INTERNATIONAL STANDARD

ISO 15796

First edition
2005-03-01

Gas analysis — Investigation and treatment of analytical bias

Analyse des gaz — Investigation et traitement des biais analytiques

Reference number
ISO 15796:2005(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 15796 was prepared by Technical Committee ISO/TC 158, *Analysis of gases*.

Introduction

Traceability is considered as one of the key items of quality assurance in gas analysis. In general, it is defined by the existence of unbroken chains of comparisons, relating the analytical result to acknowledged standards of measurement. More specifically, an analytical result is considered traceable if, by way of these comparisons, it has been demonstrated to be free of significant bias, significance referring to the specified uncertainty of the result.

As a rule, traceability is not demonstrated individually for a single analytical result but for a defined analytical procedure with specified ranges of analyte concentration and matrix composition. An analytical procedure is considered traceable if it has been demonstrated to be free of significant bias, or if significant bias has been corrected, by measurement on representative samples of known traceable composition. These may be samples of appropriate reference gas mixtures. Alternatively, other representative samples may be analysed in parallel using an accepted reference procedure.

This International Standard provides generic methods for demonstrating, or establishing, traceability of analytical procedures using reference gas mixtures or reference analytical procedures, implementing principles laid out in ISO 14111 [1] and ISO/TS 14167 [2], and respecting the principles of the *Guide to the Expression of Uncertainty in Measurement* (GUM) [3].

In this International Standard, the term “concentration” is used for two different purposes:

— as a general term for quantities measured in gas composition analysis, replacing the term “content” (see ISO 7504 [4]);

— as a generic substitute for any of the specific quantities measured in gas composition analysis such as the mass concentration or the mole fraction of a specified analyte (see ISO 7504 [4]).

Gas analysis — Investigation and treatment of analytical bias

1 Scope

This International Standard specifies generic methods for detecting and correcting bias (systematic errors) of analytical procedures for the analysis of gases, using reference gas mixtures or reference analytical procedures, as well as for estimating the correction uncertainty.

The main sources of (and parameters affecting) bias of analytical procedures are instrumental drift (time) and matrix interferences (matrix composition). Moreover, bias normally varies with analyte concentration. This International Standard therefore establishes protocols for

— detecting and correcting drift for an analytical system of limited stability,

— investigating and handling bias of a stable analytical system for a specified range of sample composition,

which are intended to be used in method development and method validation studies, either separately or sequentially.

This International Standard specifies procedures for two options, applicable to systematic effects, as follows:

a) tracing the observed pattern of deviations and correcting for their effect,

b) averaging over their effects and increasing the uncertainty,

where normally the first option entails lower uncertainty at the expense of higher effort.

For the convenience of the user, the methods specified in this International Standard are described for procedures of composition analysis, i.e. procedures for measuring the concentration of a specified analyte in a gas mixture. However, they are equally applicable to measurements of physico-chemical properties of a gas or gas mixture relevant to gas analysis, and translation into this subject field is straightforward.

2 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

2.1
bias
estimate of systematic error

NOTE Since the true value of a measurand cannot be known exactly, systematic errors cannot be determined exactly but have to be estimated using reference values.

2.2
correction
procedure by which the uncorrected result of a measurement is adjusted to compensate for systematic error

NOTE 1 Since systematic errors cannot be determined exactly, a correction can never be complete.

NOTE 2 In the VIM [5], the term correction is used with a different meaning.

2.3
uncertainty
parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand

[GUM:1993 [3], definition 2.3.2]

2.4
traceability
property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons

[VIM:1993 [5], definition 6.10]

NOTE In ISO 14111 [1], the term traceability is defined as the ability to provide evidence of the overall accuracy attributed to measurement results through documented calibrations, using measurement standards of known accuracy and comparison measurements of known performance.

2.5
reference value
estimate of a quantity, with sufficiently well established traceability and specified uncertainty, used as a reference for a specified purpose

NOTE In gas analysis, reference values of composition or physico-chemical properties are most often provided by reference gas mixtures and reference analytical procedures.

