NPL REPORT AS 55

Calibration Gas Proficiency Testing Scheme Round 5

M B Williams

NOVEMBER 2010

Calibration Gas Proficiency Testing Scheme Round 5

Matthew Williams
Operations Directorate

ABSTRACT
This report presents the results from the fifth round of the calibration gas proficiency testing scheme operated by NPL. Measurements made by participants at NPL from Monday 18th to Friday 28th October 2010.

ISSN 1754-2928

National Physical Laboratory
Hampton Road, Teddington, Middlesex, TW11 0LW

Approved on behalf of NPLML by Martyn Sene, Director of Operations, *Operations Directorate*.

CONTENTS

1 INTRODUCTION

This report describes the results of the fifth round of a calibration gas proficiency-testing scheme carried out by the National Physical Laboratory. Measurements by participants were made at NPL from Monday 18th October to Thursday 28th October.

The calibration gas Proficiency Testing (PT) scheme provides a way of assessing the performance of laboratories by a series of regular inter-laboratory comparisons. The assigned values were provided by traceable measurements of the concentrations of the gases carried out by NPL. The results of each participant's analyses were compared to assigned values of the test gases. The set of results has been reported anonymously, and in addition each participant has been made aware of their own results. In this way participants are able to assess their performance in relation to other laboratories.

2 PROFICIENCY TESTING SCHEME DESCRIPTION

2.1 ORGANISATION

Thirteen companies took part in the fifth round of the scheme, which involved the measurement of a number of traceable standard gas mixtures at NPL.

The PT scheme assesses the measurement of gaseous components using continuous emission monitoring (CEM) equipment. This has the benefit that the test samples each have a known, 'true' value, and that the analysis of these is non destructive, in the sense that the same gas mixture can be analysed by more than one laboratory. The list of species and nominal concentrations of the gas mixtures used in this round of the PT scheme is given in Table 1.

Table 1 Gases used in round 5 of the calibration gas proficiency testing scheme

Species	**Nominal Concentration**
SO_2	200 ppm
CO	200 ppm
NO	100 ppm
O_2	3 %
C_3H_8	40 ppm

The participants were told the nominal concentration of the cylinders they received as given in Table 1, but not the absolute value. The results of these analyses were returned to NPL. Participants were also requested to report the measurement uncertainty that they assigned to the results obtained. Not all participants took part in all tests and not all reported uncertainties.

Participants were requested to inject the test gases direct into the 'sample in' of their gas analyser bypassing all gas conditioning systems. The test gases are dry therefore injection through the entire sampling system would serve no purpose except to highlight any leaks in the sample system. The aim of the scheme is to act as a 'calibration check'; poor performance highlighting a problem with the calibration standards used to calibrate the instrument or an issue with the instrument itself.

2.2 DATA ANALYSIS

Participants were requested to make three repeat measurements of the test gases, these were then averaged to give a single result for each test gas. Each of these averages is expressed as a percentage difference from the assigned value. Using this method it is left up to the participant to judge how well they have performed.

A more refined interpretation of PT scheme results involves the calculation of a performance score for each result. This is usually based on comparing the results achieved against an assigned target standard deviation, σ. The simplest form of this is the 'z score'. This is calculated by dividing the deviation of each result from the true value by σ, see Equation 1.

$$z = \frac{x - T}{\sigma} \qquad \text{Equation 1}$$

where:
- z z score
- x value obtained by participant
- T true value for test sample
- σ assigned value for standard deviation

This provides a z score for each result, which can be compared with other z scores from analyses of the same sample and with analyses of different species. If a suitable value of σ is chosen for each species then the z score also provides a method of deciding decision limits for the PT scheme. In general, if all results are normally distributed about the true value of the test sample and a reasonable value of σ has been chosen, then few (< 5 %) of the z scores should lie outside ± 2. z scores lying outside ± 3 would be strongly indicative of a true bias in the reported value, rather than random uncertainty. From this it is possible to apply a classification as follows:

