Reply to the comments of Reviewer #2 on the manuscript "Greenhouse gas measurements from a UK network of tall towers: technical description and first results"

Kieran M. Stanley[1], Aoife Grant[1], Simon O'Doherty[1], Dickon Young[1], Alistair J. Manning[1,2], Ann R. Stavert[1], T. Gerard Spain[3], Peter K. Salameh[4], Christina M. Harth[4], Peter G. Simmonds[1], William T. Sturges[5], David E. Oram[5], Richard G. Derwent[6]

[1]School of Chemistry, University of Bristol, Bristol, United Kingdom
[2]Met Office, Exeter, Devon, United Kingdom
[3]Department of Experimental Physics, National University of Ireland, Galway, Ireland
[4]Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA
[5]School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom
[6]rdscientific, Newbury, Berkshire, United Kingdom

We thank the reviewer #2 for their time and effort in evaluating this manuscript and for their suggestions for improvements. All points made by the reviewer are addressed on the following pages.

The paper is describing the setup for greenhouse gases measurements at four sites in UK. The authors are providing a very detailed description of the inlet parts, analyzers, and calibration protocols. In the last section of the paper atmospheric signals (diurnal and seasonal cycles, trend) are very briefly discussed. Considering the purpose of the manuscript I am not fully convinced by the need for this very general discussion on the observed variabilities, coming after very detailed technical descriptions of the setup and protocols. From my point of view, the most problematic point of this manuscript is the lack of analysis of the measurements in terms of uncertainties and quality control. There are very few quantitative indicators which could help to justify some choices in the protocols. For CRDS measurements only precisions estimates are provided, whereas repeatability's are provided for GCs. Even more problematic is the use of the precision (based on standard deviations calculated over one minute intervals) as a 'demonstration' that measurements are compliant with WMO recommendation for compatibility of monitoring sites. But it seems that few efforts were put to characterize the possible biases (e.g. no measurement of a target gas as recommended by WMO/GAW). I would expect at least some analysis of the existing information (e.g. variability of the calibration residuals), description of the troubleshooting, etc. . . Also I am very surprised some suspicious signals are not discussed at all, even though the protocol described in the manuscript claims a significant emphasis of the data review.

Response: The aim of this paper was to predominantly describe the setup of the UK DECC network, including the instrumentation used, sampling and calibration protocols and data processing methods to help guide the setup of other future sites and networks. The addition of the data since the network starts help show what data there is available and the potential uses for it. No in depth analysis was presented in this paper of the data as the

authors felt that this would be better done in a paper where interesting signals and patterns can be looked at in conjunction with inverse modelling results.

No quantification of uncertainty within the data has been presented in this paper as this is something that the authors want to work on to include with the data. This has been outlined in the final paragraph of the summary and conclusions, where we have added in the future improvements for the network. We realise the added benefit of using target tanks as an independent quality control measure; however, when the network was set up, we used protocols based on GC measurements within AGAGE, which has not historically measured target tanks. The inclusion of target tanks within the network is another aspect that we are looking into to include within the network.

We recognise the issues related to the use of precision information for CRDS measurements and repeatability values for GC measurements. We have changed our precision data to short-term precision and included repeatability data for the CRDS instruments, which is outlined in the specific comments below. In addition, we have removed the comparison with the WMO recommendations as we realise that our precision data does not fully encapsulate the error from measurements in the network.

Three extra sections have been added to this manuscript to improve information on the reader, including a troubleshooting section (section 5.4) and a section on maintenance (section 3.8) and a recommendations section for readers wishing to create future sites or stations.

Page 1, line 14: "A network of . . .three. . . tall tower"

Response: Done.

Page 1, line 24: "The long-term 1-minute mean precisions (1s)" You should show an indicator comparable to GC

Response: This has been changed to short-term precision (1σ of 1 minute mean data per standard injection) and a repeatability of standard injections (1σ of 20-minute mean injections).

Page 3, line 2: "the accuracy of the inversion is limited by the number and distribution of measurement locations available": I would suggest to mention also the capacity of models to properly represent the observed time series, which lead to favor tall towers over flat terrains.

Response: Done.

Page 3, line 9: "Measurements. . .constrained global or hemispheric scale fluxes": I would not say that atmospheric measurements constrained the fluxes (also on the next line of the same paragraph). Hopefully they can give some constraints to the estimation of the fluxes by inverse method.

Response: Text altered to give clarity for both sentences.

Page 4, line 3: I do not think the list of species measured by Medusa is needed in Table 1. Please refer to a publication.

Response: Medusa species removed, apart from SF_6, and Miller et al. (2008) and Arnold et al. (2012) papers were referred to.

