-234	326.67	8.20	14/11/2013	5
DDI	H-235	326.56	8.13	14/1/2014	
BDL	H-222	326.23	8.05	2/10/2015	

Site	Time period	Drying method	Water mole fraction at CRDS	<i>Maximum</i> CO ₂ error μmol mol ⁻¹	<i>Maximum</i> <i>CH</i> ₄ error nmol mol ⁻¹
	2014-01-01 - 2015-06-17	Nafion drying with instrument specific water correction	0.05 - 0.2 %	- 0.1	- 2
Bilsdale	2015-06-18 - 2016-10-13	Instrument specific	0-0.2 %	- 0.1	- 1.5
(BSD)		water correction	> 0.2 %	± 0.05	± 1
	2016-10-14 - Current	Instrument specific	0 - 1 %	- 0.2	- 2
		water correction	1 – 3.5 %	± 0.05	± 1
	2013-12-01 - 2015-09-30	Nafion drying with instrument specific water correction	0.05 - 0.2 %	- 0.1	- 2
Heathfield	2015-10-01 - 2016-08-23	Instrument specific	0-0.2 %	- 0.1	- 1
(HFD)		water correction	> 0.2 %	± 0.05	± 1
	2016-08-23 – Current	Instrument specific	0 – 1.7 %	- 0.1	- 3
		water correction	> 1.7 %	± 0.05	± 1

Table 5 – Estimates of the maximum error associated with the measurement of ambient CO_2 and CH_4 mole fraction samples using the given drying and/or water correction method for the BSD and HFD sites

Supplementary

S1: Assessing the CRDS water correction

The CRDS water correction was assessed through a series of simple Dew Point Generator (DPG; Licor LI-610 Portable Dew Point Generator, USA) experiments. Here four cylinders of air with varying CO₂

- 5 and CH₄ mole fractions (Table 4) were humidified to a range of set dew points between 2.5 and 30 °C (0.6 to 3.5 % H₂O) and measured, with and without cryogenic drying, using a Picarro G2301 CRDS at the University of Bristol (UoB) laboratory. Cylinder delivery pressure was controlled using single stage high purity stainless steel Parker Veriflow regulators (95930S4PV3304, Parker Balston, USA) or TESCOM regulators (64-2640KA411, Tescom Europe).
- 10 In brief, the output of the cylinder regulator was plumbed to the input of the DPG. A T-piece connected prior to the DPG input vented any excess gas via a flow meter (F1, Fig. S2) ensuring that the DPG input remained at close to ambient atmospheric pressure throughout the experiment. The output of the DPG passed through a second T-piece with the over flow outlet also connected to a flow meter (F2) to ensure that the CRDS input pressure remained near ambient. A third flowmeter (F3) was placed on the outflow
- 15 of the Nafion counter purge. Flow meters F1 and F2 had a range of 0.1–1 L min⁻¹ (VAF-G1-05M-1, Swagelok, UK) while F3 had a smaller flow range 0.5–0.5 L min⁻¹ (FR2A12BVBN-CP, Cole-Palmer, USA). Typical output flows were 0.1, 0.3 and 0.3 L min⁻¹ for F1, F2 and F3 respectively. After F2 the sample flow was further split using a T-piece, with half the flow passing through a cryogenic water trap before reaching a 4-port 2-position valve, V1 (EUDA-2C6UWEPH, VICI Valco AG International,
- 20 Switzerland, actually a 6-port valve configured as a 4-port valve). The other half bypassed the water trap and connected directly to V1. One of the outputs of V1 went via the Nafion to a second identical valve, V2, while the second output went directly to V2. The first output of V2 connected directly to the input of the CRDS while the second connected to a pump (PICARRO Vacuum pump S/N PB2K966-A) set to a flow rate matching that of the CRDS (0.3 L/min) to ensure uniform flow through both branches of the
- 25 system. These valves were controlled manually using a VALCO electronic controller and universal actuator.

