The EC QA/QC programmes for inorganic gas pollutants testing

M. Barbiere[1], A. Borowiak[1], F. Lagler[1], M. Gerboles[1], M. Kapus[2] & C. Belis[1]
[1]*European Commission – Joint Research Centre, Institute for Environment and Sustainability, European Reference Laboratory for Air Pollution (ERLAP), Italy*
[2]*Danfoss d.o.o., Slovenia*

Abstract

The Air Quality Directive (2008/50/EC) asks for the organisation of quality assurance programmes for air quality assessment methods at European level. Since the early 1990s the European Reference Laboratory for Air Pollution (ERLAP) of the EC's Joint Research Centre (JRC) has carried out Inter-comparison Exercises (IE) for air pollution measurements on a regular basis for Member States of the EU. All European National Reference Laboratories (NRLs), joined together in the AQUILA Network, are obliged to participate in IE. More than 45 laboratories and institutes, coming from 35 European countries, have participated in the IE during the last 15 years. The results of the most recent IE which took place from 2005 to 2010 are described. Gas mixtures with some concentrations of CO, SO_2, NOx, and O_3 were generated and measured by the participants. With the results of the participants' z'-score, En number, repeatability and reproducibility, outlier through the test of Grubb were evaluated.

Keywords: air quality, intercomparison, ozone, sulphur dioxide, nitrogen monoxide, nitrogen dioxide, carbon monoxide.

1 Introduction

With the adoption of Directive 2008/50/EC [1] on ambient air quality and cleaner air for Europe, a framework for a harmonized air quality assessment in Europe was set. This Directive specifies, among others, the reference methods

WIT Transactions on Ecology and the Environment, Vol 147, © 2011 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/AIR110171

for measuring specific atmospheric pollutants and sets data quality objectives (DQO) for the uncertainty, minimum data capture and time coverage. It establishes limit and target values for sulphur dioxide (SO_2), nitrogen dioxide (NO_2) and nitrogen oxide (NO), particulate matter, lead, benzene, carbon monoxide (CO) and ozone (O_3), not to be exceeded to avoid negative effects on human health and the environment.

The European Commission (EC) has supported the development and publication of standard measurement methods for CO [2], SO_2 [3] , NO-NO_2 [4] and O_3 [5] as European standards. Appropriate calibration methods [6–8] have been standardized by the International Organization for Standardization (ISO).

As foreseen in the Air Quality Directive the ERLAP organizes inter-laboratory comparison exercises (IE) to assess and improve the status of comparability of measurements of National Reference Laboratories (NRL) of the Member States of the European Union. Through the IE ERLAP promotes information and know how exchange among the expert laboratories. Currently, a more systematic approach has been adopted, in accordance with the Network of NRLs for Air Quality (AQUILA) [9], aiming both at providing an alert mechanism for the purposes of implementation of legislation and at supporting the operation of quality schemes by NRLs. The protocol for the organization of IEs was developed by ERLAP in collaboration with the WHO CC and the AQUILA Network, collecting all the experiences of the previous IE for gaseous air pollutants [10]. This evaluation scheme was adopted in December 2008 by the AQUILA Network and WHO CC and is applied to all IEs since then. It contains common criteria to alert on possible performance failures which do not rely solely on the uncertainty claimed by participants. The evaluation scheme implements the z'-score [11] and E_n method [11] with the uncertainty requirements for calibration gases stated in the European standards [2–5], which are consistent with the DQOs of European Directives. Beside the proficiency of participating laboratories, the repeatability and reproducibility [13] of standardized measurement methods [12–14] are evaluated as well. These group evaluations are useful indicators of trends in measurement quality over different IE.

2 Inter-comparison exercises

In this report the results of nine IE that were organized between 2005 and 2010 in three European facilities (Langen (D), Essen (D) and Ispra (I)) are described. In Langen the inter-comparison facility of the German Federal Environment Agency (UBA) Pilot station was used and the IE was carried out under the supervision of the World Health Organization Collaborating Centre for Air Quality Management and Air Pollution Control, Berlin (WHO CC) in collaboration with JRC. In 2007 the JRC organized an IE in Essen at the facility of the North Rhine-Westphalian State Agency for Nature, Environment and Consumer Protection (LANUV) in cooperation with WHO CC [17]. All the others IE took place at the ERLAP laboratory of the JRC in Ispra (IT). In Table 1 the list of IE evaluated is shown.

