Calibration and measurement capabilities for nitrogen monoxide in nitrogen[†]

Adriaan M.H. van der Veen, Heleen Meuzelaar, Iris de Krom, and Gerard Nieuwenkamp

VSL, Dutch Metrology Institute, Thijsseweg 11, 2629 JA Delft, the Netherlands

15 February 2018

Abstract

In 2016, the Gas Analysis Working Group of the Consultative Committee on Amount-of-Substance (GAWG) introduced an extrapolation scheme to be used for translating the measurement uncertainty reported in a key comparison to a measurement capability. This extrapolation scheme has a tipping point at $10\,\mu\mathrm{mol}\,\mathrm{mol}^{-1}$. Below this amount-of-substance fraction level, the expanded uncertainty is assumed to be independent of the amount-of-substance fraction, and above the relative expanded uncertainty is assumed to constant. So, a relative expanded uncertainty of $1.0\,\%$ at $10\,\mu\mathrm{mol}\,\mathrm{mol}^{-1}$ translates to an expanded uncertainty of $1.0\,\%$ at $0.1\,\mathrm{mol}\,\mathrm{mol}^{-1}$, and into a relative expanded uncertainty of $10\,\%$ at $1.0\,\mu\mathrm{mol}\,\mathrm{mol}^{-1}$.

This work revisits the calibration and measurement capabilities (CMCs) for nitrogen monoxide and outlines the calculations necessary for a customised extrapolation scheme. Key inputs include the measurements performed to support the national measurement standards, the stability studies run on representative gas mixtures to understand their long-term behaviour and last but not least, the uncertainty calculations performed in the participation in key or other international comparisons. Meta-analysis is used to determine a reproducibility componemt of uncertainty, which is combined with the standard uncertainty calculated in accordance with ISO 6143.

The re-evaluation of the long-term stability study data for calibration gas mixtures of nitrogen monoxide in nitrogen shows that there is good agreement between the expanded uncertainty calculated from these data and the expanded uncertainties submitted in two international comparisons. It is shown that the relationship between the amount-of-substance fraction and the expanded uncertainty can be satisfactorily be described by a straight line when taking the logarithm on both axes to get an acceptable distribution of the data points along the axes. the proposed extrapolation scheme outperforms the scheme proposed by the GAWG, especially for low and high amount-of-substance fractions. At the same time, the re-evaluation defies the approach using the performances demonstrated in 'track A' key comparisons for CMCs for nitrogen monoxide.

1 Introduction and rationale

In 2016, the Gas Analysis Working Group of the Consultative Committee on Amount-of-Substance (GAWG) introduced an extrapolation scheme to be used for translating the measurement uncertainty reported in a key comparison to a calibration and measurement capability (CMC) [1]. This

---

[†]Filed as report S-CH.17.11

extrapolation scheme has a tipping point at $10\,\mu\mathrm{mol\,mol^{-1}}$. Below this amount-of-substance fraction level, the expanded uncertainty is extrapolated absolute, and above relative. So, a relative expanded uncertainty of $1.0\,\%$ at $10\,\mu\mathrm{mol\,mol^{-1}}$ translates to an expanded uncertainty of $1.0\,\%$ at $0.1\,\mathrm{mol\,mol^{-1}}$, and into an absolute expanded uncertainty of $0.1\,\mu\mathrm{mol\,mol^{-1}}$ at $1.0\,\mu\mathrm{mol\,mol^{-1}}$.

The introduction of this extrapolation scheme for calibration and measurement capabilities (CMCs) had some unfortunate side effects for VSL, especially for the services related to ppm-levels of nitrogen monoxide and propane. The claimed calibration and measurement capabilities for these services had to be increased for no other reason than that the extrapolation scheme [1] predicted something different from what can be justified on the basis of, e.g., the multipoint calibration of the analyser and propagating the uncertainty from gas mixture preparation [2, 3] and analysis [4], among other uncertainty sources.

For the measurement standards and related calibration and reference material services, there is another isuse. The use of expanded uncertainties reported during the last three 'track A' key comparisons (assessing the 'core capabilities' of a national metrology institutes [1]) provided $U_{\mathrm{rel}}(x_0) = 0.09\,\%$ at amount-of-substance fraction $x_0 = 10\,\mu\mathrm{mol\,mol^{-1}}$, a CMC that is not supported by any evidence involving this group of measurement standards, such as the calculation of the measurement uncertainty using a calibration function or the evaluation of relevant long-term stability study data. Notwithstanding that the CMC is too low at x_0, it is much too high at ambient level, considering the participation in key comparison CCQM-K26a [5]. The predicted relative expanded uncertainty is $4\,\%$, whereas the reported relative expanded uncertainty was $1.1\,\%$. The degree-of-equivalence indicated satisfactory performance in CCQM-K26a.

The decision was taken to revisit the data collected over the years concerning the status of the national measurement standards (in most cases suites of Primary Standard gas Mixtures, PSMs) and related measurement and preparative facilities and to develop for all groups of measurement standards a customised extrapolation scheme that relates the expanded uncertainty for calibrations and reference materials to the amount-of-substance fractions for the entire range for which services are provided. A model for calculating CMCs is presented, as well as an extrapolation scheme.

