NPL REPORT AS 57

Gas measurement proficiency testing scheme, Round 2.

M B Williams

JANUARY 2011

Gas measurement proficiency testing scheme, Round 2

Matthew Williams
Operations Directorate

ABSTRACT
This report presents the results of the second round of the gas measurement proficiency testing scheme operated by NPL which was completed in October 2010.

ISSN 1754-2928

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

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

CONTENTS

1 INTRODUCTION

This report describes the results of the second round of a gas measurement proficiency-testing scheme carried out by the National Physical Laboratory.

The gas measurement proficiency testing (PT) scheme provides a way of assessing the performance of laboratories by a series of regular inter-laboratory comparisons. The scheme was conducted at the stack simulator facility located at the National Physical Laboratory from the 19th to 28th October 2010. The results have 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

Eleven stack testing laboratories took part in the second round of the scheme, which involved measurement of combustion gases in five test conditions.

The PT scheme assessed the measurement of five test gases (NO_x, SO_2, CO, O_2 and VOC) at five simulated stack conditions. For each test, each sample gas has a known assigned value, determined by NPL. Up to four test teams operated on the stack simulator at one time and assigned values are specific to the day tested.

To determine the NO_x, SO_2, CO and O_2 assigned values a Horiba PG250 was run from a Baldwin Testerschoice gas conditioner. For VOCs a Sick Maihak FID was used. Both instruments use the standard reference methods (SRM) except for SO_2. However for SO_2 the Horiba PG250 has been proven equivalent to the SRM (NPL Report AS 26 – Validation of an Alternative Method for the measurement of SO_2 emissions using instrumental methods). An ABB 9100 FTIR was used to monitor simulator stability. . In addition check gases were introduced into the analyser between test runs to control drift.

Based on a theoretical dilution in the stack simulator the nominal concentrations for each test are shown in table 1.

Table 1. Stack simulator nominal test concentrations.

Test	NO_x (ppm)	SO_2 (ppm)	CO (ppm)	VOC (ppm)	O_2 (%)	H_2O (%)
1	280	400	100	10	12	10
2	28	40	50	5	12	10
3	140	200	10	1	8	10
4	14	20	5	0.5	6	10
5	0	0	0	0	12	10

2.2 DATA ANALYSIS

There are a number of ways in which the results of PT schemes can be interpreted. The most straightforward technique is to examine the percentage differences of the reported results from the true value, as has been applied above to this PT scheme. It is then left to the participants to gauge 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 either by other participants or by the same participant on different rounds of the scheme. If a suitable value of σ is chosen 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. 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 as soon as possible.

The target standard deviation is usually taken to be a value, which is fit for purpose for the measurements being made. The value for σ in the calculation of the z score was derived with reference to LCPD, WID and the Environment Agency's M21 (SO_2 Instrumental Method). These provide requirements on allowable uncertainties of the measurement of NO_x, SO_2, CO and VOCs as a percentage of an ELV. The relevant CEN standard methods also give guidance on the allowable uncertainties as a percentage of emission limit values (ELVs). A nominal ELV for the stack simulator was chosen at each test concentration for the calculation of the required sigma; see table 2.

Table 2. Nominal ELV and standard deviation for proficiency testing at each test condition

Gas	Unit	Uncertainty requirement (95% expanded)	Sigma (% of measured)	Test 1 ELV	Test 1 sigma	Test 2 ELV	Test 2 sigma	Test 3 ELV	Test 3 sigma	Test 4 ELV	Test 4 sigma	Test 5 ELV	Test 5 sigma
SO_2	ppm	10	5	400	20	40	5	200	10	20	5	0	5
CO	ppm	6	3	100	3	50	1.5	10	0.75	5	0.75	0	0.75
NO_x	ppm	10	5	280	14	28	3.5	140	7	14	3.5	0	3.5
VOC	ppm	15	7.5	10	0.8	5	0.6	2	0.2	1	0.2	0	0.2
O_2	%	6	3	13	0.4	13	0.4	10	0.4	8	0.3	13	0.4

However, if the standard uncertainty of the assigned value (σ_{av}) is greater than 0.3 times the standard deviation for proficiency testing (σ_{pt}) then a corrected value of sigma must be used which takes into account the standard uncertainty of the reference measurement. In this case equation 2 is used for a calculation of a modified standard deviation for proficiency testing (σ'). See table 9 for standard deviation for proficiency testing used in calculation of z scores.

$$\sigma' = (\sigma_{pt}^2+\sigma_{av}^2)^{0.5} \qquad \text{Equation 2}$$

3 DESCRIPTION OF STACK SIMULATOR FACILITY

The NPL stack simulator is able to reproduce a wide range of simulated stack gases under controlled conditions. The simulator is a recirculating system, which recreates a cross section of a 1.5 m duct. Four standard 5 inch ports are available for sampling probes or cross stack instruments to be installed. In addition a number of extractive gas analysers may be connected to gas extraction ports. Figure 1 shows a schematic of the stack simulator.