Table 1: List of IE from 2005 till 2010.

Inter-comparison	Site	N. of participants
June 2005	Ispra (IT)	11
June 2007	Ispra (IT)	11
October 2007	Essen (DE)	16
April 2008	Ispra (IT)	9
October 2008_1	Ispra (IT)	10
October 2008_2	Ispra (IT)	9
September 2009	Langen (DE)	8
October 2009	Ispra (IT)	9
June 2010	Ispra (IT)	10

The participants were required to participate in the IE with their own measurement instrumentation, data acquisition equipment and working standards to be used for calibrations during the IE.

2.1 The preparation of test mixtures

During the IE, gas mixtures were prepared for SO_2, CO, O_3, NO and NO_2 at concentration levels around European Air Quality limit and target.

The test mixtures were prepared by the dilution of gases from cylinders containing high concentrations of NO, SO_2 or CO using thermal mass flow controllers [8]. O_3 was added using an ozone generator and NO_2 was produced applying the gas phase titration method [16] in the conditions of excess NO. Several different concentrations steps were generated, each lasting roughly 2 hours.

Participants were required to report three half-hour-mean measurements for each concentration level in order to evaluate the repeatability of their measurements. Zero concentration levels were generated for one hour and one half-hour-mean measurement was reported. The sequence program of generated test gases is given in Figure 1. In order to test simultaneous gas mixtures under homogeneous experimental conditions a calibration bench [15] was used to generate the different pollutant mixtures (Figure 1).

The calibration bench allows through a dynamic dilution the generation of complex gas mixtures by dilution of high concentration gas cylinders.
The system is further equipped with an ozone generator for the implementation of the Gas Phase Titration (GPT) and with a water vapor generator for the preparation of humid gas mixtures.

All the functions of the bench are programmable and controlled by computer, so that automated and unattended operation is possible.

The gas mixture is supplied to the workbenches. During each IE its reference value is given by the ERLAP laboratory monitor who is connected to one workbench.

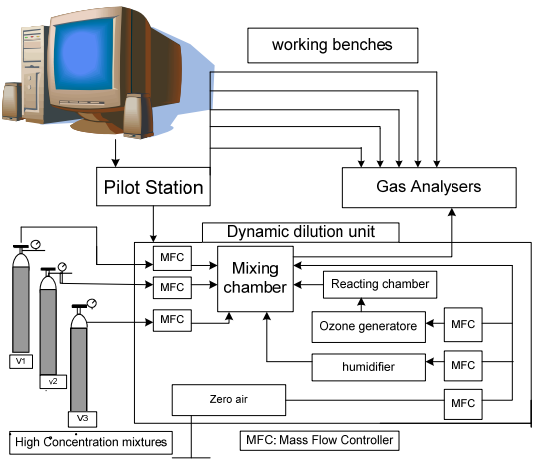

Figure 1: Calibration bench scheme.

Table 2: Example of the sequence program of generated test gases.

day	start time	duration (h)	operation	zero air ppb	NO ppb	NO_2 ppb	O_3 ppb	CO ppm	SO_2 ppb
1st	12:00	6	installation						
2nd	8:00	3	calibration						
2nd	11:00	1	NO-NO2-O3	0					
2nd	12:00	2	NO-NO_2		500	0			
2nd	14:00	2	NO-NO_2		380	120			
2nd	16:00	2	O_3				120		
2nd	18:00	2	NO-NO_2		250	0			
2nd	20:00	2	NO-NO_2		146	104			
2nd	22:00	2	O_3				104		
3rd	0:00	2	NO-NO_2		150	0			
3rd	2:00	2	NO-NO_2		90	60			
3rd	4:00	2	O_3				60		
3rd	6:00	2	NO-NO_2		50	0			
3rd	8:00	2	NO-NO_2		29.1	20.9			
3rd	10:00	2	O_3				20.9		
3rd	12:00	2	NO-NO_2		15.7	0			
3rd	14:00	2	NO-NO_2		2.1	13.6			
3rd	16:00	2	O_3				13.6		
3rd	< 18:00	2	calibration						
3rd	20:00	1	CO-SO_2	0					
3rd	21:00	2:30	CO-SO_2					8.6	132
3rd	23:30	2	CO-SO_2					6	47
4th	1:30	2	CO-SO_2					4.3	18.8
4th	3:30	2	CO-SO_2					2	7.5
4th	5:30	2	CO-SO_2					1	3
4th	7:30	1		0					
4th	8:30	END							