2   Default extrapolation scheme

The amount-of-substance fraction range described in the GAWG Strategy [1] for a component runs in principle from the purity analysis up to $50\,\%$ The tipping point in the extrapolation scheme is at an amount-of-substance fraction $x_0 = 10\,\mu\mathrm{mol\,mol^{-1}}$. For smaller amount-of-substance fractions, the expanded uncertainty is assumed to be the same as for x_0, and for greater amount-of-substance fractions, the relative expanded uncertainty is assumed to be constant, viz., [1]

$$U(x) = \frac{U(x_0)}{x_0} \cdot x \qquad\qquad \text{for } x > x_0 \tag{1}$$
$$U(x) = U(x_0) \qquad\qquad \text{for } x \leq x_0$$

According to the same strategy document, calibration and measurement capabilities can be supported in two different ways by key comparisons [1]:

1. a dedicated key comparison (track A, track B, or track C)

2. for selected components and ranges, the combination of the last three track A key comparisons

Where possible, VSL seeks to support its calibration and measurement capabilities though the default scheme, i.e., using dedicated key comparisons. The flexible scheme, involving the last three track A key comparisons, is used in exceptional cases. In 2017, VSL was requested to resubmit their CMCs based on the performance in the last three track A key comparisons. The pooling of the past three participations in track A key comparisons gave an relate expanded uncertainty of 0.09 % at $x_0 = 10 \, \mu\text{mol} \, \text{mol}^{-1}$.

3   Uncertainty calculations in key comparisons

Uncertainty calculations in key comparisons largely follow the usual calculations used in the maintenance of the national measurement standards, and the production of Primary and Certified Reference Materials (PRMs and CRMs). PSMs are generally maintained with only the amount-of-substance fraction and standard uncertainty from gravimetry (calculated in accordance with ISO 6142 [2, 6]).

In preparative key comparisons, such as CCQM-K53 [7] and CCQM-K54 [8], an allowance is made for, e.g., gas sampling and adsorption effects, often based on the standard uncertainty from verification [2, 4].

In analytical key comparisons, such as CCQM-K76 [9], the basis for the uncertainty calculation is usually the one given in ISO 6143 [4]. As a rule, the mean value of the measurements is reported, together with the expanded uncertainty for a single measurement. ISO 6143 requires obtaining the responses under *reproducibility conditions*, which is rarely done, and moreover, difficult to achieve in practice. In practice, instrument responses are obtained under *repeatability conditions*, and an allowance is made for reproducibility effects in the analysis. Initially, this approach took the standard deviation of the three (or more) measurements as additional uncertainty component. The standard deviation is assessed against the standard uncertainty calculated for a single measurement, to evaluate to what extent the standard uncertainty could be explained from the dispersion of the measurements. More recently, statistical methods known from meta-analysis [10] are used instead, which formalises the basic idea.

Another adjustment made is to adapt the standard uncertainty calculated for the PSMs. One of the first instances was in CCQM-K76 [11]. The purpose of including these extra uncertainties is first and foremost to ensure that the calculated calibration function meets the requirements of ISO 6143. Usually the additional uncertainty contribution is combined with the standard uncertainty calculated from gravimetry, and assumed to be independent (i.e., no systematic effect for all PSMs).

4   Calculation of CMCs

The calculation of CMCs is based on the following mixed effects model

$$y_i = \mu + At_i + B_i + \varepsilon_i \tag{2}$$

where y_i denotes the amount-of-substance fraction of the ith data point, μ the expected value at time 0, A the degradation rate, t_i the time of the ith data point, B_i the bias in data point i, and the random measurement error ε_i. The mixed effects model in equation (2) is an extension of the simple mixed effects model for stability study in reference material production [12, 13]. The term B_i enables accounting for an excess variance [10] due to a reproducibility effect in the stability study measurements. For stable gas mixtures, $A = 0$ and the model in equation (2) reduces to a

random effects model as known from between-bottle homogeneity studies [14] and meta-analysis [10].

In the following, the excess variance $\tau^2 = \mathrm{var}(B_i)$ and the variance computed from the ith measurement $\sigma^2 = \mathrm{var}(\varepsilon_i)$. The latter variance is computed using the procedure of ISO 6143 [4]. For evaluating the long-term stability study data, the variance computed for the last measurement is used in the subsequent meta-analysis. If there would be an improvement in the performance of the measurement method, then this would result in an understatement of the standard uncertainty of older measurements. Consequently, the value obtained for τ would be larger. For the data for nitrogen monoxide, the assumption that σ did not change during the stability study is justified, for there have not been any meaningful improvements in the instrumentation and methods used.

The data from the gas mixtures used for long-term stability monitoring have been used. At regular time intervals, these gas mixtures are analysed using a suite of Primary Standard gas Mixtures (PSMs). The data of the stability study is then assessed for a trend [12]. If no trend is observed, the data is then fitted to the DerSimonian-Laird model to assess whether there is an excess variance. This excess variance is a reproducibility component, and calculated as [15]

$$\tau^2 = \frac{\sum_i w_i (y_i - \bar{y})^2 - (a-1)}{\sum_i w_i - \sum_i w_i^2 / \sum_i w_i} \tag{3}$$

where $w_i = 1/s_i^2$, s_i denotes the within-group standard deviation, a the number of groups, and $\bar{y} = \sum_i w_i \bar{Y}_i / \sum_i w_i$. Similar to traditional ANOVA, in meta-analysis $\tau^2 = 0$ when the result of equation (3) is negative.

The calibration and measurement capability is then calculated as

$$U = k\sqrt{\tau^2 + \sigma^2} \tag{4}$$

where σ denotes the standard uncertainty computed in accordance with ISO 6143, k the coverage factor and U the CMC. This computation is made for several amount-of-substance fractions across the entire range for which measurement standards of NO in nitrogen are maintained.