$-2 \leq z \leq 2$	satisfactory
$-3 < z < -2$ or $2 < z < 3$	questionable
$z \leq -3$ or $z \geq 3$	unsatisfactory

These limits allow each participant to judge their own performance and can be used to indicate potential problems. The target standard deviation is usually taken to be a value, which is fit for purpose for the measurements being made. The value of σ used for the fifth round of the PT Scheme is shown in table 2. The value of σ has been chosen with reference to previous round results, the relevant CEN standards, Waste Incineration Directive (WID) and Large Combustion Plant Directive (LCPD).

Table 2 Standard deviation for proficiency testing

Species	Standard deviation for proficiency testing (σ) as % of value
SO_2	5
CO	3
NO	5
O_2	2
C_3H_8	7.5

A z score of 2 or greater shows significant bias in the measured result and not random noise. Participants who attain z scores of 2 or higher should investigate the cause of the performance with an aim to improving their performance in subsequent rounds. Those with z scores of 3 or higher should put in place a documented mechanism to correct any issues identified.

The assigned values and their uncertainties are shown in Table 3.

Table 3 Assigned values and their uncertainties

Species	Nominal concentration	Assigned value (AV)	Standard uncertainty (k=2)
SO_2	200 ppm	204.6 ppm	2.0 ppm
CO	200 ppm	203.7 ppm	2.0 ppm
NO	100 ppm	103.7 ppm	1.0 ppm
O_2	3 %	2.97 %	0.03 %
C_3H_8	40 ppm	41.1 ppm	0.4 ppm

3 RESULTS

The difference, percentage difference and z scores for each species are show in Table 4 to 8.

Table 4 Proficiency test results for 200 ppm SO_2

Participant	Submitted result (ppm)	Difference from AV (ppm)	% Difference from AV	Z score
1	218.2	13.6	6.65	1.33
2	241.0	36.4	17.79	3.56
3	206.9	2.3	1.12	0.22
4	214.7	10.1	4.94	0.99
5	210.3	5.7	2.77	0.55
6	204.5	-0.1	-0.05	-0.01
7				
8	208.7	4.1	1.99	0.40
9	204.8	0.2	0.10	0.02
10	214.1	9.5	4.63	0.93
11	211.8	7.2	3.52	0.70
12	188.5	-16.1	-7.87	-1.57
13				

Table 5 Proficiency test results for 200 ppm CO

Participant	Submitted result (ppm)	Difference from AV (ppm)	% Difference from AV	Z score
1	201.8	-1.9	-0.95	-0.32
2	193.0	-10.7	-5.25	-1.75
3	198.6	-5.1	-2.50	-0.83
4	205.3	1.6	0.79	0.26
5	204.1	0.4	0.20	0.07
6	197.8	-5.9	-2.90	-0.97
7				
8	198.8	-4.9	-2.43	-0.81
9	202.7	-1.0	-0.49	-0.16
10	199.1	-4.6	-2.25	-0.75
11	202.9	-0.8	-0.39	-0.13
12	200.2	-3.5	-1.70	-0.57
13	206.5	2.8	1.37	0.46

Table 6 Proficiency test results for 100 ppm NO

Participant	Submitted result (ppm)	Difference from AV (ppm)	% Difference from AV	Z score
1	102.5	-1.2	-1.19	-0.24
2	102.0	-1.7	-1.64	-0.33
3	102.8	-0.9	-0.87	-0.17
4	101.3	-2.4	-2.31	-0.46
5	96.6	-7.1	-6.84	-1.37
6	102.6	-1.1	-1.06	-0.21
7				
8	113.0	9.3	8.97	1.79
9	102.2	-1.5	-1.45	-0.29
10	103.9	0.2	0.20	0.04
11	102.5	-1.2	-1.16	-0.23
12	102.3	-1.4	-1.35	-0.27
13	103.7	0.0	0.00	0.00

