TAbMEP Assessment: ICARTT *n*-Butane Measurements

1. Introduction

Here we provide the assessment for the *n*-butane (C_4H_{10}) measurements during the summer 2004 ICARTT field campaign [*Fehsenfeld et al.*, 2006, *Singh et al*., 2006]. This assessment is based upon the four wing-tip-to-wing-tip intercomparison flights conducted during the field campaign. Recommendations provided here offer TAbMEP assessed biases for each of the measurements and a systematic approach to unifying the ICARTT *n*-butane data for any integrated analysis. These recommendations are directly derived from the instrument performance demonstrated during the ICARTT measurement comparison exercises and are not to be extrapolated beyond this campaign.

2. ICARTT *n*-Butane Measurements

Three whole air sampler instruments were deployed on three aircraft. Table 1 summarizes these techniques and gives references for more information.

Table 1. *n*-Butane measurements deployed on aircraft during ICARTT

Aircraft	Instrument	Reference
NASA DC-8	Whole Air Sampler (WAS)	*Colman et al.* [2001]
NOAA WP-3D	Whole Air Sampler (WAS)	Contact PI: eatlas@rsmas.miami.edu
FAAM BAe-146	Whole Air Sampler (WAS)	*Hopkins et al.* [2003]

3. Summary of Results

Table 2 summarizes the assessed biases as well as PI reported uncertainties for each of the three *n*-butane measurements involved in the intercomparisons. More detailed descriptions are provided to illustrate the process for the bias assessment in Section 4.1. The TAbMEP-prescribed IEIP procedures cannot be applied to the ICARTT *n*-butane measurements for precision assessment. This is because the reported data have large time gaps and a small data population (see Section 3.1 of the introduction). The assessed bias reported in Table 2 (see Section 4.1 for details) can be applied to maximize the consistency between the data sets, by subtracting the value from the reported data to ‘unify’ the data sets. If one assumes instrument performance remained constant throughout the mission, the assessed bias may be extrapolated to the entire mission although it is derived from intercomparison periods only. No assessed bias is included for the FAAM BAe-146 because there were only two overlapping points both platforms reported non-LOD (limit of detection) values. Although good agreement is shown for these points, there are two other overlapping points where DC-8 reported LOD but BAe-146 gave values around 6 pptv. Given these mixed results, no definitive assessment can be made with reasonable level of confidence.

The DC-8 and WP-3D uncertainties reported by PIs are a percentage. This may not be adequate at very low end concentration levels. Ideally, the measurement uncertainty may be better represented in the form of x pptv or y%. Based on the intercomparison data, the TAbMEP analysis cannot provide such assessment. Data users should contact the respective PIs about the proper uncertainties when dealing with the low end of measurements, e.g., < ~20 pptv.

Table 2. Recommended ICARTT *n*-butane measurement treatment

Aircraft/Instrument	Reported 2σ Uncertainty	Assessed Bias (pptv)
NASA DC-8 WAS	10%	$-2.631 + 0.0705\ C_4H_{10\ DC\text{-}8}$
NOAA WP-3D WAS	10%	$3.063 - 0.0821\ C_4H_{10\ WP\text{-}3D}$
FAAM BAe-146 WAS	Point by Point, average: 48%[a]	N/A

[a] The average encompasses only the comparison period for DC-8/BAe-146.

Figures 1 a and b display the PI reported uncertainties and recommended biases for the three *n*-butane instruments.