5   Results and discussion

The first subrange is the range of amount-of-substance fractions from 0.1 % to 1.0 %. These analyses are performed using a non-dispersive infrared analyser (NDIR). The calibration function is a third-order polynomial. Figure 1 shows the calibration data (top left) , the calibration function (bottom left), and the residuals in both directions (right-hand side). All residuals meet the criteria of ISO 6143.

The calibration function for the amount-of-substance fraction range from $100\,\mu\mathrm{mol\,mol^{-1}}$ to $1000\,\mu\mathrm{mol\,mol^{-1}}$ is also cubic (see figure 2). With one exception, all residuals meet the criteria of ISO 6143. The small discrepancy in the y-residual of the PSM at $200\,\mu\mathrm{mol\,mol^{-1}}$ has negligible influence on the calculation of the CMC.

The calibration function for the amount-of-substance fraction NO range from $10\,\mu\mathrm{mol\,mol^{-1}}$ to $100\,\mu\mathrm{mol\,mol^{-1}}$ is also cubic (see figure 3). All residuals are meeting the criteria of ISO 6143.

The calibration function for the amount-of-substance fraction NO range from $1\,\mu\mathrm{mol\,mol^{-1}}$ to $10\,\mu\mathrm{mol\,mol^{-1}}$ is a straight line (see figure 4). All residuals are meeting the criteria of ISO 6143, save the y-residuals at 1000 a.u., 5000 a.u. and 10 000 a.u.. Previous measurements in the same range showed full consistency with the criteria of ISO 6143 for the residuals.
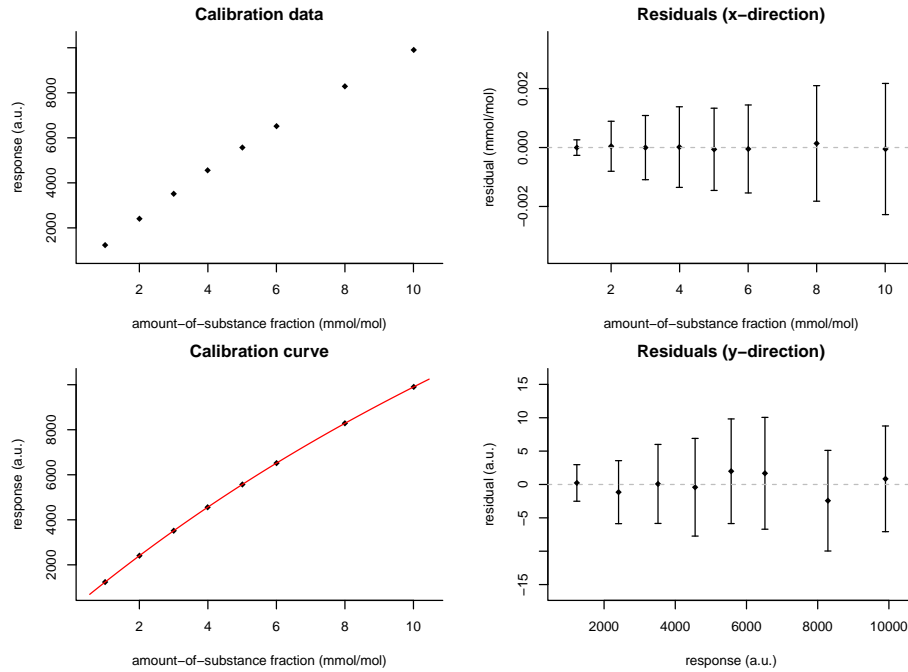
Figure 1: Calibration data, function and residuals for nitrogen monoxide in nitrogen on a non-dispersive infrared analyser (NDIR) calibrated in accordance with ISO 6143 [4]. The uncertainty bars represent 95 % coverage intervals.
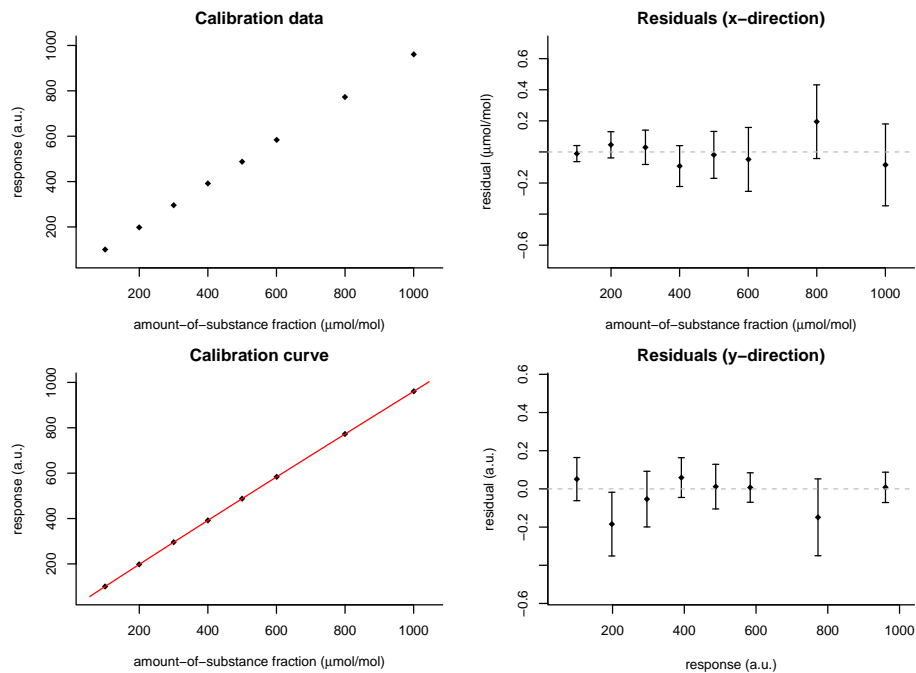