3 The evaluation of the laboratory's measurement proficiency

To evaluate the participants measurement proficiency the methodology described in ISO 13528 [11] was applied. For the IEs organized in Ispra (IT) and in Essen (DE) the measurement results of ERLAP were considered as the reference values according to the AQUILA protocol [10] while in Langen (DE), the results of UBA (DE) were used as the reference values.

The proficiency of the participants was assessed by calculating two performance indicators. The first performance indicator (z'-score) evaluates if the difference between the participant measured value and the reference value remains within the limits of a common criterion. The second performance indicator (E_n-number) tests if the difference between the participant measured value and reference value remains within the limits of a criterion, which is calculated individually for each participant from its declared uncertainty of measurement and the uncertainty of reference value.

3.1 Assigned values

The assigned values of tested concentration levels were derived from ERLAP or UBA–measurements which were calibrated against their certified reference material gases under strict condition of traceability to international standards. In this perspective the assigned values are reference values as defined in the ISO 13528 [11]. SO_2, CO and NO analyzers were calibrated using primary calibration gas mixtures prepared according to the methodology described in the ISO 6144 [7] (UBA) and ISO 6143 [6] (ERLAP). Gas mixtures for the calibration experiment were produced from the reference mixtures by dynamic dilution method using mass flow controllers [8]. All flows were measured with certified devices (Brooks vol-U-meter or Molbox/Molbloc systems).

In Ispra since 2008 O_3 calibration measurements during IE were carried out using as primary standard the NIST (US National Institute of Standards and Technology) Standard Reference Photometer SN42 (SRP) [18].

Assigned values were validated by comparison to the group statistics (x* and s*) for every parameter and concentration level of the IE. These statistics are calculated from participants, applying the robust method described in ISO 13528 [11]. The validation is taking into account reference laboratory measurement result (X) and its standard uncertainty ($u_{X'}$) as given in equation (1) [11]:

$$\frac{\left|x^* - X\right|}{\sqrt{\frac{\left(1{,}25 \cdot s^*\right)^2}{p} + u_{X'}^2}} < 2 \tag{1}$$

where 'x*' and 's*' represent robust average and robust standard deviation respectively and 'p' is the number of participants.

The homogeneity of test gas mixtures throughout the working bench was evaluated by comparison of measurements at the beginning and at the end of the distribution line. From the relative differences between beginning and end measurements, average and standard deviation (*s*) were calculated, and the uncertainty of test gas due to lack of homogeneity was calculated as the sum of squares of these average and standard deviation.

The upper and lower limits of bias homogeneity was evaluated to be smaller than 0.5% which constitutes the relative standard uncertainty of 0,3% of each concentration level assuming a rectangular distribution of the biases.

The standard uncertainties of reference values (u_X) were calculated with equation (2).

$$u_X^2 = u_{X'}^2 + \left(X \cdot u_{r\,\mathrm{homogeneity}}\right)^2 \tag{2}$$

3.2 z'-score

The z'-score statistic is calculated according to ISO 13528 [11] with equation (3).

$$z' = \frac{x_i - X}{\sqrt{\sigma_p^2 + u_X^2}} \tag{3}$$

where 'x_i' is a participant's run average value, 'X' is the reference value, 'σ_p' is the 'standard deviation for proficiency assessment' and 'u_X' is the standard uncertainty of assigned value. In the European standards [2–5] the uncertainties of calibration gases used in ongoing quality control are prescribed. In fact, it is stated that the maximum permitted expanded uncertainty for calibration gases shall be 5% and that 'zero gas' shall not give instrument reading higher than defined limit. The assessment of results in the z'-score evaluation is made according to the following criteria:

- $|z'| \leq 2$ are considered acceptable score.
- $2 < |z'| \leq 3$ are considered warning score.
- $|z'| > 3$ are considered not acceptable score. Scores falling in this range are very unusual and are taken as evidence that an anomaly has occurred that should be investigated and corrected.