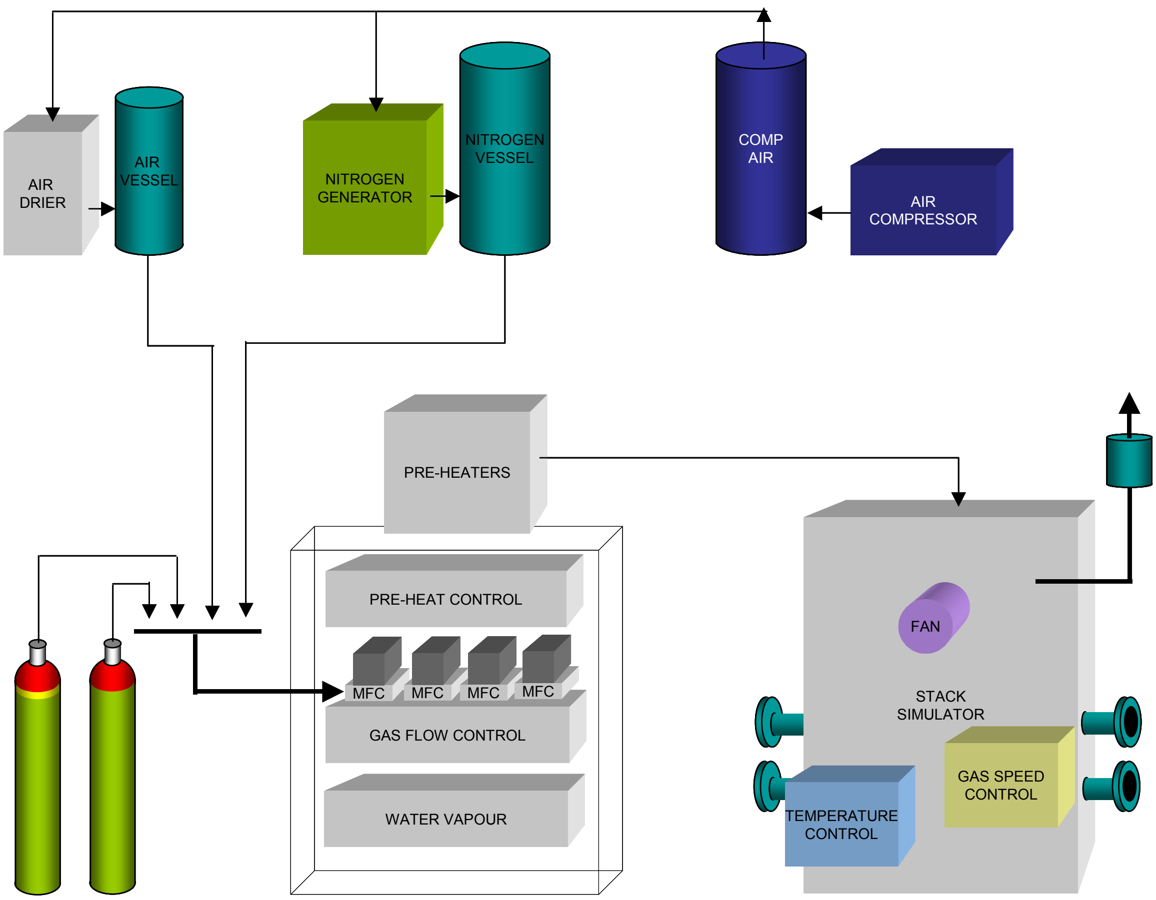

Figure 1. Schematic of stack simulator

The range of stack simulator conditions is shown in Table 3.

Table 3. Range of stack conditions achievable in the NPL Stack Simulator.

Performance Summary	
Temperature	up to 180 °C controlled to ± 2 °C measured to ± 0.5 °C
Gas vertical velocity	up to 12 metres per second
Water vapour	up to 25% by volume
Oxygen content	up to 20% by volume
Input gas flow rate	up to 300 litres per minute

The stack simulator is able to generate a mixture of test gases by blending test gases from cylinders with nitrogen and zero air from reservoirs. The bulk gas delivery is controlled by two 150 l/min mass flow controllers, allowing a blend of nitrogen and zero air. This enables reduced oxygen concentrations to be generated, simulating combustion gas conditions. Two further mass flow controllers allow the introduction of test gases, typically these consist of gas mixtures representative of combustion or waste incineration processes, i.e. SO_2 CO, NO_x. Complex gas mixtures are generated by using source gas cylinders containing a mixture of gases. The stack simulator has two low flow (~1 l/min) mass flow controllers in usual operation, though additional gas generation facilities can be added for more complex matrices.

The homogeneity of the stack simulator has been characterised according to BS EN 15259 and has been found to be homogenous across the two sample planes. A Testo 350XL was used as the stationary reference measurement and Horiba PG250 coupled to Testerschoice gas conditioner for the grid measurements; figure 2 and 3 show the grid measurements across the two sampling planes and peripheral measurements of flow and temperature.

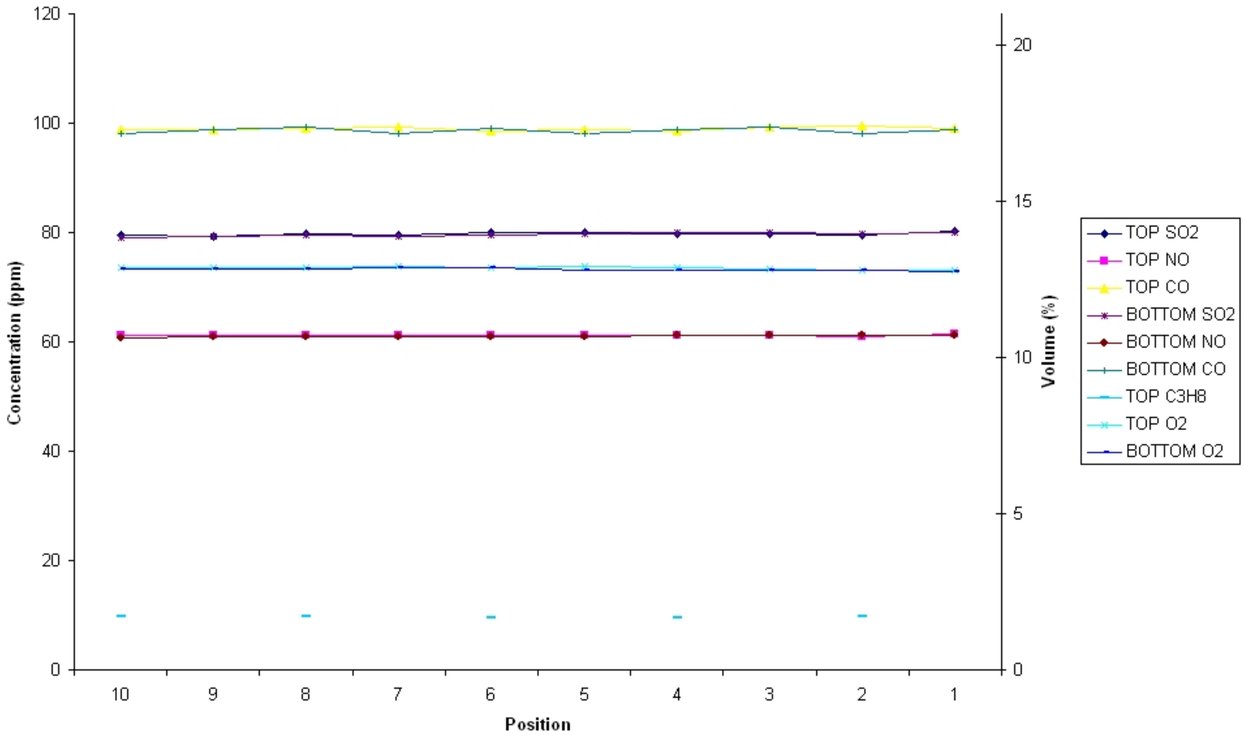

Figure 2. Grid measurements of SO_2, NO, CO, O_2 and C_3H_8 across the top and bottom sampling planes of the stack simulator. Only a limited number of points could be measured for VOCs but as same injection port is used it is expected to be homogenous.

4 RESULTS

The following tables and charts illustrate the results for each participant. Tables 4, 5, 6, 7 and 8 give the assigned values and expanded uncertainties (95% confidence; k=2) for each test against which participants were compared.

Participants were requested to submit NO_x, SO_2 and CO data as ppm on a dry gas basis, uncorrected for oxygen. O_2 data as percentage volume on a dry gas basis. VOC data to be expressed as ppm propane equivalent on a wet gas basis, uncorrected for oxygen. Participants were requested to provide a measurement uncertainty for each result.