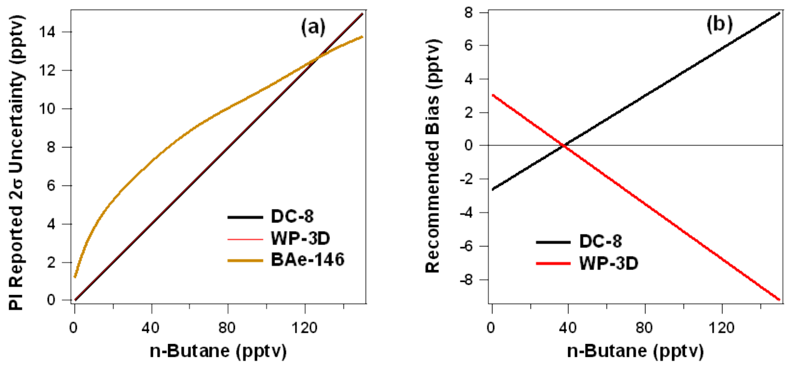

Figure 1. PI reported 2σ uncertainty (panel a) and recommended bias (panel b) for DC-8 (black), WP-3D (red), and BAe-146 (gold) as a function of *n*-butane level. Values were calculated based upon data shown in Table 2. The BAe-146 PI reported uncertainty was calculated using a function derived from the 60 second merge file.

4. Results and Discussion

4.1 Bias Analysis

Section 3.3 in the introduction describes the process used to determine the best estimate bias. Figures 2 and 4 show the time series plots for the DC-8/WP-3D and DC-8/BAe-146 comparisons. The DC-8 is generally lower than the WP-3D by 1 pptv on average, with a majority of the variance about 4 pptv. Figure 4 shows four overlapping points. For two of them at the beginning of the intercomparison period, both BAe-146 and DC-8 reported values and have agreement within 0.5 pptv; while the difference for the other two overlapping points can be as large as 6 pptv where DC-8 consistently reported LOD values. Figure 5 shows the magnitude of the bias for the intercomparison and Figure 6 shows the corresponding relative residuals. Both plots suggest that the PI reported uncertainties may not be adequate for the lowest part of data set, e.g. < ~20 pptv.

For 2 of the 3 DC-8/WP-3D flights, there are only 3 or 4 overlapping points with a small range of variation (about 10 - 60 pptv). It is not statistically significant to show the linear regression for these flights. Therefore, linear regression is performed over the data combined from all three DC-8/WP-3D flights. The linear relationships listed in Table 3 were derived from the regression equation found in Figure 3. The reference standard for comparison (RSC), as defined in the introduction, is constructed by averaging the NOAA WP-3D and NASA DC-8 measurements with equal weights. The FAAM BAe-146 is not included in the calculation of the reference standard for comparison. As discussed earlier, there were only four comparison points and LOD values were involved with two of the four points. Both factors restricted our ability to perform a robust assessment between BAe-146 and DC-8 measurements. The resulting RSC can be expressed as a function of the DC-8 C_4H_{10} measurements as the following:

$$RSC_{C4H10} = 2.631 + 0.929\ C_4H_{10\text{-}DC8}$$

The RSC is then used to calculate the best estimate bias as described in Section 3.3 of the introduction. It should be noted that the initial choice of the reference instrument (DC-8 WAS) is arbitrary, and has no impact on the final recommendations. Table 3 summarizes the assessed measurement bias for two of the ICARTT *n*-butane measurements. Note that additional decimal places were carried in the calculations to ensure better precision. It is also noted that the intercept in the equations listed in Table 3 should not be viewed as an offset. These linear equations are used to best describe the linear relation between the measurements.

The WAS technique for measuring VOCs presents some challenges in analyzing the data. The DC-8 data have an integration time of approximately 60-70 seconds, while the WP-3D data have an integration time between 6-11 seconds. For these measurements to be considered simultaneous and correlated, the start and stop times of the WP-3D data must fall within the start and stop times of the DC-8 data. In order to maximize the data coverage for statistical analysis, one exception is made to this rule. If the shorter (WP-3D) integration time falls outside the longer integration time by no more than two seconds, the data points are also considered to be simultaneous. BAe-146 integration times range from approximately 30-60 seconds. Since the DC-8 and BAe-146 have similar integration times, the measurements are considered correlated if the midpoint of DC-8 or BAe-146 fall within the start and stop time of the other measurement. In the case of the *n*-butane DC-8/BAe-146 comparison, several points were below the LOD and are not used for comparison analysis. Only the PI reported data are used in this assessment, and no interpolation is included. It is noted here the integration time difference may potentially be another factor leading to the difference between measurements.