2.6
reference gas mixture
calibration gas mixture whose composition is sufficiently well established and stable to be used as a reference standard of composition from which other composition data are derived

[ISO 7504:2001 [4], definition 4.1.1]

2.7
reference analytical procedure
analytical procedure which is capable of providing traceable results with sufficiently well established uncertainty for use as reference values

2.8
drift
slow change of output, at constant input, of a measuring system

2.9
stability
〈of a measuring system〉 absence of significant drift

2.10
matrix interference
change in analytical response for a specified analyte, caused by variations in matrix composition

3 Symbols

A, B	specified drift-control mixtures
b_i	parameters of a bias correction model
δ	deviation (from a reference value)
I	interferent under consideration
M	gas mixture under consideration
m, n, N	number of data in a data series
p	number of correction parameters
Q	recovery (with respect to a reference value)
r	number of reference gas mixtures used for bias investigation
R_i	reference gas mixtures used for bias investigation
s	standard deviation of a data series
s_r	relative standard deviation of a data series (coefficient of variation)
s^2	variance of a data series
t	time
$u(x)$	standard uncertainty of an estimated quantity x
$u^2(x)$	variance of an estimated quantity x
$u_r(x)$	relative standard uncertainty of an estimated quantity x
$u(x, y)$	covariance between two estimated quantities x and y
x, y	quantity under consideration
$\langle x \rangle$	mean value of several quantities x_i
X	analyte under consideration
Δ^2	mean-square successive difference of a data series

4 Bias related to instrumental drift

4.1 Principle

This clause specifies procedures for investigating potential drift of an analytical system and taking corrective actions, if significant drift is encountered.

If the analytical system is expected to be stable, a “drift-control mixture” is measured on a regular basis. For each specified analyte, the results are recorded on a control chart, and the time series of control data is examined continuously. As long as these data vary at random within established control limits, the analytical system has been demonstrated to be stable. Monotonic decrease or increase in control data indicates drift. As an alternative to visual inspection, a statistical test based on successive differences may be used to detect a significant trend. As soon as drift becomes significant, e.g. when data exceed control limits, or when a significant trend is observed, the analytical system is removed from service. After adjustment and re-calibration, the analytical system is returned to service.

The “drift-control mixture” should contain all analytes currently being measured. Given sufficient information on the response behaviour of the analytical system, a representative subset of analytes may be used.

If the analytical system is known to exhibit significant drift, it can at least be expected that the drift behaviour of the system does not significantly depend upon sample composition, since this would allow for drift correction based on measurements of appropriate “drift-correction mixtures”. This expectation is tested by analysing two mixtures of distinct composition (different concentrations of the analytes under consideration, different matrix compositions) on a regular basis. The results are recorded, and for each analyte the two time-series of drift measurement data are compared in order to investigate whether drift behaviour can be expressed in a concentration-invariant manner. Given this for every analyte under consideration, the two time-series for each analyte are pooled to derive a drift correction. In fortunate cases, a joint drift correction may be used for several or even all analytes.

If the system exhibits significant drift, and if moreover drift characteristics depend upon concentration, drift correction needs to be integrated with calibration. This topic is beyond the scope of this International Standard.

4.2 Stability monitoring

4.2.1 General considerations

A drift-control mixture is required which is typical of the gases for which the analytical procedure is used. The composition of the mixture shall be stable, but the concentration of the specified analytes used for stability monitoring does not have to be established in advance with high accuracy.

An analysis of the drift-control mixture should be carried out with each batch of samples. Its composition is unvarying so the results of this analysis can be used as an indication as to whether the procedure is no longer working satisfactorily or re-calibration of the instrument is necessary, or both.

Stability can be monitored using either concentration data or the corresponding response data.

4.2.2 Use of Shewhart control charts

Before first use, the drift-control mixture is analysed at least 10 times in order for precision data to be calculated. For each specified analyte in the drift-control mixture, the mean concentration (or response) and its standard deviation are calculated. If the within-day standard deviation is less than the between-day standard deviation, then these precision data need to be collected one a day for 10 days.