Table 4. Assigned values and their uncertainties from 19th October 2010

Test	SO_2 (ppm)	SO_2 uncertainty (ppm)	CO (ppm)	CO uncertainty (ppm)	NO_x (ppm)	NO_x uncertainty (ppm)	VOC (ppm)	VOC uncertainty (ppm)	O_2 (%)	O2 uncertainty (%)
1	370.9	17.0	89.2	8.9	260.4	12.7	9.0	1.1	12.8	0.45
2	38.7	2.1	51.2	5.1	25.5	1.9	5.2	1.1	13.0	0.45
3	183.3	8.4	11.5	1.2	131.4	6.5	1.7	1.1	10.1	0.41
4	17.7	1.4	6.6	0.7	12.6	1.5	1.3	1.1	8.6	0.39
5	0.0	1.1	0.0	0.7	0.0	1.4	0.7	1.1	13.0	0.45

Table 5. Assigned values and their uncertainties from 20th October 2010

Test	SO_2 (ppm)	SO_2 uncertainty (ppm)	CO (ppm)	CO uncertainty (ppm)	NO_x (ppm)	NO_x uncertainty (ppm)	VOC (ppm)	VOC uncertainty (ppm)	O_2 (%)	O2 uncertainty (%)
1	338.2	15.6	82.9	2	232.7	11.4	8.3	1.1	13.4	0.46
2	36.2	2.1	46.8	1.2	23.3	1.8	5.0	1.1	13.6	0.46
3	159.5	7.4	9.8	0.5	113.3	5.7	1.6	1.1	11.5	0.43
4	15.9	1.6	5.5	0.4	11.1	1.5	1.2	1.1	10.5	0.41
5	0.0	1.5	0.0	0.4	0.0	1.4	0.7	1.1	13.6	0.46

Table 6. Assigned values and their uncertainties from 26th October 2010

Test	SO_2 (ppm)	SO_2 uncertainty (ppm)	CO (ppm)	CO uncertainty (ppm)	NO_x (ppm)	NO_x uncertainty (ppm)	VOC (ppm)	VOC uncertainty (ppm)	O_2 (%)	O2 uncertainty (%)
1	358.6	17.5	94.1	2.2	246.5	12.1	8.7	1.1	12.8	0.45
2	38.6	3.2	52.4	1.3	24.4	1.8	5.2	1.1	12.9	0.45
3	176.7	9	11.2	0.5	124.3	6.2	1.7	1.1	10.2	0.41
4	17.1	2.7	6.3	0.4	11.9	1.5	1.3	1.1	8.8	0.4
5	0.0	2.5	0.0	0.4	0.0	1.4	0.7	1.1	12.9	0.45

Table 7. Assigned values and their uncertainties from 27th October 2010

Test	SO_2 (ppm)	SO_2 uncertainty (ppm)	CO (ppm)	CO uncertainty (ppm)	NO_x (ppm)	NO_x uncertainty (ppm)	VOC (ppm)	VOC uncertainty (ppm)	O_2 (%)	O2 uncertainty (%)
1	357.5	16.2	88.8	2.5	246.3	12.1	8.7	1.1	12.9	0.45
2	39.9	2.0	50.3	1.7	25.9	1.9	5.2	1.1	13.0	0.45
3	181.4	8.2	12.2	1.3	125.2	6.3	1.8	1.1	10.1	0.41
4	18.8	1.3	7.6	1.3	13.2	1.5	1.3	1.1	8.5	0.39
5	0.0	1.1	0.0	1.2	0.0	1.4	0.7	1.1	13.0	0.45

Table 8. Assigned values and their uncertainties from 28th October 2010

Test	SO_2 (ppm)	SO_2 uncertainty (ppm)	CO (ppm)	CO uncertainty (ppm)	NO_x (ppm)	NO_x uncertainty (ppm)	VOC (ppm)	VOC uncertainty (ppm)	O_2 (%)	O2 uncertainty (%)
1	356.8	16.2	86.9	2.2	256.2	12.6	8.9	1.1	12.6	0.44
2	36.2	1.8	48.8	1.3	24.8	1.9	5.3	1.1	12.8	0.45
3	176.4	8	11.2	0.6	127.3	6.4	1.8	1.1	10.1	0.41
4	16.3	1.3	6.5	0.5	11.8	1.5	1.4	1.1	8.8	0.39
5	0.0	1.2	0.0	0.5	0.0	1.4	0.8	1.1	12.9	0.45

The uncertainties in the assigned value were significant in most cases compared to the standard deviation for proficiency testing i.e. $\sigma_{av} > 0.3 \times \sigma_{pt}$. In these cases a modified standard deviation for proficiency testing ($\sigma^{'}$) was used. Table 9 gives the standard deviation for proficiency testing used for calculation of z scores.

Table 9. Standard deviation for proficiency testing

Date	Test	SO_2 (ppm)	CO (ppm)	NO_x (ppm)	VOC (ppm)	O_2 (%)
19/10/2010	1	26.2	9.4	18.9	1.4	0.6
	2	5.4	5.3	4.0	1.3	0.6
	3	13.1	1.4	9.6	1.1	0.6
	4	5.0	1.0	3.8	1.1	0.5
	5	5.0	1.0	3.8	1.1	0.6
20/10/2010	1	25.4	3.6	18.1	1.4	0.6
	2	5.4	1.9	3.9	1.3	0.6
	3	12.4	0.9	9.0	1.1	0.6
	4	5.2	0.9	3.8	1.1	0.5
	5	5.2	0.9	3.8	1.1	0.6
26/10/2010	1	26.6	3.7	18.5	1.4	0.6
	2	5.9	2.0	3.9	1.3	0.6
	3	13.5	0.9	9.4	1.1	0.6
	4	5.7	0.9	3.8	1.1	0.5
	5	5.6	0.9	3.8	1.1	0.6
27/10/2010	1	25.7	3.9	18.5	1.4	0.6
	2	5.4	2.3	4.0	1.3	0.6
	3	12.9	1.5	9.4	1.1	0.6
	4	5.0	1.5	3.8	1.1	0.5
	5	5.0	1.4	3.8	1.1	0.6
28/10/2010	1	25.7	3.7	18.8	1.4	0.6
	2	5.3	2.0	4.0	1.3	0.6
	3	12.8	1.0	9.5	1.1	0.6
	4	5.0	0.9	3.8	1.1	0.5
	5	5.0	0.9	3.8	1.1	0.6

Table 10, 11, 12, 13 and 14 give the z scores for each participant grouped by measurand.