Figure 2: Calibration data, function and residuals for nitrogen monoxide in nitrogen on a Non-Dispersive Ultraviolet analyser (NDUV) calibrated in accordance with ISO 6143 [4]. The uncertainty bars represent 95 % coverage intervals.
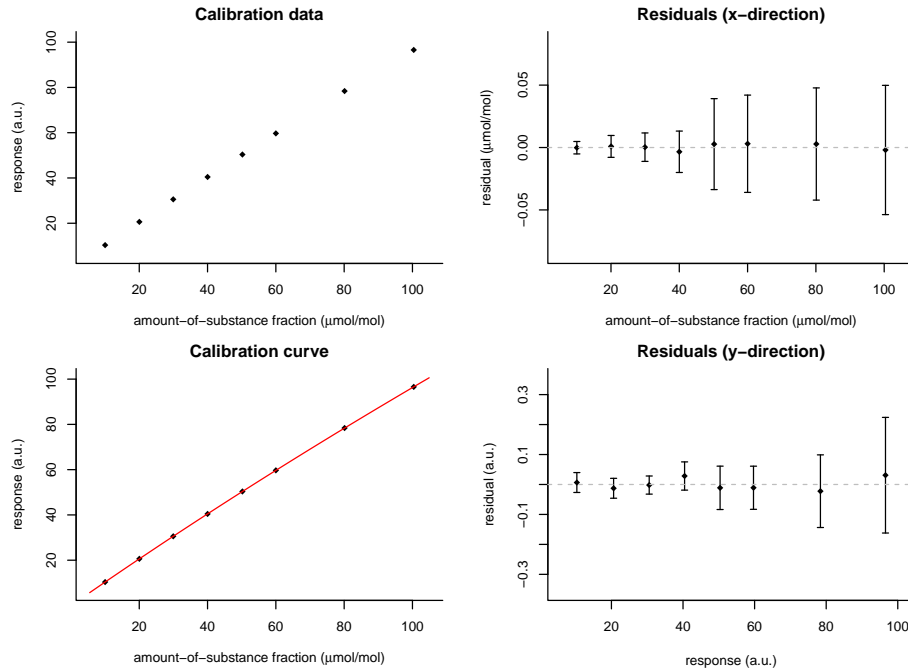
Figure 3: Calibration data, function and residuals for nitrogen monoxide in nitrogen on a Non-Dispersive Ultraviolet (NDUV) analyzer configured in accordance with ISO 6143 [4]. The uncertainty bars represent 95 % coverage intervals.
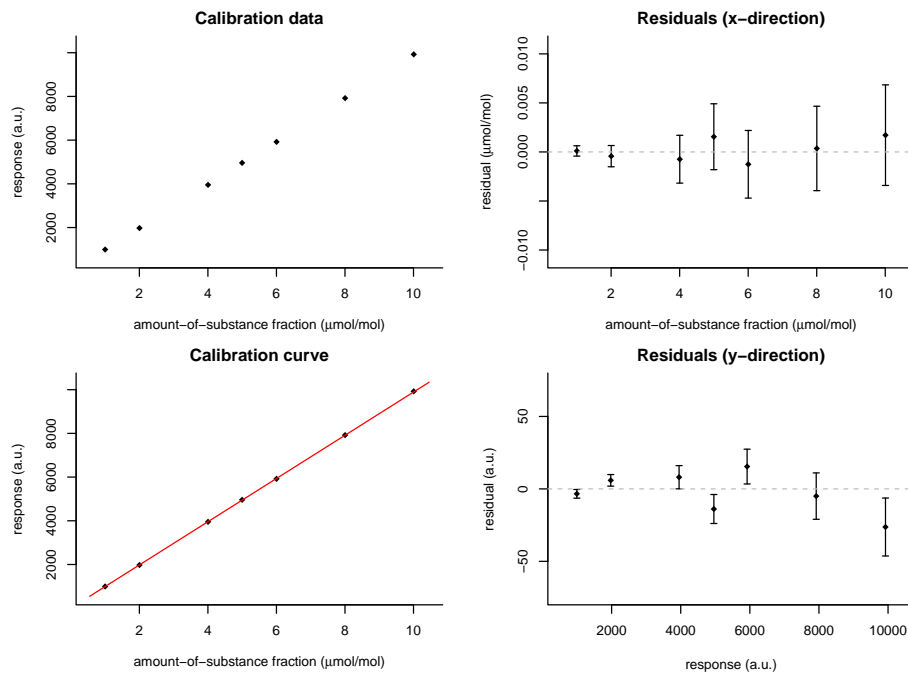