Table 3. ICARTT *n*-Butane bias estimates

Aircraft/ Instrument	**Linear Relationships**	**Best Estimate Bias ($a + b\ C_4H_{10}$) (pptv)**
NASA DC-8 WAS	$C_4H_{10\ DC\text{-}8} = 0.00 + 1.000\ C_4H_{10\ DC\text{-}8}$	$-2.631 + 0.0705\ C_4H_{10\ DC\text{-}8}$
NOAA WP-3D WAS	$C_4H_{10\ WP\text{-}3D} = 5.26 + 0.859\ C_4H_{10\ DC\text{-}8}$	$3.063 - 0.0821\ C_4H_{10\ WP\text{-}3D}$
FAAM BAe-146 WAS	N/A	N/A

As a part of ICARTT intercomparison standard exchange exercises, University of California, Irvine (UCI) prepared the common VOC samples that were sent to University of Miami (Miami), University of New Hampshire (UNH), and University of York (York) for their lab analyses. Some of these same institutions had instruments on the following planes during ICARTT: UCI on the DC-8, Miami on the WP-3D, and York on the BAe-146. The comparison incorporated 9 species, which included *n*-butane. We believe that the inclusion of this comparison result will help the readers better understand the airborne intercomparison analysis. The difference in this lab comparison between the DC-8 and WP-3D instruments was 6 pptv, WP-3D being higher, at a

DC-8 instrument reading of 434 pptv. From the same lab comparison, the difference between the DC-8 and BAe-146 was 7 pptv, BAe-146 being higher. Comparing the ICARTT flights to this lab comparison shows fairly similar results even though the flights only provide a few comparison points.

4.2 Precision Analysis

A detailed description of the precision assessment is given in Section 3.1 of the introduction. The IEIP precision, expected variability, and adjusted precision could not be calculated for *n*-butane because of the small number of points and large time gaps between measurements.