Table 10. SO_2 z scores

Participant	A	B	C	D	E	F	G	H	I	J	K
Test 1	0.23	-0.34	N/A	-3.18	0.66	-0.19	0.22	2.23	-0.68	0.23	1.80
Test 2	-1.86	-0.18	N/A	-3.89	0.35	-0.47	-0.69	-1.76	-2.03	-0.92	0.29
Test 3	0.13	-0.28	N/A	0.66	0.91	0.10	0.51	1.30	-0.73	0.44	1.16
Test 4	-0.79	0.13	N/A	0.05	0.11	0.09	-0.18	-1.37	-0.77	-0.39	-0.38
Test 5	0.41	0.38	N/A	0.21	-0.11	0.63	0.00	0.00	0.80	0.02	-0.08

Table 11. CO z scores

Participant	A	B	C	D	E	F	G	H	I	J	K
Test 1	-0.86	-0.07	0.15	-0.82	-0.21	-1.69	-0.34	-0.63	-0.54	-1.97	-0.14
Test 2	-0.94	-0.30	1.21	-1.18	-0.66	-0.47	-0.63	-0.53	-0.77	-0.99	0.07
Test 3	-0.25	0.88	0.44	0.44	-0.34	0.10	-0.35	-0.71	-0.94	-0.88	-0.82
Test 4	-0.14	0.96	0.96	0.61	0.02	0.09	0.59	-0.70	-0.91	-0.51	-0.73
Test 5	1.33	1.18	0.00	1.53	-0.59	0.63	0.71	-0.08	0.00	0.07	-0.11

Table 12. NO_x z scores

Participant	A	B	C	D	E	F	G	H	I	J	K
Test 1	-1.19	-0.12	0.29	0.15	1.74	-0.35	0.13	0.46	-0.18	-0.48	-0.54
Test 2	-0.30	-0.04	0.19	-0.12	0.39	0.21	-0.93	-0.31	-0.48	-1.11	-0.03
Test 3	-0.81	-0.33	0.00	-0.37	0.26	0.54	0.20	0.42	0.06	-0.45	0.21
Test 4	-0.14	-0.03	0.02	-0.11	-0.03	0.11	-0.81	-0.20	-0.30	-0.99	0.03
Test 5	0.27	0.02	-0.01	-0.03	-0.11	0.37	0.00	0.01	0.00	0.11	0.03

Table 13. VOC z scores

Participant	A	B	C	D	E	F	G	H	I	J	K
Test 1	0.38	0.22	0.67	0.00	0.74	-0.46	-0.02	-0.74	-0.53	-0.91	0.45
Test 2	0.29	-0.01	0.31	-0.09	1.11	-0.32	-0.16	-0.54	-0.48	-0.80	-0.32
Test 3	0.21	-0.22	0.14	-0.49	0.23	-0.60	0.21	-0.58	-0.43	-0.59	-0.63
Test 4	0.12	-0.21	0.06	-0.03	0.15	-0.62	0.10	-0.53	-0.54	-0.66	-0.60
Test 5	0.03	-0.12	-0.03	0.14	0.77	-0.64	0.07	0.02	-0.61	-0.46	-0.81

Table 14. O_2 z scores

Participant	A	B	C	D	E	F	G	H	I	J	K
Test 1	-0.20	0.08	0.04	0.36	0.12	0.04	1.70	0.17	0.05	0.75	0.18
Test 2	-0.12	0.08	0.02	0.35	-0.01	0.06	2.06	0.19	0.05	0.71	0.26
Test 3	-0.17	0.01	-0.13	0.38	-0.20	0.09	1.48	0.16	-0.24	0.58	0.35
Test 4	-0.32	-0.05	-0.18	0.21	-0.28	0.08	1.48	0.13	-0.04	0.67	0.42
Test 5	-0.09	0.09	0.04	0.37	-0.24	0.11	1.60	0.22	0.05	0.72	0.12

Figures 3, 5, 7, 9 and 11 give the z scores for each measurand grouped by test number. Figures 4, 6, 8, 10 and 12 give the ppm or % volume difference from the assigned value plus reported uncertainty bars grouped by participant. The gaps in results for some measurands are a result of participants choosing not to participate or not reporting uncertainties.