Table 7 Proficiency test results for 3% O_2

Participant	Submitted result (%)	Difference from AV (%)	% Difference from AV	Z score
1	3.05	0.08	2.69	1.35
2	3.00	0.03	1.01	0.51
3	2.98	0.01	0.34	0.17
4	2.94	-0.03	-1.08	-0.54
5	2.98	0.01	0.34	0.17
6	2.97	0.00	0.00	0.00
7				
8	2.96	-0.01	-0.27	-0.13
9	2.99	0.02	0.67	0.34
10	3.00	0.03	1.01	0.51
11	2.91	-0.06	-2.02	-1.01
12	2.97	0.00	0.00	0.00
13	2.90	-0.07	-2.36	-1.18

Table 8 Proficiency test results for 40 ppm C_3H_8

Participant	Submitted result (ppm)	Difference from AV (ppm)	% Difference from AV	Z score
1	40.7	-0.4	-0.92	-0.12
2				
3	40.2	-0.9	-2.26	-0.30
4	53.2	12.1	29.44	3.93
5	42.6	1.5	3.72	0.50
6				
7	44.6	3.5	8.52	1.14
8	41.4	0.3	0.73	0.10
9	40.2	-1.0	-2.31	-0.31
10	38.9	-2.2	-5.35	-0.71
11	45.3	4.2	10.22	1.36
12	40.4	-0.7	-1.70	-0.23
13	42.9	1.8	4.38	0.58

Figures 1 to 5 illustrate the % difference between the assigned value and the submitted result with error bars showing the measurement uncertainty attributed by each participant. Figure 6 summarises the z scores for all participants and species.

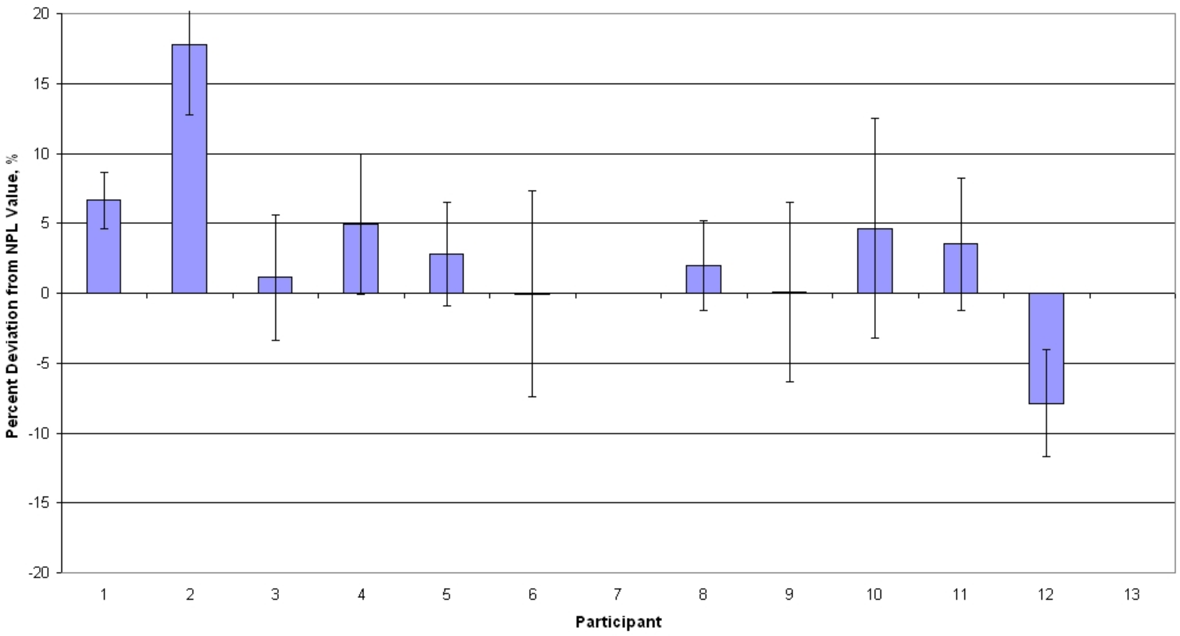

Figure 1
200 ppm SO_2 percentage difference from assigned value and uncertainty error bars attributed by participants