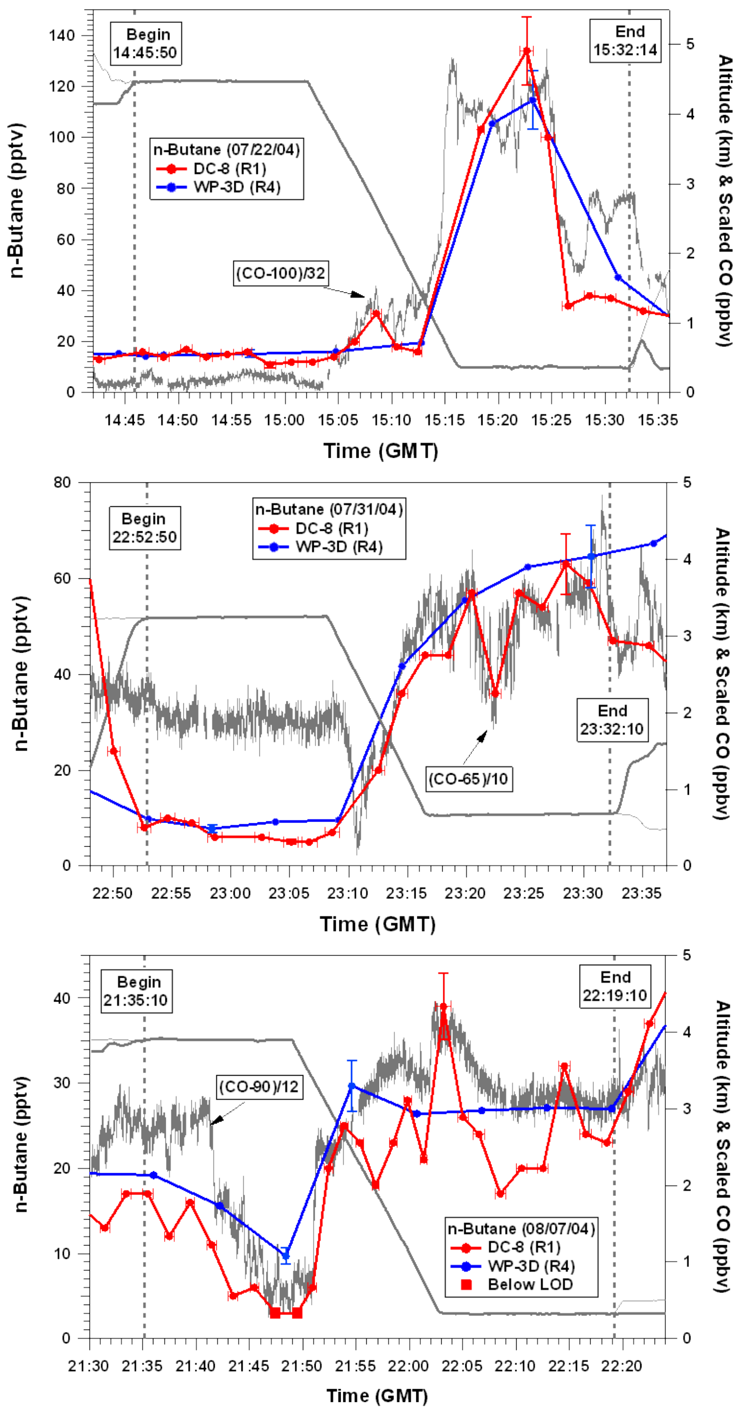

Figure 2. Time series of *n*-butane measurements and aircraft altitudes from two aircraft on the three intercomparison flights between the NASA DC-8 and the NOAA WP-3D. Error bars represent the PI reported uncertainty. In parenthesis next to the plane is the data version number.

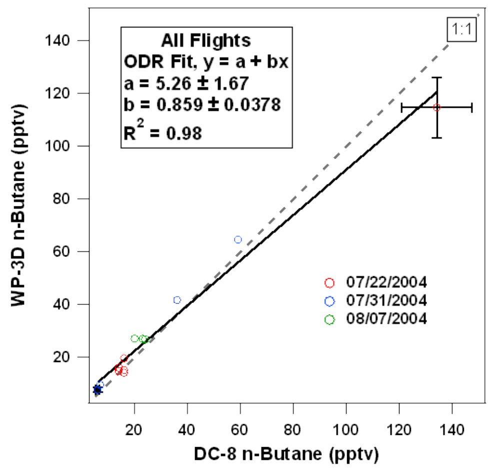

Figure 3. Combined correlation for the *n*-butane measurements on NASA DC-8 and the NOAA WP-3D for 7/22, 7/31, and 8/07 2004. Error bars represent the PI reported uncertainty.

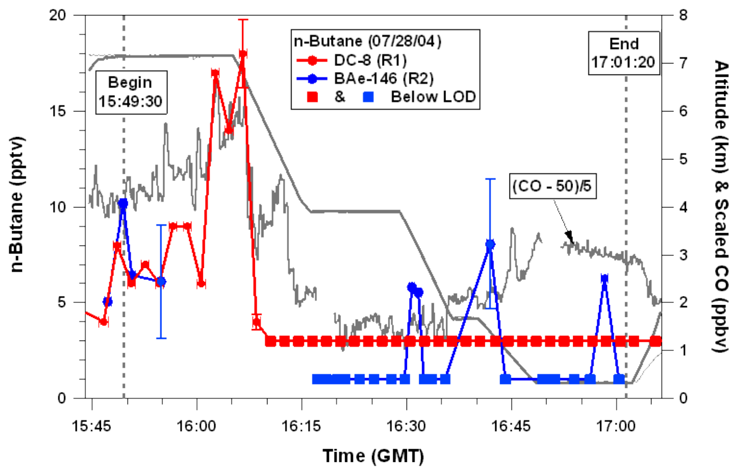

Figure 4. Time series of *n*-butane measurements and aircraft altitudes from the intercomparison flight between the NASA DC-8 and the FAAM BAe-146. Error bars represent the PI reported uncertainty. In parenthesis next to the plane is the data version number.