Page 5, line 4: 51% of marine air at MHD: please give a reference

Response: Reference added

Page 6, line 2: Is there a reason why you chose to sample N2O, SF6, CO at 100m a.g.l. and not at the highest point of the tower (185m). Also, why did not you adopt the same sapling strategy at TTA where you have a single inlet for CO2, CH4, whereas you set up 2 or inlets at the other tall towers ? Could you please explain your choices ?

Response: The reason for sampling N_2O, SF_6 and CO at the 100 m.a.g.l. inlet is purely historical. When the site was setup in 2012, the 185 m.a.g.l. inlet was not installed. A sentence has been added about the later installation of the 185m inlet. Multiple inlets were chosen at sites except TTA to try and assess boundary layer stratification and is a protocol used by a number of different networks with tall tower sites, including ICOS. The added advantage of having multiple inlets is that if a line pump fails or a filter blocks on an inlet, the affected line can be taken out of the sampling strategy. Additionally, data from other inlets can be used to help diagnose issues in data by looking at patterns between each inlet (see section 5.4 for more details). The sampling line set up at TTA was different as the site was inherited by the University of Bristol and the cost of adding additional sampling lines was too much for that site.

Page 7, line 10: What do you mean by: 'Horizontal sections of tubing at the base of the tower were kept to a minimum' ? Please clarify.

Response: From the base of the towers, the tubing runs horizontally along trunking until it reaches the laboratory. We tried to minimise these long horizontal stretches to prevent water from accumulating and the potential for contaminants to alter the composition of the air drawn down from the inlet.

Page 7, line 17: Have you tested the Perspex H2O decanting bowls to ensure the noncontamination of the measurements ? Why did you changed from the previous system used at TTA ?

Response: Yes, the Perspex H_2O decanting bowls were tested in the laboratory before being installed at. A steady gas stream of ambient air (whole air compressed into a cylinder) was passed through Synflex 1300 tubing and analysed on a CRDS with and without the H_2O decanting bowls in line to observe for differences in CO_2 and CH_4 mole fractions. No effect was found. A sentence has been added to state that the bowls were tested before being installed. Perspex H_2O decanting bowls were used in preference over the stainless steel bowls to enable site operators to see if H_2O had accumulated within the bowl.

Figure 2: Can you precise the meaning of TOC on the figures.

Response: TOC (zero air generator) has been added to the figure legend.

Page 8, line 20: "This has the advantage of eliminating sample contamination from the pump, reducing the likelihood of a torn diaphragm introducing laboratory air into the sample and improving the performance of Nafion dryers." Can you please elaborate on this sentence ? The first point is clear, but I understand that the inlet is slightly under pressure which is not an advantage to avoid the contamination from lab air.

Response: The Nafion section was removed as it was found to be incorrect.

Page 10, line 20: "The correction applied is minimized due to the removal of most H2O using the Nafion dryer." In return you introduced a possible CO2 bias due to the nafion. How have you quantified this bias ?

Response: The passing of calibration and standard gases through the Nafion dryer makes bias associated with CO_2 permeation across the membrane negligible (Welp et al., 2013; Andrews et al., 2014). Laboratory tests by Andrews et al. (2014) showed that losses of CO_2 from sample air and calibration gases were the same.

Page 12, line 17: "Each cylinder is now individually analysed as a sample to check for contamination prior to use". Can you precise how frequently you get problem with the purity of the carrier gas ?

Response: The frequency of SF_6 contamination within P5 varies greatly as it depends on the cylinder age. SF_6 was used to pressure test cylinders nearing the end of their inspection date to check for leaks and cylinder stability. On average, about one in every six cylinders were contaminated with trace amounts of SF_6. A section was also added in the recommendations about testing carrier gases before use (section 7.4).

Page 15: "4.1 Sampling sequence" : Is there any justification of the different configurations at the sites. For example why 20 or 30 min sampling time ? Is it because of the number of sampling levels ? Also could you clarify if the GCs are measuring at only one level ?

Response: As stated within the text, the justification for different configurations in CRDS sampling times is to ensure that all inlet heights are measured within one hour and is purely based on the number of sampling heights at the site. A sentence has been added to clarify that the GC are only sampled from one height.

Page 16, line 5: Can you precise the lifetimes of the standard and calibration gases like you have done for the GCs standards ? Are the standards mixing ratios reevaluated at the end ?

Response: A sentence has been added to the previous paragraph (first paragraph of section 4.2.1) stating the lifetimes of CRDS calibration and standard gases. A sentence has also been added stating that cylinders are recalibrated once removed from site.

Page 16, line 9: "line flushing": are you choices in duration, number of cycle and frequency of the calibration gases measurements based on specific tests ?

Response: Choices were made according to suggestions made by ICOS when setting up the network rather than specific tests. Specifications used by ICOS are available in Hazan et al. (2016).

Page 16, line 16: "long-term precision": what about the measurement repeatability ?

Response: This has been changed to long-term repeatability based on standard injections on the CRDS. A short-term precision has also been included.