Figure 4: Calibration data, function and residuals for nitrogen monoxide in nitrogen on a chemi-luminescence detector (CLD) configured in accordance with ISO 6143 [4]. The uncertainty bars represent 95 % coverage intervals.
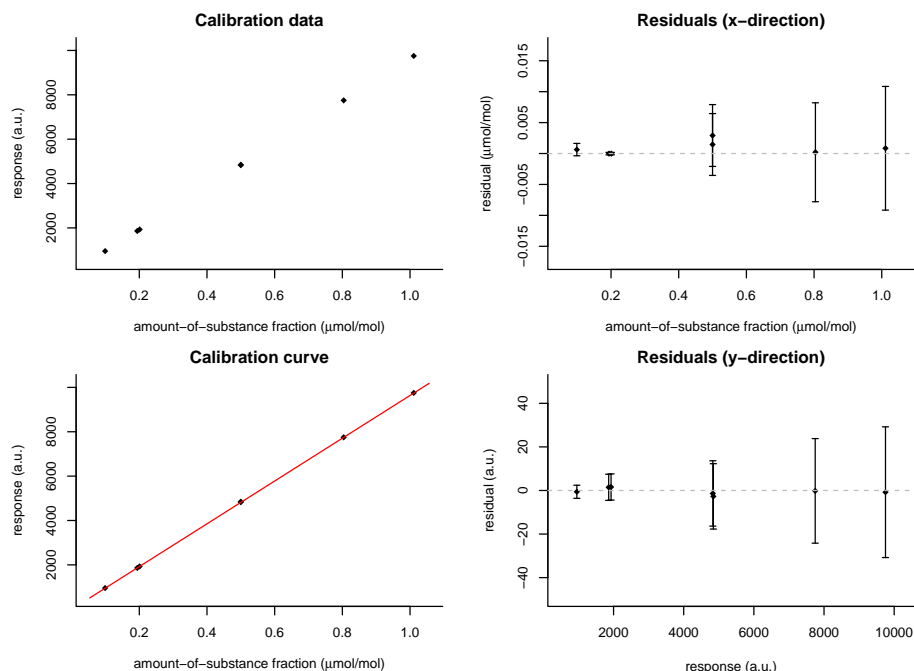
Figure 5: Calibration data, function and residuals for nitrogen monoxide in nitrogen on a chemiluminescence detector (CLD) configured in accordance with ISO 6143 [4]. The uncertainty bars represent 95% coverage intervals.

The calibration function for the amount-of-substance fraction NO range from $0.1\,\mu\text{mol}\,\text{mol}^{-1}$ to $1\,\mu\text{mol}\,\text{mol}^{-1}$ is a straight line (see figure 5). All residuals are meeting the criteria of ISO 6143. the multipoint calibration has been performed by diluting a single PSM. The results are much better than those obtained using a calibration with PSMs (see figure 6).

The results of the meta-analysis on the long-term stability study data are shown in table 1. The between-group standard deviations (expressed as coefficients of variation) decrease with increasing amount-of-substance fraction. The gas mixtures VSL128500 and PRM126817 show exceptionally good behaviour. The within-group coefficient of variation also decreases with increasing amount-of-substance fraction. The results for gas mixtures VSL128500 and PRM126817 are exceptionally good and have been discarded in the development of the relationship between the expanded uncertainty of the calibration of a gas mixture and the amount-of-substance fraction NO.

The results of the international comparisons are summarised in table 2. The results for the method using the molblocs to dilute a PSM are given in table 3. This method is used for $100\,\text{nmol}\,\text{mol}^{-1} \leq x \leq 1000\,\text{nmol}\,\text{mol}^{-1}$ where x denotes the amount-of-substance fraction NO.

The relationship between the amount-of-substance fraction and the standard uncertainty is visualised in figure 7. In the figures, also the results from CCQM-P73 [16] and CCQM-K26a [5] are included. It is important to note that (1) the results in these comparisons were consistent, and (2) the calculated standard uncertainty matches well with the standard uncertainties summarised in table 1.

From figure 7a, in can be seen that in first approximation the data can be described by a model of the kind

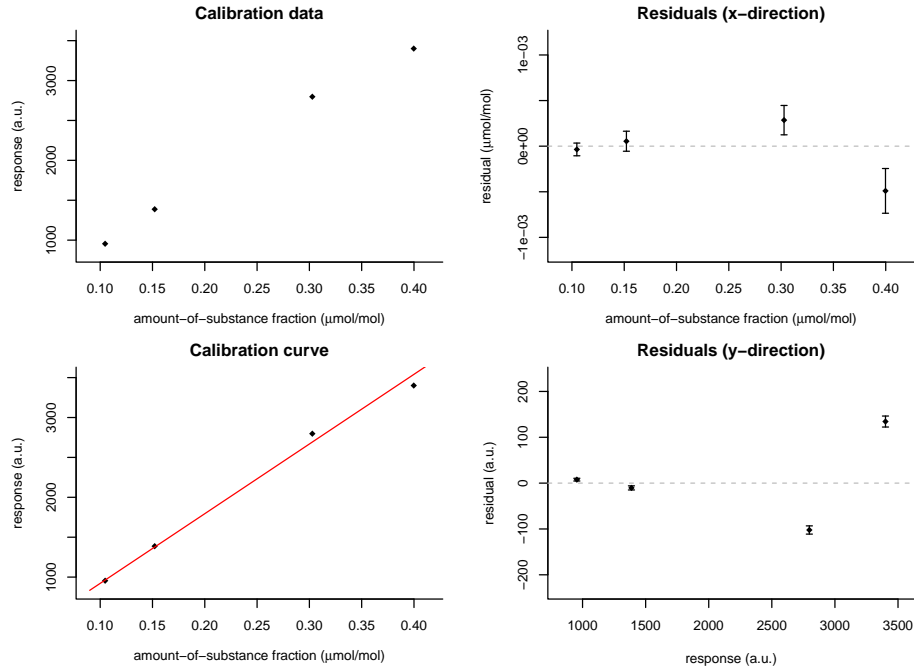$$\log U = a_1 \log x + a_0 \tag{5}$$

Figure 6: Calibration data, function and residuals for nitrogen monoxide in nitrogen on a chemi-luminescence detector (CLD) configured in accordance with ISO 6143 [4]. The uncertainty bars represent 95% coverage intervals.