For each specified analyte in the drift-control mixture, a control chart is constructed with points marked on the Y-axis representing

a) the mean concentration (or response),

b) the mean ± 1 standard deviation,

c) the mean ± 2 standard deviations (warning limits), and

d) the mean ± 3 standard deviations (action limits).

Lines parallel to the X-axis are drawn from these points. Each time the drift-control mixture is analysed, the value is plotted using the X-axis as a time scale. As more information becomes available the means and standard deviations can be updated. This assumes that the analytical system has remained stable. Data which clearly indicate some fault shall not be used to revise the control limits.

The plotted values from the analysis of the control gas are compared with the mean value and the ± 1, ± 2 and ± 3 standard deviation lines. It is assumed that the composition of the drift-control mixture is stable and that the analytical results for this follow the normal distribution. If this is true, then while the system is behaving normally, any individual results for the components of the drift-control mixture may fall outside the warning limits on 1 occasion in 20. This means that if individual results fall outside the warning limits more than just occasionally, this can indicate that either there is a systematic tendency for the results to be too high (or too low) or the random error for measurement of that component has increased. Likewise, individual results may fall outside the action limits on 3 occasions out of 1 000.

ISO 8258 [6] contains the following tests which can be used for indicating the presence of variation:

— one point exceeding ± 3 standard deviations;

— nine points in a row on one side of the mean;

— six points in a row steadily increasing or decreasing;

— fourteen points in a row alternating up and down;

— two out of three points in a row exceeding +2 standard deviations or −2 standard deviations;

— four out of five points in a row exceeding +1 standard deviation or −1 standard deviation;

— fifteen points in a row above and below the mean, but not exceeding ± 1 standard deviation;

— eight points in a row above and below the mean, but all exceeding ± 1 standard deviation.

If any of these tests indicates the presence of variation, then that shall be diagnosed and corrected. If this investigation indicates that there is no fault with the measuring procedure, then re-calibration of the instrument is required.

4.2.3 Statistical trend test

As an alternative to monitoring monotonic increase or decrease, drift-control data may be investigated for trends using a statistical test based on successive differences. Given a time-series of drift-control data x_1, x_2, …, x_N with mean value $\langle x \rangle$, the mean-square successive difference Δ^2 is determined according to

$$\Delta^2 = \left[(x_1 - x_2)^2 + (x_2 - x_3)^2 + \ldots + (x_{N-1} - x_N)^2 \right] / (N - 1) \quad (1)$$

This quantity is compared with the variance s^2 given by

$$s^2 = \left[(x_1 - \langle x \rangle)^2 + (x_2 - \langle x \rangle)^2 + \ldots + (x_N - \langle x \rangle)^2 \right] / (N - 1) \quad (2)$$

If successive values in the series are independent (and moreover from a normal distribution), then $\Delta^2 \approx 2\,s^2$. In case of a trend $\Delta^2 < 2\,s^2$ because successive values are closer than to be expected for values drawn at random from a normal distribution.

For a significance test, the test statistic Δ^2/s^2 is compared with the critical value for the specified length N of the series under investigation and the specified significance level. Values of the test statistic below the critical value indicate a significant trend. In this International Standard a significance level of 95 % or 99 % is recommended. Critical values for these significance levels are given in Table A.1, Annex A.

EXAMPLE Consider a series of drift-control data (carbon monoxide in nitrogen, expressed in mmol/mol): 1,28; 1,30; 1,30; 1,28; 1,26; 1,24; 1,27; 1,27; 1,24; 1,26. For these data the mean-square successive difference Δ^2 is $38 \times 10^{-4}/9$, while the variance s^2 is $40 \times 10^{-4}/9$. Hence the test statistic Δ^2/s^2 takes a value of 0,95. For $N = 10$, the critical value is 0,751 8 for a significance level of 99 % and 1,062 3 for a significance level of 95 %. Therefore, under the assumptions of independence and normality, the drift-control data exhibit a significant trend on the 95 % level, while the trend is not significant at the 99 % level.