Page 16, line 19: "CRDS precision within the UK DECC network is within the WMO compatibility guidelines for CO2 (± 0.1 µmol mol-1) and CH4 (± 2 nmol mol-1) (WMOGAW, 2014)": Indeed the precision is lower than the WMO recommendation for compatibility, but this is comparing apples and oranges. The WMO compatibility has to take into account the measurement repeatability, non-linearity, calibration uncertainties, H2O correction, dryer and inlet biases. Please remove or rephrase correctly this sentence. Regarding WMO recommendations it should be noted that you are not following the recommendation of measuring a target tank ("Each analysis system must include at least one 'target tank' which is a very important quality control tool ", WMO/GAW report n ◦ 229, 2015)

Response: Sentence removed.

Page 17, line 16: "Calibration scales vary depending on the gas species": the CH4 scale for GC-FID measurements is not given. Is it a different scale compared to CRDS measurements ? If so what about the compatibility of the two scales ? Also the CH4 is not mentioned at all in paragraph 5.2 (GC data processing).

Response: CH_4 measured on the MHD GC-FID is on the Tohoku University scale. Based on in situ measurements made at MHD by AGAGE and NOAA flask measurements since 1993, there is an average difference of 1.01 nmol mol^{-1} (NOAA-AGAGE) (Krummel, 2018). CH_4 has been added into section 5.2.

Page 17, line 24: "periodically in the field": Can you precise the frequency of the nonlinearity tests for N2O and SF6 ?

Response: Tests were conducted approximately every year. The sentence has been altered to reflect this.

Page 19, line 4: "H2O level too high": what is the reason of rejecting values with high H2O level ? what is the typical threshold values ?

Response: Data was filtered out when H_2O values were > 6% and was rejected as this was the highest mixing ratio used in the H_2O correction in Rella (2010) . Additionally, data with such a high H_2O concentration was indicative of liquid water passing the H_2O decanting bowl. Table 7 was added to show filter parameters.

Page 19, line 14: "A second order function can then be fitted to the data to provide a non-linearity correction": CRDS instruments have the reputation to be quite linear. Why do you apply a non-linear correction? Have you estimated the importance of this term ?

Response: For all of the UK DECC network sites there is a small non-linear response on the CRDSs. The non-linear response is not thought to be due to calibration gases used within the network as the non-linear effect has been seen when NOAA standards have been

used to calibrate instruments that use GCWerks. An extra figure has been added to shown an example of the non-linear fits used within the network (Figure 4) and histograms to show the offsets between data with and without the non-linear corrections (Figure 5). Additional information has been put in the paragraph to exemplify what the effect would be on the data if the non-linear correction wasn't applied.

Page 20, line 11: "using sample integrated height or area": can you precise which one you are using for the different species ?

Response: Done

Page 23, line 3: "There is an approximate 0.8 nmol mol-1 northern hemisphere midlatitude seasonal trend": if you give an estimate of the trend for N2O, I would suggest to do the same for all species, and gather this information in a table.

Response: This sentence has been altered to show the amplitude of the seasonal cycle.

Figure 5: How do you explain the low N2O and SF6 concentrations measured at RGL in the first year (2012) ? Those data look suspicious considering that such event do not appear later on. According to your description of the final data processing (5.3) you are evaluating the time series by comparing the stations with MHD. From my understanding this is typically a case where such comparison could lead to flagging the first year of measurements. Can you provide a possible explanation of this atmospheric signal ?

Response: Within the UK DECC network, data is not manually flagged out unless there is a specific reason for the spurious data, such as instrumental issues or leaks, even if the data looks odd against MHD. Instead, if no reason can be found, the data is left in the time series. The N_2O and SF_6 data shown in Figure 5 (now Figure 7) from the start of the time series to 11/07/2012 is noisier due to poorer instrument precision, which improved in July 2012 when more insulation was added to around the inlet of the post-column. Hence why both gases have periods with lower mole fractions than MHD.

Figure 7: Please precise if you are using local or UTC time on this figure.

Response: Done, UTC.

Figure 7c: CH4 diurnal cycle at RGL shows a very high standard deviation in midday/summer. Outliers which could cause such anomaly on the 2012/2015 average signal are not visible in the time series on figure 4.

Response: Plot altered to show all of data.

Page 25, line 5: "Fig. 8(b) demonstrates a regionally polluted period at RGL for CH4 on 30/11/2014, where air has passed over Europe and the south of the UK before arriving at the site": in such a case we could expect higher CH4 concentrations at MHD compared to TAC, due to the additional contribution from south UK. Do you such 'reverse' gradients sometimes ?