Table 1: Results from the meta-analysis performed on the calibration gas mixtures used to assess the long-term behaviour of nitrogen monoxide in nitrogen mixtures. τ denotes the coefficient of variation between-groups, and σ that within-groups. x denotes the amount-of-substance fraction nitrogen monoxide, as computed using ISO 6142-1.

| Mixture	x	τ	σ	$\sqrt{\tau^2 + \sigma^2}$
PRM280190	2.00×10^{-6}	0.26	0.12	0.28
PRM280208	5.01×10^{-6}	0.12	0.11	0.16
VSL167407*	2.00×10^{-5}	0.09	0.05	0.10
PRM105634	2.52×10^{-5}	0.13	0.04	0.14
VSL174483	6.10×10^{-5}	0.16	0.03	0.16
VSL128500	3.93×10^{-4}	0.01	0.01	0.02
PRM126817	8.74×10^{-4}	0.01	0.01	0.02
PRM240045	1.78×10^{-3}	0.11	0.01	0.11
PRM229387	5.01×10^{-3}	0.04	0.01	0.05

Table 2: Reported expanded uncertainty (including verification) in CCQM and Euramet projects; all results were consistent with the assigned value

Project	x $\mathrm{mol\,mol^{-1}}$	$U(x)$ $\mathrm{mol\,mol^{-1}}$	$U_{\mathrm{rel}}(x)$
CCQM-P73	4.70×10^{-5}	1.6×10^{-7}	0.34 %
CCQM-P73	6.30×10^{-5}	2.2×10^{-7}	0.35 %
CCQM-K26a	7.19×10^{-7}	8.0×10^{-9}	1.1 %
Euramet-K26a	4.32×10^{-7}	3.6×10^{-9}	0.83 %

Table 3: Calculation of the standard uncertainty for the dilution method ($100\,\mathrm{nmol\,mol^{-1}} \leq x \leq 1000\,\mathrm{nmol\,mol^{-1}}$)

Mixture	x $\mathrm{nmol\,mol^{-1}}$	σ %	τ %	$\sqrt{\sigma^2 + \tau^2}$ %
PRM427305	297.97	0.53	0.87	1.02
PRM337335	148.99	0.36	0.81	0.88
PRM527291	102.88	0.35	0.72	0.80
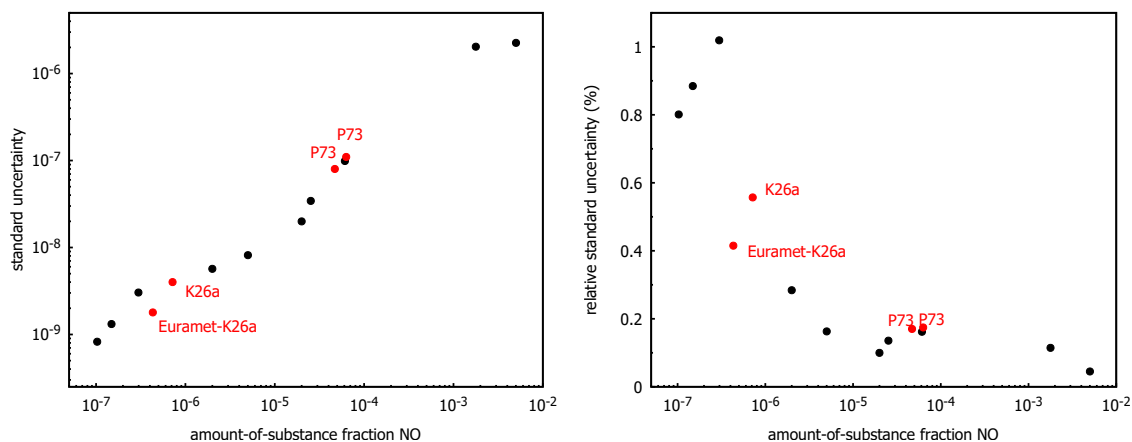
(a) Standard uncertainty as a function of the amount-of-substance fraction NO

(b) Coefficient of variation as a function of the amount-of-substance fraction NO