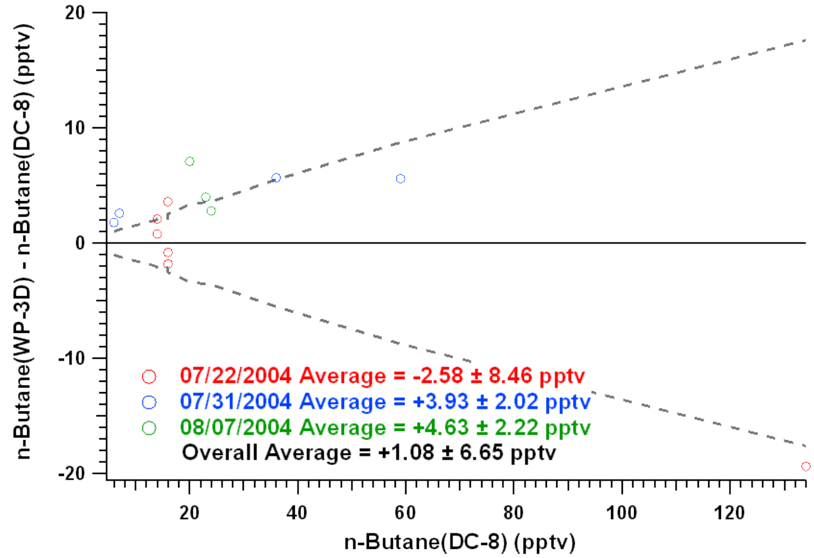

Figure 5. Difference between *n*-butane measurements from the three DC-8/WP-3D intercomparison flights as a function of DC-8 *n*-butane. The dashed lines indicate the range of results expected from the reported 2σ measurement uncertainties.

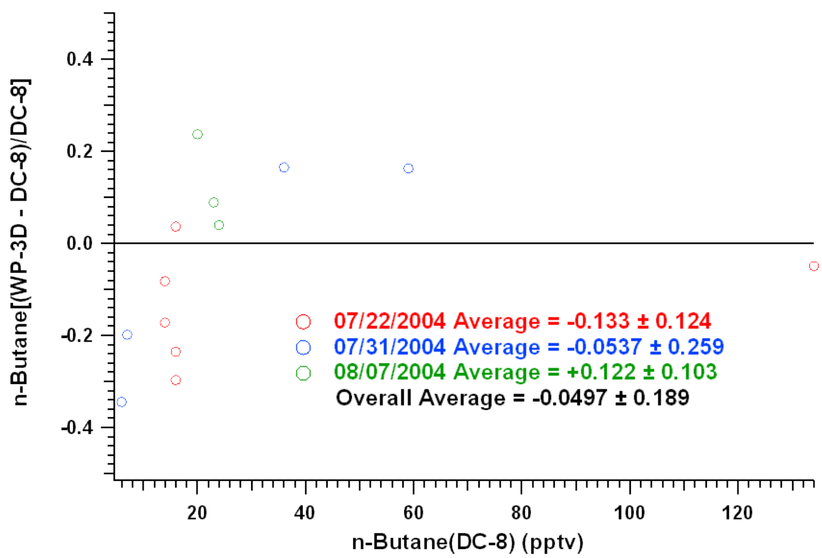

Figure 6. Relative difference between *n*-butane measurements from the three DC-8/WP-3D intercomparison flights as a function of DC-8 *n*-butane. A correction was made to account for bias.

References

Colman, J.J., et al. (2001), Description of the analysis of a wide range of volatile organic compounds in whole air samples collected during PEM-Tropics A and B, *Anal. Chem* ., 73, 3723-3731.

Fehsenfeld, F. C., et al. (2006), International Consortium for Atmospheric Research on Transport and Transformation (ICARTT): North America to Europe—Overview of the 2004 summer field study, *J. Geophys. Res., 111*, D23S01, doi:10.1029/2006JD007829.

Hopkins, J. R., K. A. Read, and A. C. Lewis. (2003), A Column Method For Long-term Monitoring Of Non-Methane Hydrocarbons (NMHCs) and Oxygenated Volatile Organic Compounds, *Journal of Environmental Monitoring*, 5 (1), 8-13.

Singh, H. B., et al. (2006), Overview of the summer 2004 Intercontinental Chemical Transport Experiment-North America (INTEX-A), *J. Geophys. Res., 111*, D24S01, doi:10.1029/2006JD007905.