After more than 15 years of IE the high level of expertise reached by the NRL is confirmed by the high percentage of acceptable results (above 90%).

Table 3: Percentage of z'-score results in IE from 2005 till 2010.

IE	Acceptable score %	Warning score %	Not Acceptable score %
June 2005	95.5	2.3	2.2
June 2007	97.8	1.9	0.3
October 2007	93.2	4.6	2.2
April 2008	93.8	2.1	4.1
October 2008_1	92.9	4.2	2.9
October 2008_2	97.0	3.0	0
September 2009	94.3	4.7	1
October 2009	98.2	1.8	0
June 2010	97.0	3.0	0

3.3 E_n - number

The normalized deviations [11] (E_n) were calculated with the following equation:

$$E_n = \frac{x_i - X}{\sqrt{U_{x_i}^2 + U_X^2}} \tag{4}$$

where 'X' is the reference value with an expanded uncertainty 'U_X' and 'x_i' is the participant's average value with an expanded uncertainty 'U_{Xi}'.

Results are acceptable when $|E_n| \le 1$.

3.4 Discussion about z'-score and En-number

For a general assessment of the quality of each result a decision diagram was developed (Figure 2) sorting the results according to the following seven categories.

a1 measurement result is completely acceptable
a2 measurement result is acceptable (z'-score acceptable and En-number ok) but the reported uncertainty is too high
a3 measured value is acceptable (z'-score acceptable) but the reported uncertainty is underestimated (En-number not ok)
a4 measurement result is warning (z'-score warning) but due to a high reported uncertainty can be considered valid (En-number ok)
a5 measurement result is warning (z'-score warning and En-number not ok)
a6 measurement result is not acceptable (z'-score not acceptable) but due to a high reported uncertainty can be considered valid (En-number ok)
a7 measurement result is not acceptable (z'-score not acceptable and En-number not ok).

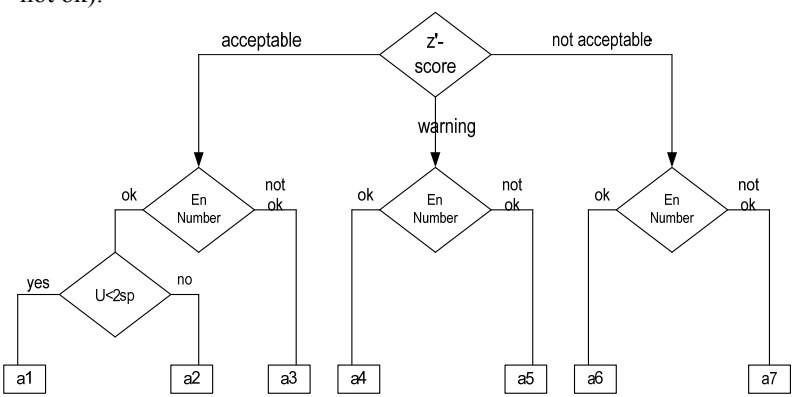

Figure 2: Diagram for general assessment of proficiency results.

