**Atmospheric
Measurement
Techniques
Discussions**

Interactive comment on "Results from the first national UK inter-laboratory calibration for very short-lived halocarbons" *by* C. E. Jones et al.

C. E. Jones et al.

jones@tmu.ac.jp

Received and published: 22 April 2011

The authors thank both reviewers for their constructive comments. We address the specific comments from anonymous referee #2 below. All page and line numbers refer to those in the revised manuscript.

1) One of the greatest vulnerabilities concerns the NOAA standard that was circulated among the laboratories for this calibration effort. As the paper notes, the NOAA laboratory was willing only to assign provisional values to this standard canister, and do so only to two significant digits, because of concerns regarding stability. Yet the NOAA values are described in the text as "certified", while at end of the manuscript (page 780, lines 24-26) recalibrating this canister at NOAA is described as something that should be done in the future. In my view, it should have been done before this paper was writ-

ten and submitted for publication. This is a major flaw in the manuscript that is noted also by Anonymous Referee #1 (in his/her comment #3). It is standard procedure in this business, especially when dealing with trace gases that are known not to store easily, to measure "out" and "in" values for a field standard before publishing anything.

NOAA assigned "provisional" mixing ratios only for those iodocarbons lacking long-term canister stability tests (since these compounds - such as CH2ICl and CH2I2, are not routinely monitored by NOAA, their long-term stability in these canisters has not been studied). In contrast, the mixing ratios assigned by NOAA for the VSLH species studied for this inter-calibration, namely CH2Br2, CHBr3 and CH3I should not be considered as just "provisional", but quantified according to the NOAA scale, albeit with the quoted uncertainties. We do note that the lack of a post inter-comparison re-analysis by NOAA is indeed a limitation, however the stability of CH3I, CH2Br2 and CHBr3 in NOAA gas standards has been monitored over periods of several years and these gases were found to experience a drift of <2 - 5 % yr-1 (see also response to reviewer 1, point 3). We have added the following sentence on Page 5 L144-151, to clarify this: "Stability studies of NOAA halocarbon gas standards prepared and contained under the same conditions as SX-3570 indicate that CH3I, CH2Br2 and CHBr3 mixing ratios are generally relatively stable over this timescale, with drift rates of <2 - 5 % yr-1 (Butler et al., 2007, suppl. mat.). In addition, NOAA-ESRL have carried out longer term stability tests on six such canisters containing continental background mixing ratios of CHBr3, CH2Br2, and CH3I (< 1 ppt) over a 3 year period (2004-2007; during the International HALocarbons in Air Comparison Experiment (IHALACE), manuscript in preparation), and all three gases showed no significant changes over this time."

2) There is mention of "the NOAA scale" but the paper doesn't say which NOAA scale. Such calibration scales are not absolute, but rather are defined within NOAA and in other groups that do primary calibrations as being related to a specific suite of primary standards. These scales are generally given a name (e.g. NOAA-2005) and have a citable reference or website where details are given. In short, the paper needs to

say which NOAA scale is used, and it needs to be clear whether the UEA and UoC measurements used the same NOAA scale or a different one. Finally, although the units of the NOAA scale (dry air mole fraction) can be found in the text, these units should also be given in the caption for Table 1 where the data are listed.

Apologies, this was an oversight on our part not to indicate the specific NOAA scales used. We have added the following sentence (Page 5, L137-138) detailing the scales used by the NOAA laboratory: "CHBr3 was calibrated based upon the NOAA-2003 scale, whilst CH3I and CH2Br2 were calibrated according to the NOAA-2004 scale." We have also added the following sentences detailing the NOAA scales used by Cambridge and UEA, respectively: "The inter-calibration standard SX-3570 was compared to the Cambridge working standard (NOAA cylinder no. SX-3568; CHBr3 calibrated according to the NOAA 2003 scale and CHÂň3I and CH2Br2 calibrated based upon the 2004 scale)..." (Page 7 L192-195) and "The latter was brought in line with NOAA calibration scales (the 2003 scale for CHBr3 and 2004 scale for CH3I and CH2Br2) by measuring it against two tertiary standards...." (Page9, L253-255).

3) Instrumental non-linearities and blank corrections are not mentioned anywhere in the discussions of instrumental methods for the five British laboratories in the study. Each laboratory should say whether blanks and instrumental non-linearities were addressed. This is a fundamental part of analytical chemistry.