Consider now the modified data series generated by interchange of the 3rd and 9th datum: 1,28; 1,30; 1,24; 1,28; 1,26; 1,24; 1,27; 1,27; 1,30; 1,26. For these data, the mean square successive difference Δ^2 is now $98 \times 10^{-4}/9$, while the variance s^2 still is $40 \times 10^{-4}/9$. Hence the test statistic Δ^2/s^2 takes a value of 2,45 which means that the modified data series does not exhibit any indication of drift, under the above assumptions.

For stability monitoring based on regular drift-control measurements, a moving window comprising 10 to 20 data is recommended.

If any of these tests indicates the presence of variation, then that shall be diagnosed and corrected. If this investigation indicates that there is no fault with the measuring procedure, then re-calibration of the instrument is required.

4.3 Drift correction

4.3.1 General considerations

This clause specifies a general method for post-processing analytical data to correct for instrumental drift. For this purpose an analytical system is treated as a “black box”. Here the input is the value of the measurand, i.e. the (true) concentration of the analyte under consideration in the analysed sample, and the output is the measured value of this analyte concentration.

This clause is applicable for absolute methods, i.e. analytical methods where analyte concentration is determined directly, or relative methods where the relationship between measured response and analyte concentration is known. It is also applicable for comparison methods, i.e. analytical methods where the relationship between measured response and analyte concentration is determined empirically by calibration.

Two drift-correction mixtures are required which are typical of the gases for which the analytical procedure is used. Each analyte to be determined by the procedure shall be present in both mixtures, at different levels bracketing an appropriate concentration range. The composition of the mixtures shall be stable, and the concentration of the analytes used for drift investigation shall be known with specified uncertainty.

4.3.2 and 4.3.3 specify two complementary approaches, based on additive and multiplicative drift modelling respectively. Usually only one or neither of these procedures will work. If in a particular case both approaches work, the one with the better performance, i.e. with lower correction uncertainty, should be used.

NOTE 1 For an analytical comparison method, drift correction is sometimes better performed using measured responses instead of analyte concentrations.

NOTE 2 Using reference gas mixtures for drift control has the advantage of providing an “absolute” drift correction, i.e. relative to reference values of analyte concentration. As an alternative, less well-characterized gas mixtures could be used for drift correction relative to analyte concentrations measured at a specified time t_0. The latter approach would, in addition, require a proof of stability of the drift-control mixtures. Secondly, the concentrations measured at t_0 would have to be investigated for bias in a later stage.

4.3.2 Additive drift correction

In this subclause, instrumental drift is presented as an additive bias according to Equation (3):

$$x(t) = x + \delta(x,t) \tag{3}$$

where

$x(t)$ is the measured concentration of the analyte under consideration at time t,

x is the true analyte concentration,

$\delta(x,t)$ is the bias at analyte level x, due to drift at time t.

If, for a given analyte, this bias should be the same for different levels of analyte concentration, it can be corrected by determining the bias obtained on a sample with known concentration of the analyte as a function of time, and subtracting the applicable bias from the results obtained on other samples.

To this end, the concentrations of two mixtures (called A and B) as specified above are measured on a regular basis, and the time-series of results are recorded. For the given analyte, these are x_{A1}, x_{A2}, …., x_{AN} and x_{B1}, x_{B2}, …., x_{BN}. The two time-series are smoothed by interpolation or regression, yielding two curves (or functions) $x_A(t)$ and $x_B(t)$. Given concentration-invariant additive bias for that analyte, these curves should be parallel, with a distance of $x_A(t) - x_B(t) = x_{A,ref} - x_{B,ref}$ where $x_{A,ref}$ and $x_{B,ref}$ are the reference values given for mixture A and B.