Table 4: General assessment of proficiency results. "nv": no values reported.

	unit	2007_1	2007_2	2008_1	2008_2	2008_3	2009_1	2009_2	2010
a1	%	64	55	54	37	27	55	85	80
a2		30	30	14	40	31	27	6	8
a3		4	6	6	14	19	3	8	4
a4		1	1	1	1	1	4	0	1
a5		1	3	1	3	2	1	1	2
a6		0	0	2	1	0	1	0	0
a7		0	2	1	2	0	0	0	0
nv		0	3	21	2	20	9	0	5

In Table 4 is presented a summary of z'-score and En Number evaluation of all IE beside 2005 in which this discussion was not carried out. As described above a1 indicates the percentage of acceptable results as measured value and calculated uncertainty. Generally the results have good measured values (a1+a2+a3) but in 2008_2 and 2008_3 the percentage of uncertainty values too high (a2) and too low (a3) are considerably above all the other IE. Beside 2009_2 and 2010 the percentage of a2 category shows a tendency to overestimate the uncertainty. In order to investigate this issue in the future would be interesting asking the participants to provide the method used to calculate the uncertainty.

3.5 Reproducibility and repeatability

Reproducibility (R) and Repeatability (r) were determined [13] with equation respectively (5) and (6). In equation (5) p is the number of participants after discarding outliers, s_i is the standard deviation of the measurements of each participant for each sample, y_m is the mean of the measurements of each participant, m is the reference value of each sample and n is the number of repeated measurements.

$$R = 2.8\sqrt{\frac{1}{p-1}\sum_{p}(y_m - m)^2 + \frac{n-1}{n}\frac{\sum_{p} s_i^2}{p}} \tag{5}$$

$$r = 2.8\sqrt{\frac{\sum_{p} s_i^2}{p}} \tag{6}$$

From Table 5 to Table 14 repeatability and reproducibility of two levels of concentrations for NO, NO_2, O_3, CO and SO_2 are represented.

Reproducibility values are quite higher than repeatability and this could be a sign of a possible non homogeneous calibration procedure and reference material used.

Table 5: NO (conc. 500ppb) reproducibility and repeatability in all IE.

Parameter	r (ppb)	R (ppb)	IE
NO	0.006	0.036	2005
NO	0.003	0.056	2007_1
NO	0.004	0.053	2007_2
NO	0.002	0.022	2008_1
NO	0.003	0.012	2008_2
NO	0.0045	0.090	2008_3
NO	0.005	0.142	2009_1
NO	0.003	0.048	2009_2
NO	0.007	0.065	2010

Table 6: NO (conc. 50ppb) reproducibility and repeatability in all IE.

Parameter	r (ppb)	R (ppb)	IE
NO	0.0048	0.062	2007_1
NO	0.009	0.086	2007_2
NO	0.005	0.102	2008_1
NO	0.011	0.042	2008_2
NO	0.006	0.133	2008_3
NO	0.013	0.118	2009_1
NO	0.004	0.098	2009_2
NO	0.005	0.083	2010

Table 7: NO_2 (conc. 100ppb) reproducibility and repeatability in all IE.

Parameter	r (ppb)	R (ppb)	IE
NO_2	0.008	0.095	2005
NO_2	0.007	0.068	2007_1
NO_2	0.007	0.090	2007_2
NO_2	0.006	0.049	2008_1
NO_2	0.012	0.097	2008_2
NO_2	0.005	0.043	2008_3
NO_2	0.009	0.084	2009_1
NO_2	0.006	0.116	2009_2
NO_2	0.006	0.108	2010

Table 8: NO_2 (conc. 20ppb) reproducibility and repeatability in all IE.

Parameter	r (ppb)	R (ppb)	IE
NO_2	0.029	0.330	2005
NO_2	0.016	0.086	2007_1
NO_2	0.010	0.166	2007_2
NO_2	0.015	0.099	2008_1
NO_2	0.025	0.137	2008_2
NO_2	0.024	0.106	2008_3
NO_2	0.027	0.061	2009_1
NO_2	0.008	0.124	2009_2
NO_2	0.008	0.143	2010

Table 9: O_3 (conc. 120ppb) reproducibility and repeatability in all IE.

Parameter	r (ppb)	R (ppb)	IE
O_3	0.014	0.084	2007_1
O_3	0.018	0.027	2007_2
O_3	0.015	0.032	2008_1
O_3	0.007	0.035	2008_2
O_3	0.017	0.037	2008_3
O_3	0.013	0.086	2009_1
O_3	0.013	0.068	2009_2
O_3	0.009	0.077	2010

Table 10: O_3 (conc. 20ppb) reproducibility and repeatability in all IE.