The following sentences have been included to outline instrument linearity and blank corrections for each group's analyses: Page 7 L184-186: "No blank corrections were applied to the data. The Medusa-MS has demonstrated linearity for concentrations of up to 12 x typical background ambient mixing ratios for CHBr3 and 8 x ambient CH2Br2 mixing ratios." Page 7 L200-203: "Blank chromatograms were run following each sample or calibration chromatogram by passing helium carrier gas through the adsorbent bed (these served as a check on desorption efficiency and system impurities and were sufficiently clean that blank corrections were not necessary)." Page 8 L220-222: "These enable the tracking of small changes in detector non-linearity, which

C427

are taken into account in the work-up of the data. For CHBr3 and CH2Br2, the response is reasonably linear (Gostlow et al., 2010)." Page 9 L257-260: "Throughout these analyses, blank signals were much smaller than the reported measurement precision, and were thus not considered. Analysis of the SX-3570 standard demonstrates that this system produces a linear response for CH3I, CH2Br2 and CHBr3 up to ~2-5 ppt." Page 10 L283-284: "The instrument response was linear for 0.5 - 2.0 L gas for all three compounds (R2 > 0.99), and no blank corrections were applied." Page 10 L291-293: "The instrument response was linear for CH3I (0 - 4 pmol L-1, R2 > 0.98), CH2Br2 (0 - 14 pmol L-1, R2 > 0.99) and CHBr3 (0 - 15 pmol L-1 R2 > 0.99)." Page 11 L 321-323: "Sample volumes of 1.0, 1.5 and 2.0 L of the NOAA inter-calibration standard were analysed to produce a linear response curve. No VSLH signals were detected in nitrogen-only blank samples, and as such blank corrections were deemed unnecessary."

4) Although NOAA-ESRL is not a British laboratory, the dependence of this paper on NOAA calibrations requires that the methods used at NOAA at least be described briefly, and that a reference to the NOAA methods be given.

Agreed. We have included the following summary of NOAA analytical method and references to further details on Page 5 L132-135 : "NOAA-ESRL determined VSLH mixing ratios in SX-3570 by gas chromatography with mass-selective detection (GC/MS) (Montzka et al., 1993 and 1996). Samples were analyzed relative to gravimetrically-prepared standards (cylinders similar to SX-3570, Butler et al., 2007)."

5) In two places in the paper the NOAA calibrations are described as "certified", yet the numbers to which this refers are actually described as "preliminary" and are given with limited precision for that reason. In fact, no such calibrations should be regarded "certified" or "correct" âÂćAÂąT no authority is absolute. It would be better to call the NOAA values "assigned" mixing ratios, or something to that effect.

We agree that the use of the word "certified" is inappropriate here, and so have

changed "certified" to "assigned" (Page 12 L360 and in the Table 1 title). However we would like to stress again that the NOAA-assigned mixing ratios for CH3I, CH2Br2 and CHBr3 are not regarded as "preliminary" (see response to point 1).

6) The paper includes many subjective statements that are inappropriate to this quantitative subject, such as that numbers are "in excellent agreement", "good agreement", reasonably good agreement", or "very close agreement". Wherever possible these subjective statements should be replaced by objective ones that relate the degree of agreement to what is expected on the basis of measurement statistics.

Agreed. The following sentences have been changed such that discussion of the level of agreement between measurements now includes some quantification. Page 13 L381-384 – "Although both institutions report CHBr3 mixing ratios ∼0.8 ppt larger than the NOAA-assigned value, and in excellent agreement with each other (within 1 %), all three values are consistent within the uncertainties of the measurements." Page 13 L388-391 – "Despite not currently using the NOAA scale for VSLH calibrations, both the Universities of York and Bristol report CHBr3 mixing ratios in excellent agreement with the NOAA-assigned value (within 2 %), and CH3I mixing ratios equivalent to the NOAA value, within the measurement uncertainties." Page 13 L399-403 – "The PML-assigned CH2Br2 mixing ratio was nevertheless in good agreement with the NOAA analyses (within ∼12 %, and equivalent within measurement uncertainties). However, the CH3I and CHBr3 mixing ratios reported by PML differed from the NOAA-assigned values by ∼25 % and ∼45 %, respectively, which is outside of measurement uncertainties." Page 14 L413-419 – "The mean and median of all the independently derived CH3I, CH2Br2 and CHBr3 mixing ratios are in reasonably good agreement (within 2 - 15 % for median, 7 – 21 % for mean) with the NOAA-assigned values. The largest discrepancy is between the average independently derived CHBr3 mixing ratio and the figure reported by NOAA, however this is skewed by the high PML mixing ratio value; removing this value gives a mean (and median) of ∼5.8 ppt CHBr3, bringing the agreement to within 8 %."

7) Statistical precisions are reported as ± one standard deviation, while estimated systematic uncertainties are reported as ± two standard deviations. This is confusing. But more importantly, the authors should recognize that many systematic errors, such as those caused by contamination or by impure reagents, are not normally distributed and often have only one sign. The use of the standard deviation to describe such errors is therefore inappropriate.