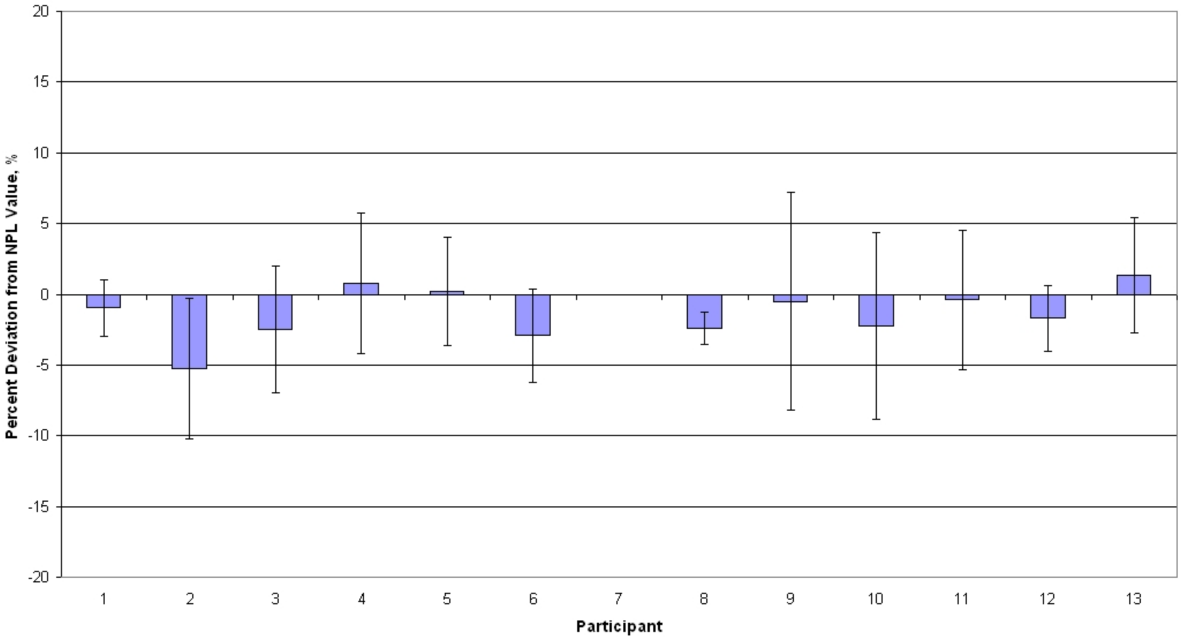

Figure 2
200 ppm CO percentage difference from assigned value and uncertainty error bars attributed by participants