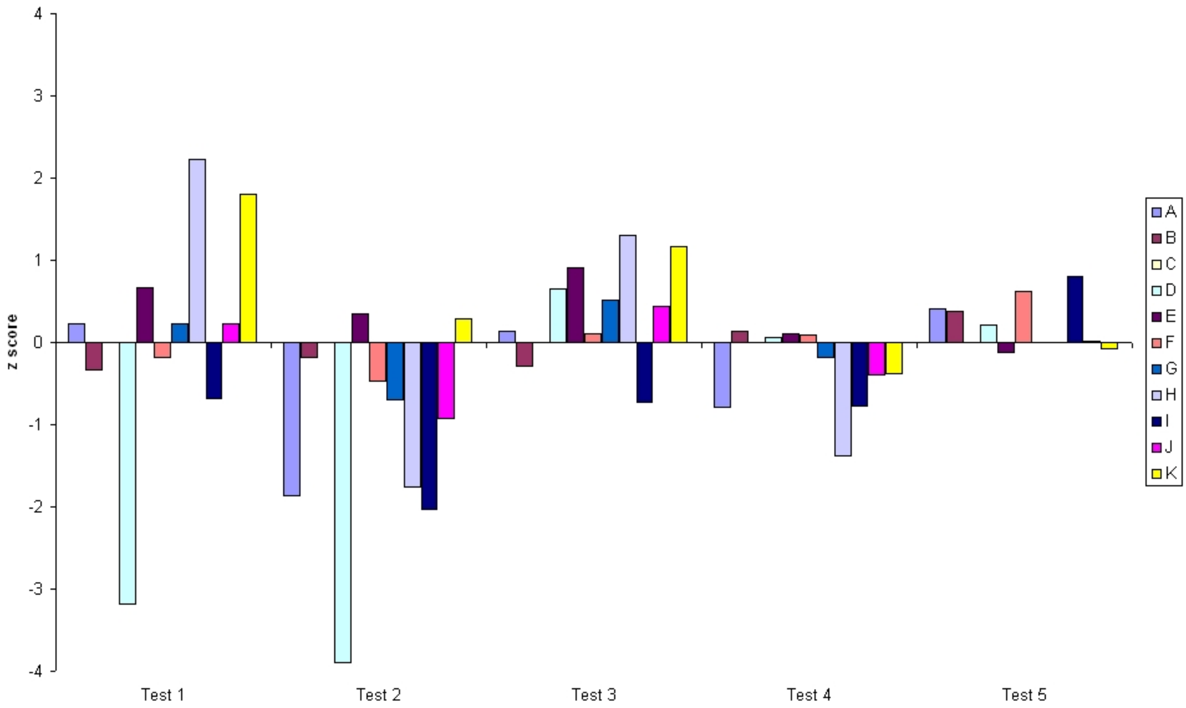

Figure 3. SO_2 z scores

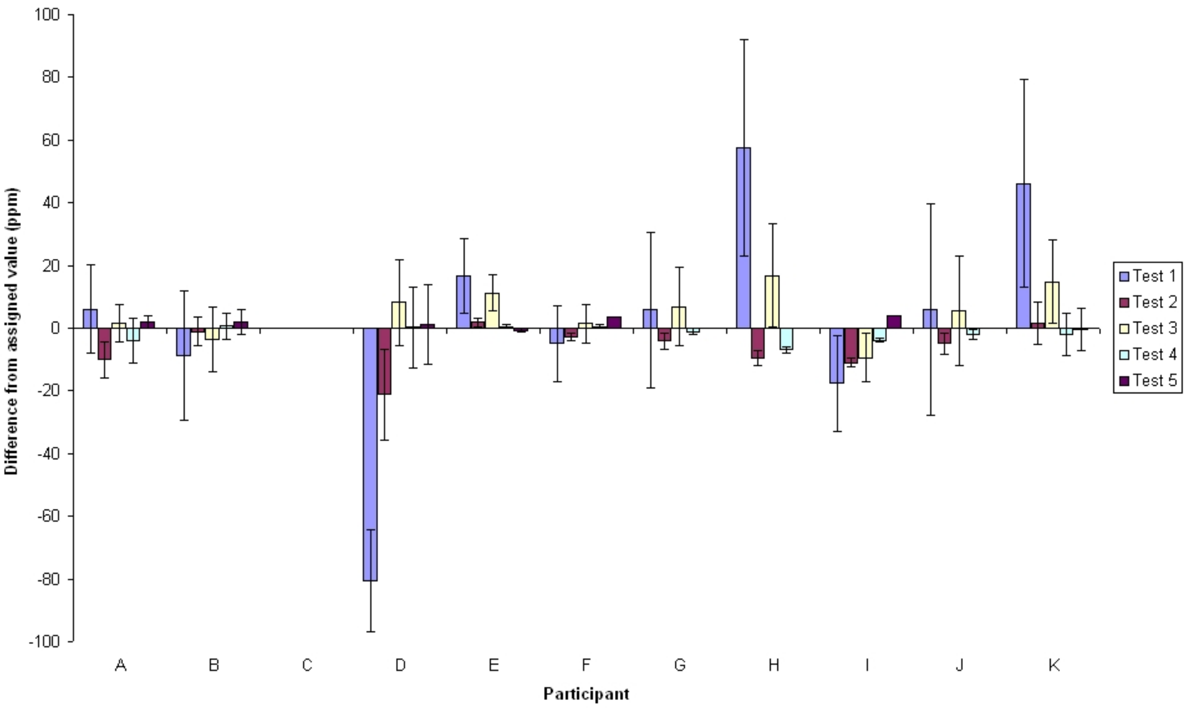

Figure 4. SO_2 ppm difference from assigned value with uncertainty bars

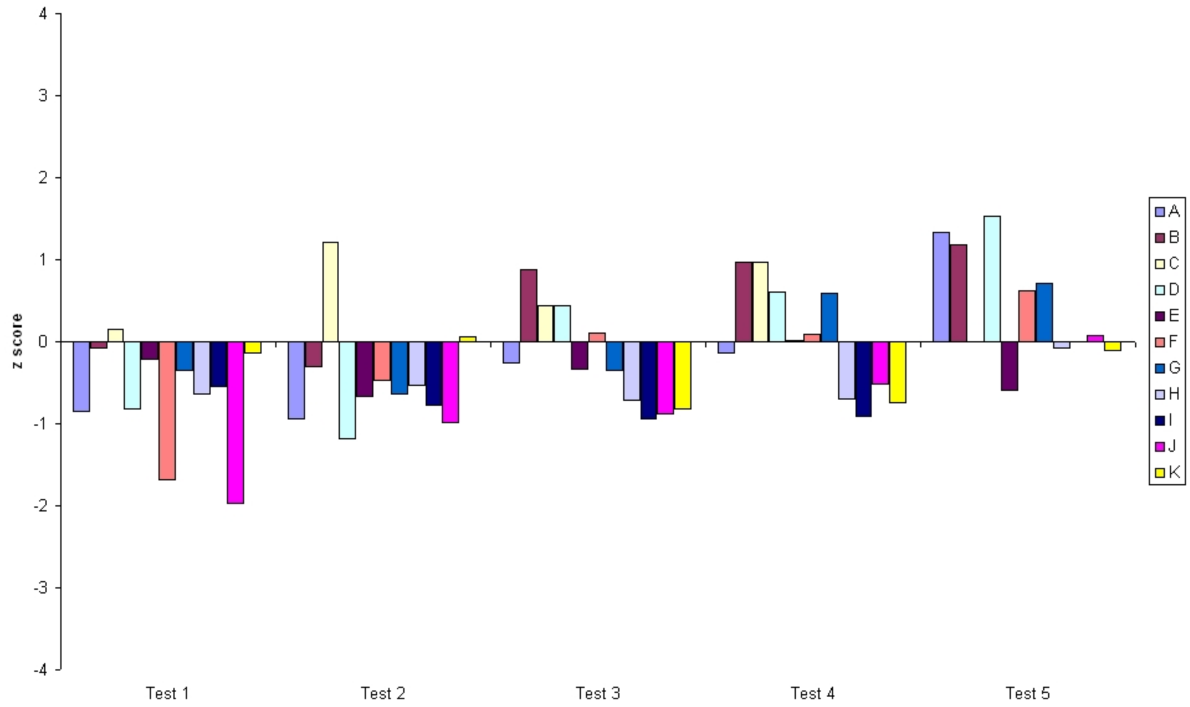

Figure 5. CO z scores

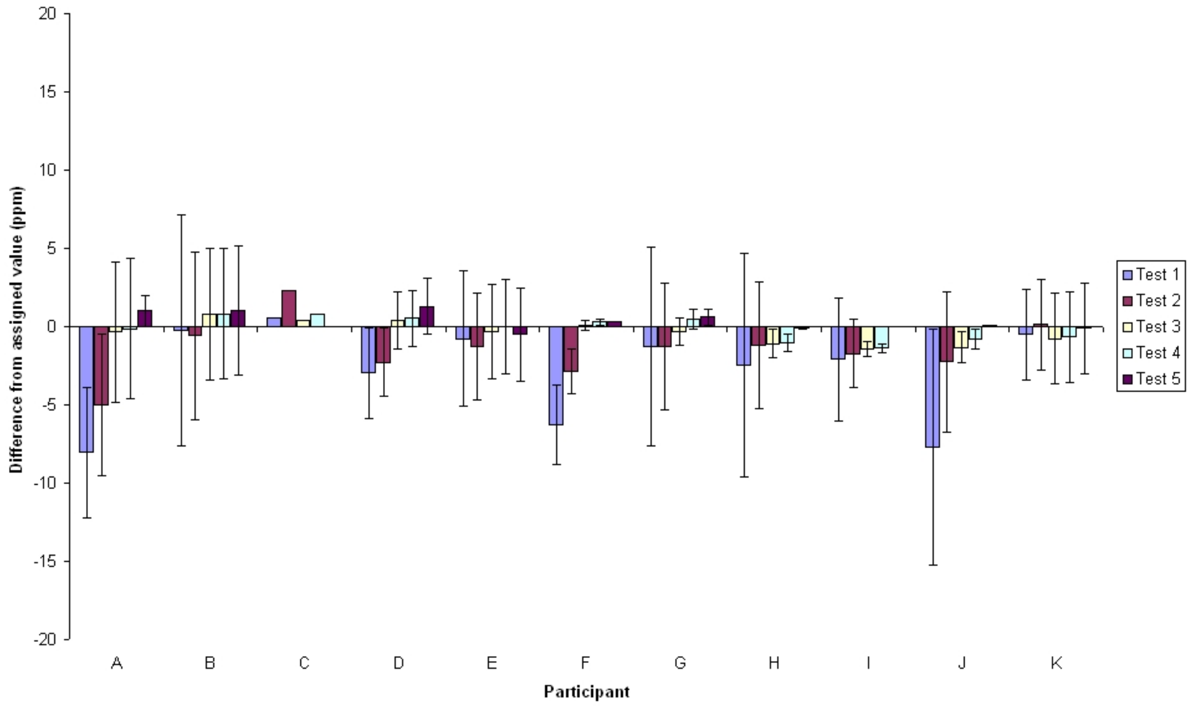

Figure 6. CO ppm difference from assigned value with uncertainty bars

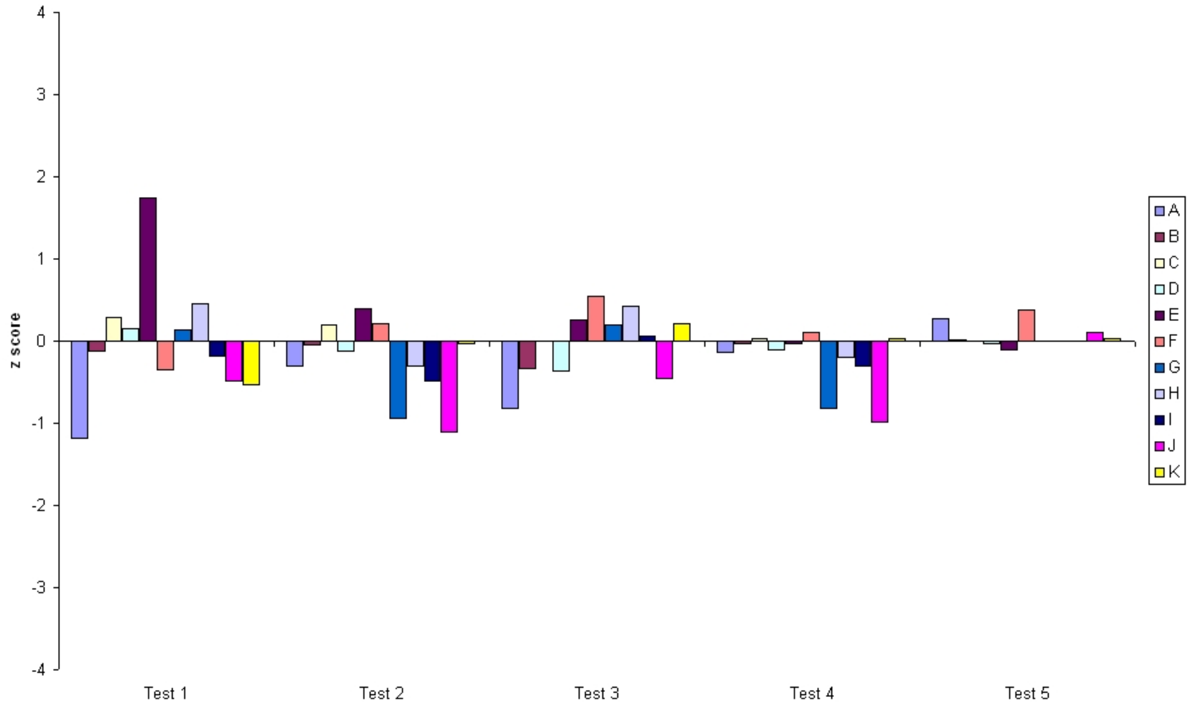

Figure 7. NO_x z scores

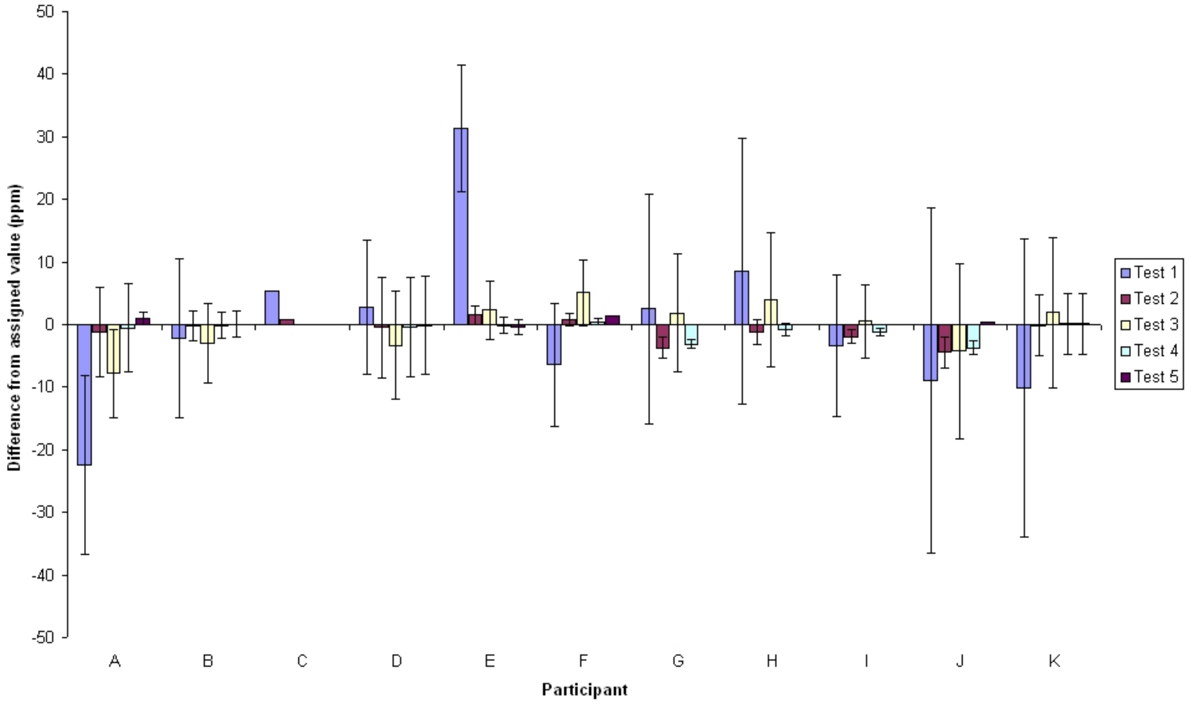

Figure 8. NO_x ppm difference from assigned value with uncertainty bars

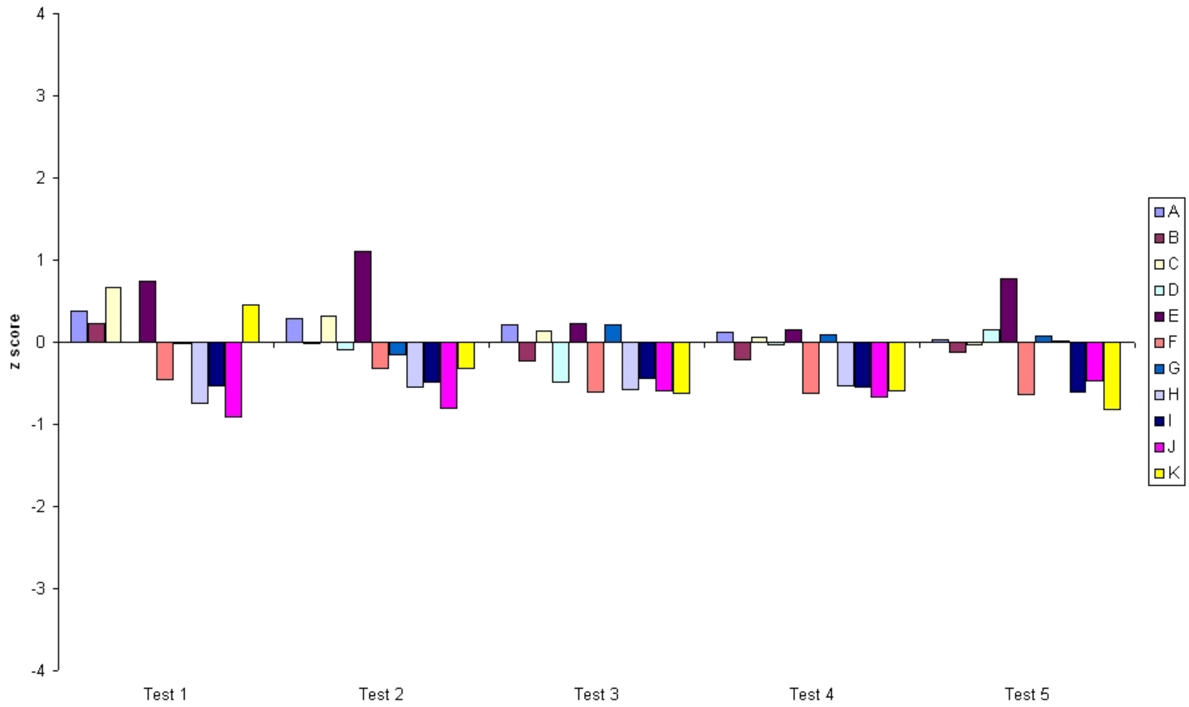

Figure 9. VOC z scores

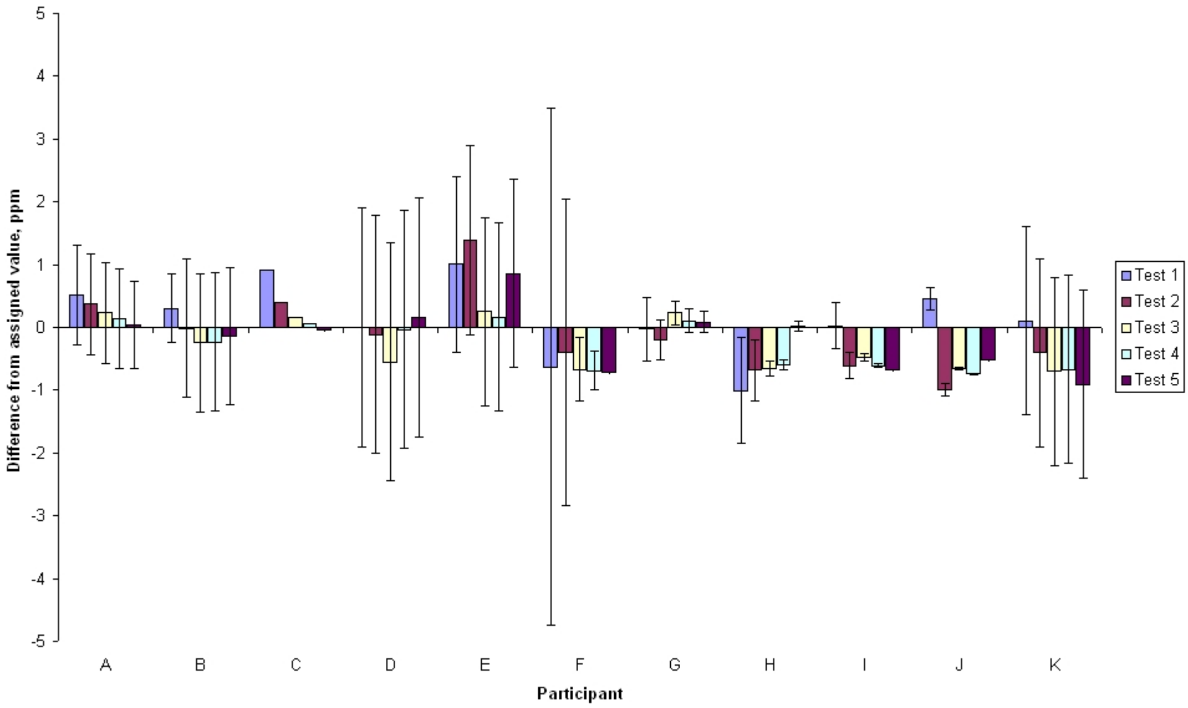

Figure 10. VOC ppm difference from assigned value with uncertainty bars

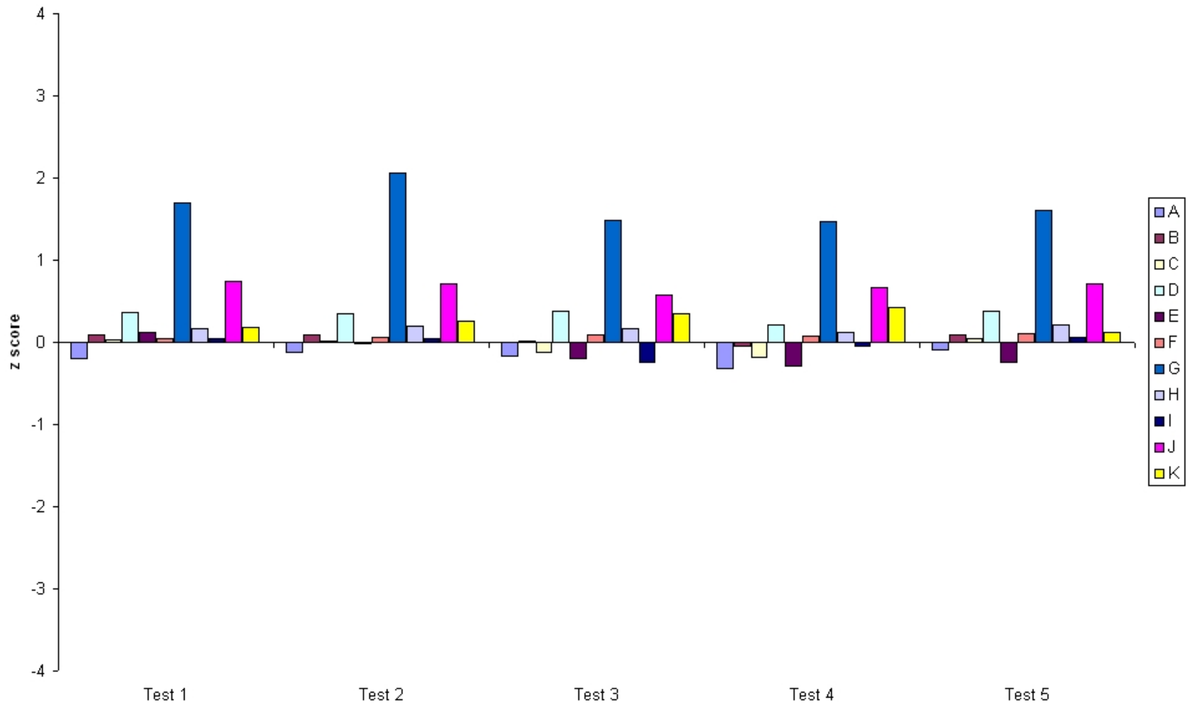

Figure 11. O_2 z scores

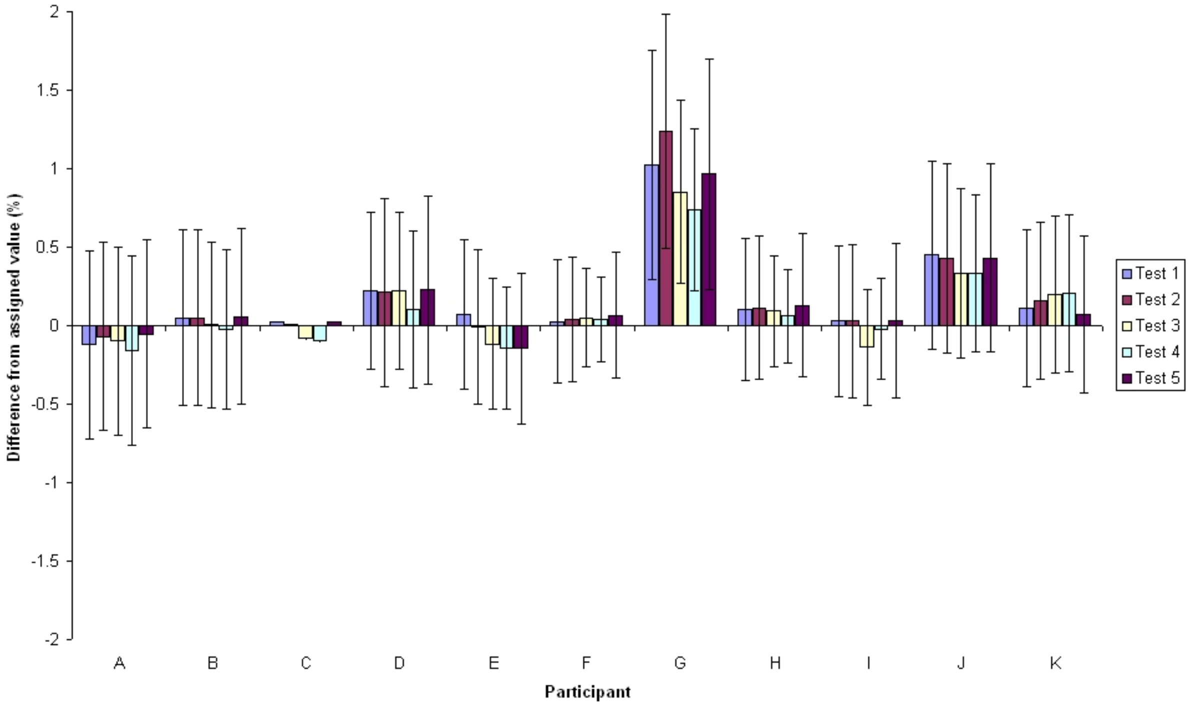

Figure 12. O_2 percentage volume difference from assigned value with uncertainty bars