Of course, a single systematic error is likely to have only one sign. However, the total errors are a combination of bias (e.g. potential wall losses/gains, flow errors, gravimetric errors and precision errors). The precision error is presumed to behave randomly, with a zero mean. Both the bias and precision are presumed to represent stationary statistical properties of a Gaussian distributed data set (a stationary process is a random process where all of its statistical properties do not vary with time). As long as the errors are un-correlated, then it is acceptable to estimate the total error from the sum of squares of all the bias and precision errors (Moffat, R. J., 1988 - for full reference please see manuscript).

It is appropriate to assume that the bias limit for fixed errors (e.g. manufacturer's stated error for mass flow meter flow rate) represents a 2σ rather than 1σ error (Moffat, 1988). However, most researchers report 1σ error for precision, and so that is what we included in the Table. In hindsight, we agree that this is confusing and have therefore reported the 2σ precision errors in Table 1. We have also changed our values for total measurement uncertainty such that U is now calculated from the root sum of squares of the 2σ precision errors and 2σ bias errors. We have added the following statement to clarify how we have estimated the measurement uncertainty (Page 12 L362-369): "The mixing ratios determined by each laboratory are reported in Table 1, together with the associated 2σ analytical precision and estimated 2σ overall measurement uncertainty. The total uncertainties are a combination of bias (e.g. potential wall losses/gains, flow errors, gravimetric errors) and precision errors. The precision error is presumed to behave randomly, with a zero mean. Both the bias and precision are presumed to rep-

resent stationary statistical properties of a Gaussian distributed data set. Assuming that the errors are un-correlated, the total uncertainty may be estimated from the sum of squares of all the bias and precision errors (Moffat, 1988)."

8) I do not agree with the assertion (page 779, lines 22-25) that calibrations are somehow tied to the methods that are used. What is true is that different methods are vulnerable to different kinds of errors, but all methods applied properly ought to give the same answers. Perhaps this statement is just poorly worded. I also agree with Anonymous Referee #1 (in his/her comment #8) that further discussion of the comparisons between gas phase and liquid phase measurements would strengthen the paper significantly. These are quite different communities, and I am sure that there is much to be gained by bringing them closer together and recognizing what steps ought to be taken to reconcile their different approaches. Indeed, some of the answers may lie in issues of blank correction and instrumental non-linearity mentioned above.

We agree that in hindsight this sentence was not well worded and has been removed. Regarding gas phase vs seawater comparison, we have added the following discussion of the discrepancies (Page 14, L405-412): "The apparent discrepancies between aqueous and gas phase CH_3I and $CHBr_3$ calibrations result in PML under-estimating CH_3I and over-estimating $CHBr_3$ mixing ratios in the SX-3570 gas standard. Over-estimation of $CHBr_3$ might potentially be explained by incomplete sparging of $CHBr_3$ from seawater standards (although purge efficiencies have previously been measured, and are taken into account in the PML calibrations, see section 2.2.4), whilst under-estimation of CH_3I could result from breakthrough when sampling SX-3570 gas onto the sorbent tubes. CH_3I is the most volatile of the three VSLH studied here, and thus is most susceptible to breakthrough."

In addition, we have added the following consideration as to how to conduct future inter-comparisons between both gas phase and aqueous VSLH calibrations (Page 15 L456-467): "Given the apparent discrepancies between aqueous and gas phase calibrations for CH_3I and $CHBr_3$ observed in this study, we suggest that future comparisons be-

C431

tween aqueous and gas phase instruments should (a) use a range of primary liquid standards when preparing working standards to assess the potential variability of using liquid compared to gaseous standards and (b) evaluate any systematic differences between gas phase and aqueous phase analyses of different methods. The latter could be achieved by using a common gaseous standard for calibration of instruments, in-conjunction with analysis of a common aqueous sample (analyses of aqueous samples would have to be carried out near simultaneously to avoid any degradation). Groups using liquid standards should be able to report the comparability of these from one campaign to the next, i.e. be able to compare to a primary standard. Likewise, groups using NOAA gaseous standards for calibration should routinely return cylinders for re-analysis. "

9) I agree with Anonymous Referee #1 (at the end of his/her comment #5) that field comparisons are not called for if laboratory comparisons are problematic.

The authors feel the main benefit of a field-based comparison is that it requires co-location of all instruments and simultaneous measurements, removing any uncertainties related to the stability of shorter-lived VSLH when stored in canisters. Perhaps this was not explained clearly in the initial submission, but has now been reworded (please see details in response to reviewer 1, point 5).

10) In the description of the UoB measurements, it should be made clearer that there is no Scripps primary calibration for these gases, and that Scripps only checked for drift in the tertiary standard used for the UoB measurements.

We have modified details on Page 6 L175-179 to clarify this: "The NOAA standard SX-3570 was compared against a tertiary standard (J-128), filled during relatively clean-air conditions at Trinidad Head using modified diving compressors (SA-6, RIX Industries), and measured against secondary standards at Scripps Institution of Oceanography (SIO) to check for drift before and after usage at Mace Head."