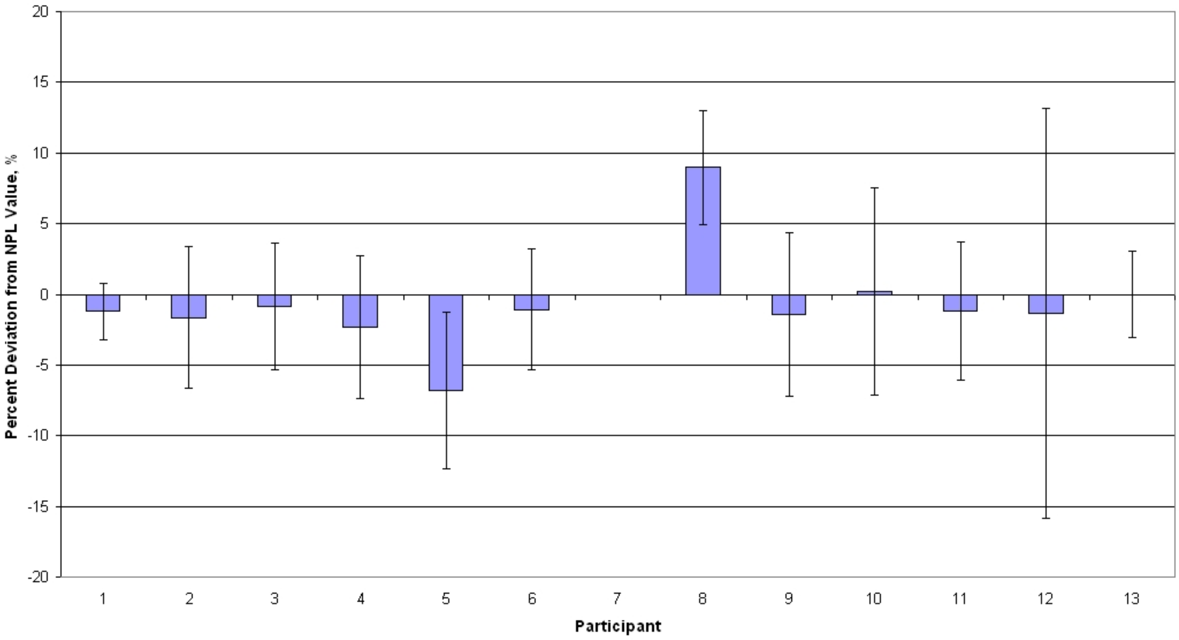

Figure 3
100 ppm NO percentage difference from assigned value and uncertainty error bars attributed by participants

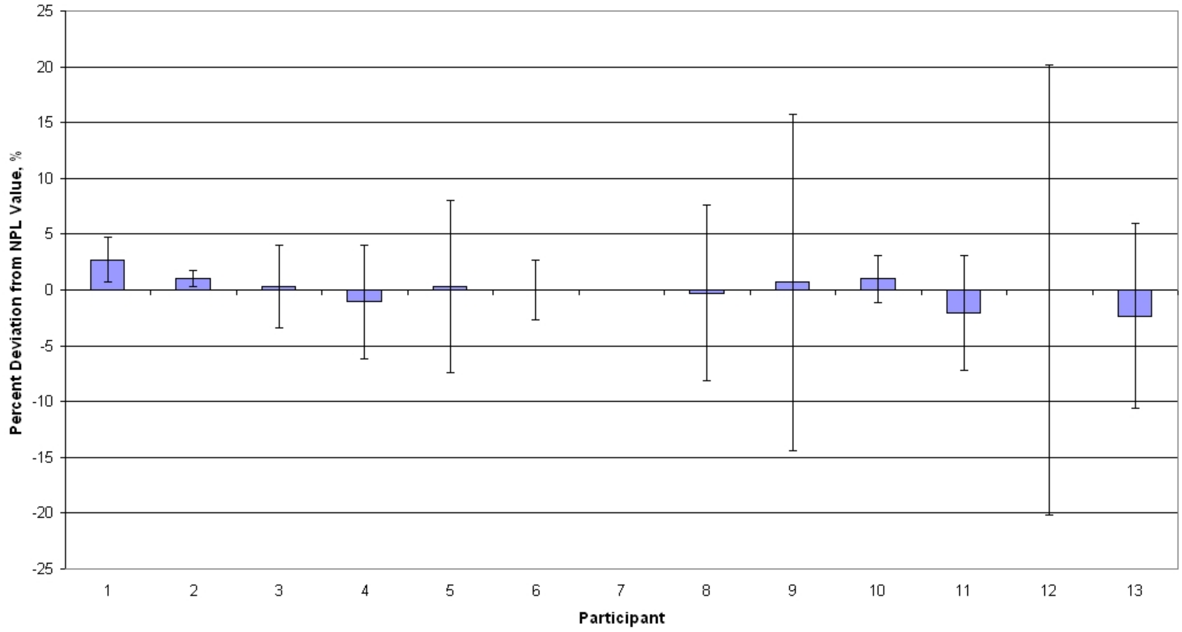

Figure 4
3 % O_2 percentage difference from assigned value and uncertainty error bars attributed by participants