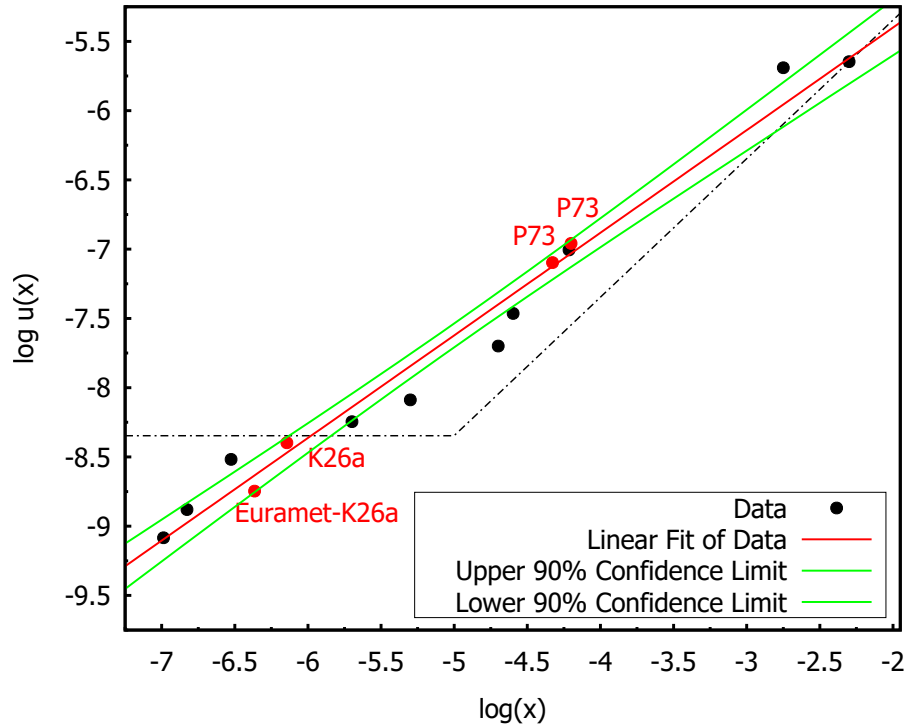Figure 7: CMCs as a function of the fraction NO

Figure 8: Regression of the standard uncertainty as a function of the amount-of-substance fraction NO. The dotted lines indicate the extrapolation scheme from the GAWG

where a_1 denotes the slope and a_0 the intercept. Comparing this model with equation (1), it has the same shape for $x \geq 10 \, \mu\text{mol mol}^{-1}$ if the slope is set $a_1 = 1$. Fitting the data using ordinary least squares (OLS) [17] yields for $a_1 = 0.741$ (see figure 8), where it should be noted that the data from gas mixtures VSL128500 and PRM126817 have been eliminated.

The GAWG extrapolation scheme (see equation (1)) can be written in a form similar to equation (5). Taking the logarithm on either side of equation (1) gives the following relationship

$$\log U(x) = \log U(x_0) - \log x_0 + \log x$$

Comparing this expression with equation (5), for the GAWG extrapolation scheme, $a_1 = 1$ and $a_0 = \log U(x_0) + \log x_0$. From figure 8, it can be seen that the regression line is flatter than the line from the GAWG extrapolation scheme for $x > x_0$.

For setting up a formula to predict the CMC as a function of the amount-of-substance fraction, the use of OLS is not entirely appropriate, as typically there are some data points $(x, u(x))$ above the line, and others below. The points above the line indicate data for which the predicted CMC might constitute an understatement of the actual uncertainty. An increase of the intercept a_0 would 'cure' this problem to the extent that the predicted CMC is not meaningfully smaller than the observed (computed) CMC.

6 Interpolation scheme for CMCs

Based on the calculations as presented, the relationship becomes

$$\log u = -3.919 + 0.741 \log x \tag{6}$$

Two datapoints lie slightly above the 90 % coverage interval, but otherwise the equation gives a good prediction of the CMCs across the entire range of amount-of-substance fractions.

Based on equation (6), the relative expanded uncertainty ranges from 1.6 % $\rightarrow$ 0.08 % for amount-of-substance fractions from $100 \times 10^{-9}\,\mathrm{mol\,mol^{-1}}$ to 1 %.

7 Conclusions

The approach developed for relating the expanded uncertainty of the analysis of nitrogen monoxide in nitrogen to the amount-of-substance fraction provides satisfactory results. The use of meta-analysis is helpful in characterising the dispersion due to, a combination of, reproducibility of the measurement method, effects of cylinder walls and regulators. The standard deviation τ from the DerSimonian-Laird model is combined with the standard uncertainty computed from the multi-point calibration from ISO 6143 to determine the calibration and measurement capability (CMC). The results of the participation in two international comparisons agree well with the CMCs determined from the long-term behaviour of the standards of nitrogen monoxide in nitrogen.

The extrapolation scheme developed by the GAWG does a modest job in characterising the CMCs. It is a crude model, especially for amount-of-substance fractions below $10\,\mathrm{\mu mol\,mol^{-1}}$ where it overstates the CMCs. It also overstates the CMCs at the high end of the range, although there the agreement between the GAWG scheme and the fitted data is much better.

The proposed model for inter- and extrapolating CMCs performs well from the ambient level of NO in air to the highest fractions in the range.

References

[1] P. Brewer and A. M. H. van der Veen. GAWG strategy for comparisons and CMC claims. GAWG, Gas Analysis Working Group, Sévres, France, October 2016.

[2] Gas analysis – Preparation of calibration gas mixtures – Gravimetric method for Class I mixtures. ISO 6142–1, International Organization for Standardization, Geneva, Switzerland, First edition 2015.

[3] Gas analysis – Preparation of calibration gas mixtures using dynamic methods. ISO 6145, International Organization for Standardization, Geneva, Switzerland. 11 parts.

[4] Gas analysis – Comparison methods for determining and checking the composition of calibration gas mixtures. ISO 6143, International Organization for Standardization, Geneva, Switzerland, 2001. Second edition.