5 OBSERVATIONS

Participant D obtained unsatisfactory scores for SO_2 on tests one and two but acceptable scores for tests three, four and five. This team observed that their results for test one and two were significantly different from the nominal concentration set and decided to re-calibrate the SO_2 channel on their instrument for tests three, four and five. In their submitted results it was stated that had this occurred on site then this data would have been treated as invalid.

Participants A, C and G submitted NO data as opposed to NO_x. No NO_2 was injected into the simulator as part of the proficiency test. Submitted NO data was treated as if it were NO_x. None of these participants achieved unacceptable scores.

In the processing of results, raw data supplied on the test day is compared with the submitted results from participants. During this process it was noted there was a significant difference between participant J's raw and submitted results for NO_x. This difference was highlighted to the participant who then re-submitted their results for NO_x. The error seen could clearly be attributed to data processing, specifically an error in unit conversion and not a measurement problem. The participant was not told the detail of the discrepancy only that there was a significant difference between their raw and submitted results for NO_x. The participant indicated an entry would be made in their quality system, an investigation and corrective actions would be put in place.

The instrumentation used for the assigned value was prone to drift; this has been quantified and corrected for. Any additional uncertainty introduced by these corrections will not have given incorrect z score classification.

The SO_2 channel of the instrument used for the assigned value had interference between air and nitrogen; this has been corrected for.

6 CONCLUSIONS

The results for this round of the scheme were in general good. Participant D being the only participant achieving unacceptable z scores for SO_2 in the first two tests.

There was a good spread of results above and below the assigned value for SO_2, NO_x, VOC and O_2. A small bias was seen in the CO results with participants under reading in comparison to the assigned value. However the magnitude of the bias is not significant, and has not pushed any participant's scores into the unacceptable range.

The participants' reported uncertainties are variable and show the difference in confidence participants have in their measurements. 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.

7 REFERENCES

Robinson, R A, Coleman, M, Williams, M, Elliott, R, Clack, M, Curtis, A. NPL Report AS 26 – Validation of an Alternative Method for the measurement of SO_2 emissions using instrumental methods.