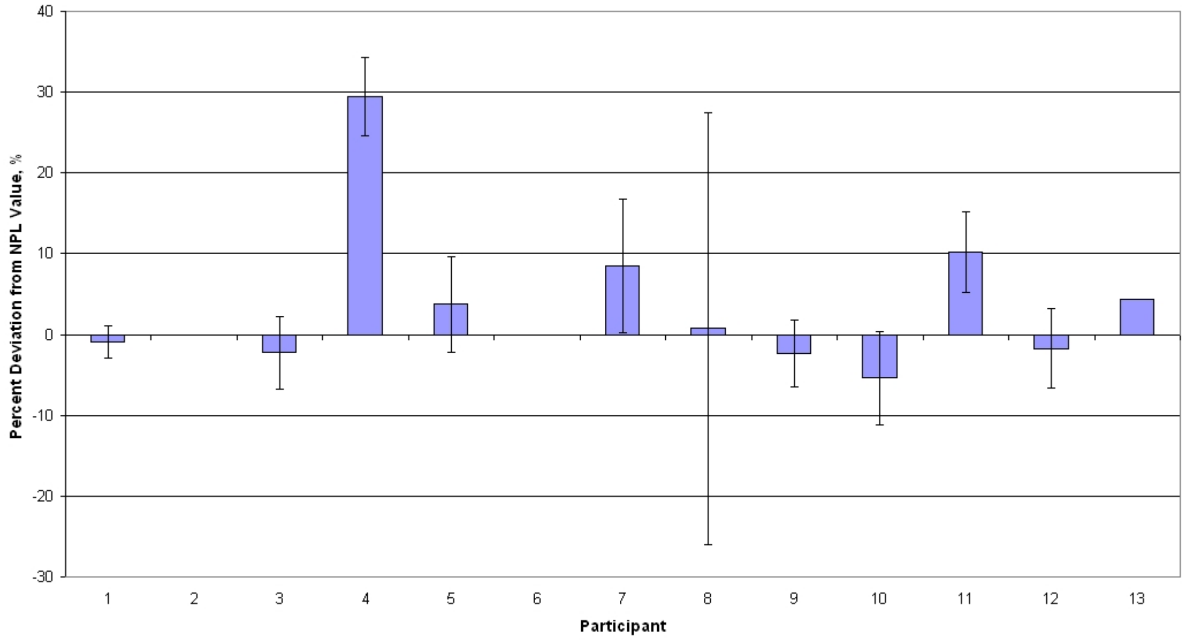

Figure 5
40 ppm C_3H_8 percentage difference from assigned value and uncertainty error bars attributed by participants