Response: If both sites are receiving the same air masses, then an increasing gradient of CH_4 mole fractions would be expected between RGL and MHD. However, in the instance shown in Figure 8(b), now figure 10(b), the air mass that RGL received was from over the midlands and south of the UK and the Benelux region and Germany; however, at this time, MHD was received air that had passed over the south west of the UK, the west of France and Spain as shown in Figure 1 below. The reverse gradient from that shown in Figure 10(b) – an increasing concentration from west to east over the UK, is frequently observed due to the UK predominant wind direction. Figure 10(a) demonstrates the increasing CH_4 mole fraction with longitude on the 4th and 5th December. A sentence has been added to section 6.3 to reflect this.

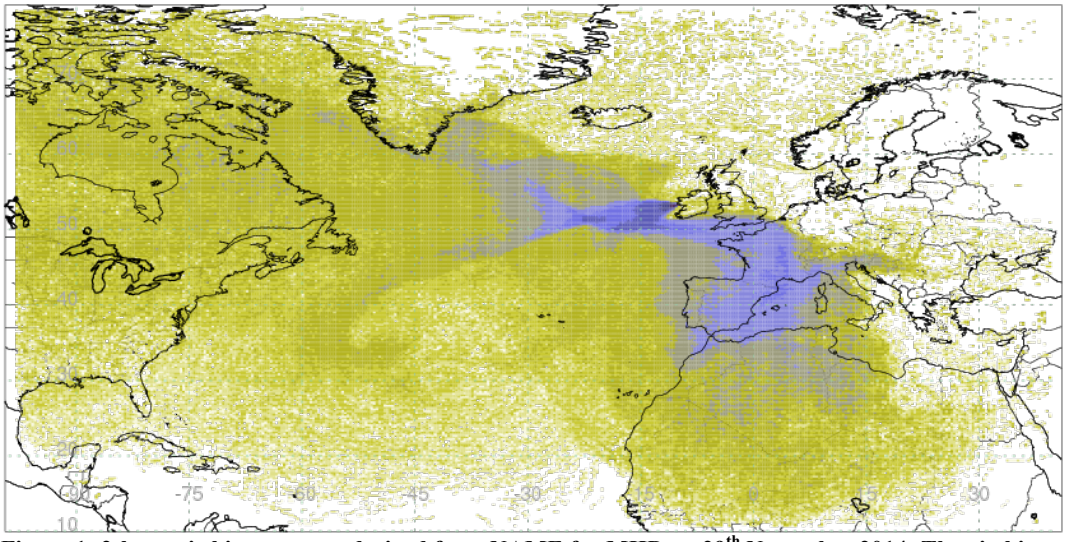

Figure 1: 2-hour air history maps derived from NAME for MHD on 29th November 2014. The air-history maps describe which surface areas (0-40m) in the previous 30-days impact the observation point within a particular 2-hour period.

Page 26, line 6: "Results from the network give good spatial and temporal coverage": I fully agree for the temporal resolution, but I would remove the comment on the spatial coverage, the relevance of which depends on the intended scientific purpose.

Response: Spatial has been removed.

References

Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., Chao, D. Y., Kitzis, D. R., Novelli, P. C., Zhao, C. L., Dlugokencky, E. J., Lang, P. M., Crotwell, M. J., Fischer, M. L., Parker, M. J., Lee, J. T., Baumann, D. D., Desai, A. R., Stanier, C. O., De Wekker, S. F. J., Wolfe, D. E., Munger, J. W., and Tans, P. P.: CO2, CO, and CH4 measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts, Atmos. Meas. Tech., 7, 647-687, 10.5194/amt-7-647-2014, 2014.

Arnold, T., Mühle, J., Salameh, P. K., Harth, C. M., Ivy, D. J., and Weiss, R. F.: Automated Measurement of Nitrogen Trifluoride in Ambient Air, Analytical Chemistry, 84, 4798-4804, 10.1021/ac300373e, 2012.

Hazan, L., Tarniewicz, J., Ramonet, M., Laurent, O., and Abbaris, A.: Automatic processing of atmospheric CO2 and CH4 mole fractions at the ICOS Atmospheric Thematic Center, Atmos. Meas. Tech. , 9, 1-34, 10.5194/amt-9-4719-2016, 2016.
Krummel, P.: RE: MHD NOAA-AGAGE difference for CH4. Personal communication to Stanley, K. M. on 15/01/2018.
Miller, B. R., Weiss, R. F., Salameh, P. K., Tanhua, T., Greally, B. R., Muhle, J., and Simmonds, P. G.: Medusa: A Sample Preconcentration and GC/MS Detector System for in Situ Measurements of Atmospheric Trace Halocarbons, Hydrocarbons, and Sulfur Compounds, Anal. Chem., 80, 1536-1545, 2008.
Rella, C.: Accurate greenhouse gas measurements in humid gas streams using the Picarro G1301 carbon dioxide/methane/water vapor gas analyzer, White paper, Picarro Inc, Sunnyvale, CA, USA, 2010.
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.