The water trap consisted of a coil of $\frac{1}{4}$ " diameter (I.D. 3.36mm) stainless steel tubing immersed in a Dewar of silicone oil (Thermo Haake SIL 100, Thermo Fisher Scientific, USA). The silicone oil was cooled using an immersion probe (CC-65, NESLAB) to less than -50 °C. Other than the water trap and two short sections (< 10 cm) of $\frac{1}{4}$ " (O.D.) plastic tubing immediately prior to and post the DPG, $\frac{1}{16}$ "

- 5 stainless steel tubing was used throughout the system.
- The experiment was conducted in a temperature-controlled laboratory at 19 °C, and thus, at temperatures lower than a number of the dew points used within the experiment. Hence, in order to avoid condensation forming on the walls of the tubing all components of the system between the cylinder, excluding the water trap, and the outputs of V2 were contained within a chamber heated to > 32 °C. Tubing between the
- heated chamber and the input of the CRDS was also heated with heating tape to > 32 °C while the internal temperature of the CRDS was > 32 °C throughout the experiment.
 Multiple measurement blocks of each sample treatment were conducted after a lengthy initial stabilisation period. This period allowed for the establishment of equilibrium between the water in the condenser block of the DPG and the sample gas and lasted at least 2 hours (sometimes up to 5 hours). The treatment blocks
- 15 varied in length depending on the time required for the concentration to stabilise. At least 15 minutes of data was collected after the concentration stabilised.

It is important to note that the DPG was not calibrated but the H₂O concentration was measured directly by the CRDS during the undried experiments. These values were used as the reference H₂O concentration in all calculations and Figure 11.

20 Flow rates, cylinder pressure, chamber temperature and H2O trap temperature were manually logged after each valve position change and when the water trap was inserted into the silicone oil bath.



Figure S1: (a) CO₂, (b) CH₄, (c) CO and (d) H₂O mole fractions of the TOC output. Note the HFD TOC was powered up just before to analysis in contrast to the other sites where the TOCs had been running for at least 12 hours prior to analysis.



Figure S2: Droplet test instrument specific water correction flow path and procedure



5 Figure S3: A schematic of the humidification system used to assess the CRDS water correction. Here the TOC is the dry gas generator. The Black arrows and lines show the direction of sample gas flow. Grey dashed lines and arrows show the flow path of the Nafion® counter purge gas. Heated zones are shown in yellow.


Figure S4: The (a) CO₂, (b), CH₄ and (c) H₂O minute mean data obtained during the Nafion[®] counter purge experiments for cylinder UoB-16 at a dewpoint of 10C. Error bars are ± 1 standard deviation of each minute mean. Purple and grey data points are the sample without and with the H₂O trap, respectively, while blue and red data points are the Nafion[®] counter purge before and after the Nafion[®], respectively.



Figure S5: Droplet test residual, wet corrected data less dry mean, plots for Bilsdale (a) 2015, (b) 2016 and (c) 2017, Heathfield (d) 2015 and (e) 2016 and University of Bristol (f) 2015, (g) 2016 and (h) 2017. The instrument specific CO₂ residual values (red) are shown in the upper plots, CH₄ (blue) in the middle plots. And CO (orange) in the lower plots. The residuals of the factory determined water correction are also shown in grey. The mean $\pm 1\sigma^2$ of the residuals are given for each plot for both the instrument specific

5 (black) and the factory (grey) corrections.

10



Figure S6: Droplet test residual, wet corrected data less dry mean, plots for (a) below ambient cylinder and (b) the above ambient cylinder at the University of Bristol. The instrument specific CO₂ residual values (red) are shown in the upper plots and CH₄ (blue) in the middle plots. The residuals of the factory determined water correction are also shown in grey. The mean $\pm 1\sigma^2$ of the residuals are given for each plot for both the instrument specific (black) and the factory (grey) corrections.