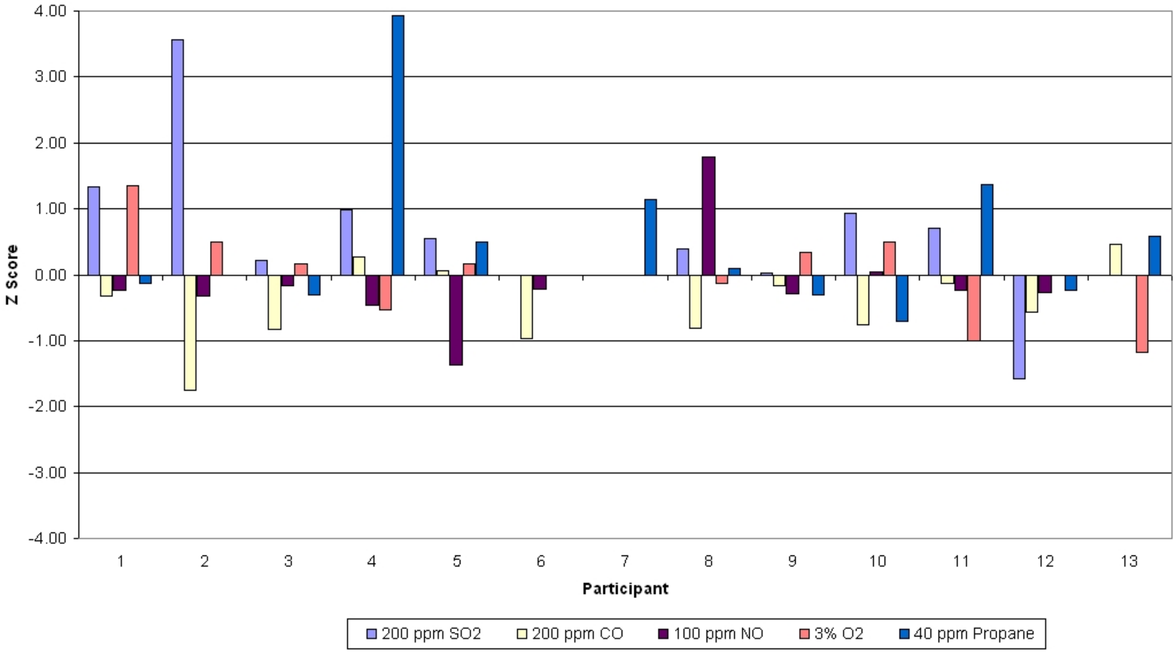

Figure 6
z scores for all species and participants

4 DISCUSSION

4.1 SO_2

Measurements of the SO_2 standard were consistently over read by all bar one participant. This is unexpected as SO_2 is a reactive species and stabilisation times are long due to the time taken to 'quench' sample lines and analyser components. It became apparent during the testing of the SO_2 cylinder that the three eight minute readings with the last four minutes averaged was not long enough for all participants' analysers to give a stable reading. It was decided that participants should submit the average of four minutes of stable readings, this was achieved within ten to fifteen minutes for most participants.

On average all participants over read the cylinder by 3.24%. Participant 2 over read the cylinder by approximately 18%; if this outlier is removed the average percentage difference is less than 2% which shows good agreement with the assigned value. All participants achieved acceptable z scores except participant 2.

4.2 CO

Measurements of the CO standard were under read by on average 1.38%. Overall there was good agreement with the assigned value and no significant systematic biases. All participants achieved acceptable z scores.

4.3 NO

Measurements of the NO standard were under read by on average 0.72%. Overall there was good agreement with the assigned value and no significant systematic biases. All participants achieved acceptable z scores.

4.4 O_2

Measurements of the O_2 standard were excellent with an average percentage difference from the assigned value of 0.03%. There was excellent agreement with the assigned value and no significant systematic biases. All participants achieved acceptable z scores.

4.5 C_3H_8

Measurements of the C_3H_8 standard were over read by on average 4.04%. All participants except one achieved acceptable z scores. Measurements of the standard were the most variable of all the standards on offer to participants and this is indicative of the more challenging nature of VOC measurements.

4.6 UNCERTAINTIES

The participants' reported uncertainties are variable and show the difference in confidence participants have in their measurements. This is unusual as the majority of participants were using the same instrumentation so you would expect participants to report broadly the same uncertainties. Those participants who report large uncertainties but are close to the assigned value might consider investigating why their uncertainty budgets are large. These participants may be able to reduce their uncertainties increasing confidence in their results.

Participants who report 'acceptable' z scores but whose uncertainties do not overlap the assigned value might consider investigating to ascertain if they are underestimating elements of their uncertainty budget.