[5] Alejandro Perez, Masaaki Maruyama, Jiri Novak, Jari Walden, Annette Borowiak, Friedrich Lagler, Jin Seog Kim, Tatiana Macé, Christophe Sutour, Franklin Guenther, Adriaan van der Veen, Gerard Nieuwenkamp, Gergely Vargha, Martin Milton, Graham Leggett, Anneliese Medem, Volker Stummer, Leonid Konopelko, Yuri Kustikov, and Natalia Pivovarova. International key comparison CCQM-K26a and Pilot Study CCQM-P50a (NO). CIPM Key Comparison Database, 2007.

[6] Gas analysis – Preparation of calibration gas mixtures – Gravimetric method. ISO 6142, International Organization for Standardization, Geneva, Switzerland, second edition, 2001.

[7] Jeongsoon Lee, Jin Bok Lee, Dong Min Moon, Jin Seog Kim, Adriaan M H van der Veen, Laurie Besley, Hans-Joachim Heine, Belén Martin, L A Konopelko, Kenji Kato, Takuya Shimosaka, Alejandro Perez Castorena, Tatiana Macé, Martin J T Milton, Mike Kelley, Franklin Guenther, and Angelique Botha. Final report on international key comparison CCQM-K53: Oxygen in nitrogen. *Metrologia*, 47(1A):08005–08005, jan 2010.

[8] Adriaan M H van der Veen, Hima Chander, Paul R Ziel, Rob M Wessel, Ed W B de Leer, Damian Smeulders, Laurie Besley, Kenji Kato, Takuro Watanabe, Jin Seog Kim, Jin-Chun Woo, Hyun Kil Bae, Yong Doo Kim, Alejandro Pérez Castorena, Francisco Rangel Murillo, Victor M Serrano Caballero, Carlos Ramírez Nambo, Manuel de Jesús Avila Salas, Leonid A Konopelko, Tatjana A Popova, V V Pankratov, M A Kovrizhnih, T A Kuzmina, O V Efremova, Yury A Kustikov, Martin J T Milton, Gergely Vargha, Frank R Guenther, and George C Rhoderick. International comparison CCQM-K54: Primary standard gas mixtures of hexane in methane. *Metrologia*, 47(1A):08019, 2010.

[9] Franklin R Guenther, Michael E Kelley, Gerald D Mitchell, Manuel de Jesús Avila Salas, Jorge Koelliker Delgado, Francisco Rangel Murillo, Victor M Serrano Caballero, Alejandro Pérez Castorena, Uehara Shinji, Dariusz Ciecior, Valnei Smarçaro da Cunha, Cristiane Rodrigues Augusto, Claudia Cipriano Ribeiro, Andreia de Lima Fioravante, Florbela Dias, Oh Sang-Hyub, Tatiana Macé, Christophe Sutour, Tamás Büki, Han Qiao, Angelique Botha, David M Mogale, James Tshilongo, Napo Ntsasa, Tshepiso Mphamo, Ian Uprichard, Martin Milton, Gergely Vargha, Chris Brookes, Prabha Johri, Miroslava Valkova, Leonid Konopelko, Yury Kustikov, V V Pankratov, D V Rumyantsev, M V Pavlov, E V Gromova, Adriaan van der Veen, Peter van Otterloo, and Rob M Wessel. International comparison CCQM-k76: Sulfur dioxide in nitrogen. *Metrologia*, 48(1A):08015–08015, jan 2011.

[10] Guido Schwarzer, James R Carpenter, and Gerta Rücker. *Meta-Analysis with R (Use R!)*. Springer, 2015.

[11] Adriaan M.H. van der Veen. Processing of the data for CCQM-K76. Technical report, VSL, Dutch Metrology Institute, Delft, the Netherlands, 2010. Report S-CH.10.09.

[12] Adriaan M. H. van der Veen, Thomas P. J. Linsinger, Andree Lamberty, and Jean Pauwels. Uncertainty calculations in the certification of reference materials. 3. Stability study. *Accreditation and Quality Assurance*, 6(6):257–263, may 2001.

[13] ISO Guide 35 reference materials – Guidance for characterization and assessment of homogeneity and stability. ISO, International Organization for Standardization, Geneva, Switzerland, 2017. Fourth edition.

[14] A. M. H. van der Veen, T. P. Linsinger, and J. Pauwels. Uncertainty calculations in the certification of reference materials. 2. Homogeneity study. *Accreditation and Quality Assurance*, 6(1):26–30, jan 2001.

[15] Rebecca DerSimonian and Nan Laird. Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3):177–188, 1986.

[16] Robert I Wielgosz, Michael Esler, Joële Viallon, Philippe Moussay, Sang Hyub Oh, Byung Moon Kim, Angelique Botha, James Tshilongo, Ipeleng S Mokgoro, Masaaki Maruyama, Tatiana Mace, Christophe Sutour, Viliam Stovcík, Miroslava Valková, Stanislav Musil, Alejandro Pérez Castorena, Victor Serrano Caballero, Francisco Rangel Murillo,

Leonid A Konopelko, Yury A Kustikov, Vladimir V Pankratov, Elena V Gromova, William J Thorn, Franklin R Guenther, Damian Smeulders, Gonçalo Baptista, Florbela Dias, Rob M Wessel, Gerard Nieuwenkamp, and Adriaan M H van der Veen. Final report on CCQM-P73: International comparison of nitrogen monoxide in nitrogen gas standards (30–70) μmol mol^{-1}. *Metrologia*, 45(1A):08002, 2008.

[17] Norman R. Draper and Harry Smith. *Applied Regression Analysis*. Wiley-Blackwell, apr 1998.