Parameter	r (ppb)	R (ppb)	IE
O_3	0.008	0.124	2007_1
O_3	0.014	0.103	2007_2
O_3	0.006	0.052	2008_1
O_3	0.011	0.118	2008_2
O_3	0.006	0.050	2008_3
O_3	0.007	0.140	2009_1
O_3	0.006	0.080	2009_2
O_3	0.008	0.077	2010

Table 11: CO (conc. 8ppm) reproducibility and repeatability in all IE.

Parameter	r (ppb)	R (ppb)	IE
CO	0.004	0.035	2005
CO	0.002	0.070	2007_1
CO	0.004	0.053	2007_2
CO	0.004	0.049	2008_1
CO	0.010	0.110	2008_2
CO	0.001	0.072	2008_3
CO	0.003	0.073	2009_1
CO	0.007	0.046	2010

Table 12: CO (conc. 2ppm) reproducibility and repeatability in all IE.

Parameter	r (ppb)	R (ppb)	IE
CO	0.034	0.145	2005
CO	0.003	0.150	2007_1
CO	0.019	0.169	2007_2
CO	0.005	0.040	2008_1
CO	0.012	0.305	2008_2
CO	0.006	0.184	2008_3
CO	0.007	0.164	2009_1
CO	0.003	0.113	2010

Table 13: SO_2 (conc. 130ppb) reproducibility and repeatability in all IE.

Parameter	r (ppb)	R (ppb)	IE
SO_2	0.003	0.077	2005
SO_2	0.004	0.084	2007_1
SO_2	0.004	0.073	2007_2
SO_2	0.003	0.061	2008_1
SO_2	0.004	0.087	2008_2
SO_2	0.004	0.131	2008_3
SO_2	0.008	0.047	2009_1
SO_2	0.005	0.104	2009_2
SO_2	0.005	0.960	2010

Table 14: SO_2 (conc. 20ppb) reproducibility and repeatability in all IE.

Parameter	r (ppb)	R (ppb)	IE
SO_2	0.018	0.100	2005
SO_2	0.028	0.184	2007_1
SO_2	0.019	0.098	2007_2
SO_2	0.003	0.050	2008_1
SO_2	0.015	0.137	2008_2
SO_2	0.014	0.082	2008_3
SO_2	0.010	0.026	2009_1
SO_2	0.015	0.211	2009_2
SO_2	0.010	0.132	2010

3.6 Grubbs' test with outlier and straggler

Tests for data consistency and statistical outliers as described in ISO 5725-2 [13] were carried out during the evaluation. Laboratories showing some form of statistical inconsistency were requested to investigate the cause of discrepancies and laboratories were allowed to correct their results in case of identification of exceptional errors. In Table 15 outliers and stragglers are presented for each IE. This table for each failing result is also shown at which level was the anomaly: underestimation (Gmin), overestimation (Gmax) or both. Generally this test didn't show any relevant situation with a great number of outlier. Per each level of concentration for each pollutant (7) and (8) was used to define if the higher or the lower value was an outlier.

$$G\max = \frac{(x_{\max} - \bar{x})}{s} \tag{7}$$

$$G\min = \frac{(\bar{x} - x_{\min})}{s} \tag{8}$$

Table 15: Statistical outliers and stragglers according to Grubb's test for all IE.

IE	Site	Pollutant	Straggler	Failing test level	Outlier	Failing test level
Jun-05	Ispra (IT)	CO	1	Gmax	/	
		NO_2	2	Gmin/Gmax	2	Gmin/Gmax
Jun-07	Ispra (IT)	CO	/		1	Gmax
		NO_2	2	Gmax	2	Gmax
		O_3	1	Gmax		
		SO_2			1	Gmax
Oct-07	Essen (DE)	NO	1	Gmin	/	
Apr-08	Ispra (IT)	NO_2	1	Gmax	2	Gmin/Gmax
		O_3	1	Gmax	/	
Oct 08_1	Ispra (IT)	CO	1	Gmin	/	
		NO_2	1	Gmin	/	
		O_3	1	Gmax	/	
Oct 08_2	Ispra (IT)	NO	1	Gmin	/	
		NO_2	1	Gmax	/	
		O_3	2	Gmin	/	
Sep-09	Langen	all	none		none	
Oct-09	Ispra	NO	1	Gmin	2	Gmin
		O_3	1	Gmax	1	Gmax
		SO_2	/		1	Gmin
Jun-10	Ispra (IT)	CO	1	Gmax	1	Gmax
		NO	2	Gmin	/	
		NO_2	2	Gmin	/	
		O_3	2	Gmax	3	Gmax
		SO_2	1	Gmax	/	

4 Intercomparison exercise as a learning process

As discussed in section 3.4, in all IE the high percentages of valid measured values confirmed the general good performance of the laboratories.

From the En number results (section 3.4) came out a need to harmonize and to define an estimation procedure of the measurement uncertainty.

All results obtained in these exercises were below the DQO of 15% of expanded measurement uncertainty requested by the directive [1]. It must be considered that the IE took place in an ideal situation: ambient temperature under control, constant relative humidity, absence of interferences and all instruments were recently calibrated and maintained. Under routine conditions existing in the networks an increase of measurement uncertainty, repeatability and reproducibility is expected to happen.

In order to evaluate the ability of NRL to meet the DQO under field condition, it would be advisable to organize an IE using real samples.

One of the most important results obtained during these intercomparison exercises can be found in the opportunity for all the experts in air quality monitoring to exchange information and technical know-how. The way in which IE were managed gave the chance to experts in young teams, to those in new EU member States and Candidate Countries to get in touch with experienced colleagues.

References

[1] Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe, L 152, 11.06.2008
[2] EN 14626:2005, Ambient air quality – Standard method for the measurement of the concentration of carbon monoxide by non-dispersive infrared spectroscopy
[3] EN 14212:2005, Ambient air quality – Standard method for the measurement of the concentration of sulphur dioxide by ultraviolet fluorescence
[4] EN 14211:2005, Ambient air quality – Standard method for the measurement of the concentration of nitrogen dioxide and nitrogen monoxide by chemiluminescence
[5] EN 14625:2005, Ambient air quality – Standard method for the measurement of the concentration of ozone by ultraviolet photometry
[6] ISO 6143:2001, Gas analysis – Comparison methods for determining and checking the composition of calibration gas mixtures
[7] ISO 6144:2003, Gas analysis – Preparation of calibration gas mixtures – Static volumetric method
[8] ISO 6145-7:2001, Gas analysis – Preparation of calibration gas mixtures using dynamic volumetric methods – Part 7: Thermal mass-flow controllers
[9] http//ies.jrc.ec.europa.eu/aquila-homepage.html
[10] AQUILA Position Paper N. 37, (2008) Protocol for intercomparison exercise. Organisation of intercomparison exercises for gaseous air pollution for EU national air quality reference laboratories and laboratories of the WHO EURO region http://ies.jrc.ec.europa.eu/aquila-project/role-and-tasks-of-national-reference-laboratories.html
[11] ISO 13528:2005, Statistical methods for use in proficiency testing by interlaboratory comparisons
[12] ISO 5725-1:1994, Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions
[13] ISO 5725-2:1994, Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method
[14] ISO 5725-6:1994, Accuracy (trueness and precision) of measurement methods and results – Part 6: Use in practice of accuracy values

[15] De Saeger E. et al., (1997) European comparison of Nitrogen Dioxide calibration methods, EUR 17661
[16] ISO 15337:2009, Ambient air – Gas phase titration – Calibration of analysers for ozone
[17] Kapus M. et al. (2009) The evaluation of the Intercomparison Exercise for SO2, CO, O3, NO and NO2 carried out in October 2007 in Essen. JRC scientific and technical reports. EUR 23788.
[18] Viallon J. et al 2009 Metrologia 46 08017. Final report, on-going key comparison BIPM.QM-K1: Ozone at ambient level, comparison with JRC, 2008. doi: 10.1088/0026-1394/46/1A/08017