If this is true (within experimental variability) the differences $x_{Ai} - x_{A,ref}$ and $x_{Bi} - x_{B,ref}$ are pooled and the combined time-series is smoothed yielding a curve (or function) $\delta(t)$. This curve is then used to correct the result obtained on another mixture M at any time t within the period covered according to

$$x_M = x_M(t) - \delta(t) \tag{4}$$

The standard uncertainty of the corrected result is determined by

$$u^2(x_M) = u^2\left[x_M(t)\right] + u^2\left[\delta(t)\right] \tag{5}$$

In this uncertainty budget, the first term is obtained from the uncertainty budget of the analytical procedure. The second term is estimated from the residual scattering of the pooled differences used to determine the correction curve $\delta(t)$. In addition, the uncertainty contributions of the drift measurements and the composition of the drift-correction mixtures are included if significant.

If bias corrections should be approximately the same for different analytes, then a joint correction may be derived for a group of such analytes from pooled time-series.

4.3.3 Multiplicative drift correction

In this subclause, instrumental drift is presented as a recovery factor according to Equation (6):

$$x(t) = Q(x,t)\,x \tag{6}$$

where

$x(t)$ is the measured concentration of the analyte under consideration at time t;

x is the true analyte concentration;

$Q(x,t)$ is the recovery factor at analyte level x, due to drift at time t.

If, for the given analyte, this recovery factor should be the same for different levels of analyte concentration, it can be corrected by determining the recovery on a sample with known concentration of the analyte as a function of time, and dividing the results obtained on other samples by the applicable recovery factor.

To this end, two mixtures (called A and B) as specified above are measured on a regular basis, and the time series of results are recorded. For the given analyte , these are x_{A1}, x_{A2}, …., x_{AN} and x_{B1}, x_{B2}, …., x_{BN}. The two time-series are smoothed by interpolation or regression, yielding two curves (or functions) $x_A(t)$ and $x_B(t)$. Given concentration-invariant recovery for that analyte, these curves should be proportional, with a proportionality factor $x_A(t)/x_B(t) = x_{A,ref}/x_{B,ref}$ where $x_{A,ref}$ and $x_{B,ref}$ are the reference values given for mixture A and B.

If this is true (within experimental variability), the quotients $x_{Ai}/x_{A,ref}$ and $x_{Bi}/x_{B,ref}$ are pooled and the combined time series is smoothed yielding a curve (or function) $Q(t)$. This curve is then used to correct the result obtained on another mixture M at any time t within the period covered according to

$$x_M = \frac{x_M(t)}{Q(t)} \tag{7}$$

EXAMPLE Consider two drift-control mixtures (carbon monoxide in nitrogen) A and B with established analyte concentrations $c_{A,ref}$ = 1,295 ± 0,006 mmol/mol and $c_{B,ref}$ = 21,65 ± 0,15 μmol/mol (where standard uncertainties are specified). Using an automated system, these gases are analysed alternately every 4 h between batches of samples with CO concentrations in the range from 10 μmol/mol to 5 mmol/mol. For a measurement campaign over 80 h, drift control data were obtained as shown in Table 1 [elapsed time $t_i - t_0$ in hours, concentration $c_A(t_i)$ for gas A measured at time t_i in mmol/mol, concentration $c_B(t_i)$ for gas B measured at time t_i in μmol/mol].

Table 1 — Time-series of drift-control data

$t_i - t_0$	0 h	4 h	8 h	12 h	16 h	20 h	24 h	28 h	32 h	36 h	40 h
$c_A(t_i)$ mmol/mol	1,28		1,30		1,30		1,28		1,26		1,24
$c_B(t_i)$ μmol/mol		21,7		21,3		21,0		21,0		21,0	
$t_i - t_0$	44 h	48 h	52 h	56 h	60 h	64 h	68 h	72 h	76 h	80 h	
$c_A(t_i)$ mmol/mol		1,27		1,27		1,24		1,26		1,25	
$c_B(t_i)$ μmol/mol	21,3		20,9		21,2		20,6		21,0		

The times-series of measurements on gas A and gas B are smoothed (sm) using ordinary linear least squares regression (as available in common software). This gives two straight-line equations

$c_{A,sm}(t) = 1{,}291 - 5{,}682 \times 10^{-4}(t - t_0)$ for gas A

$c_{B,sm}(t) = 21{,}41 - 7{,}727 \times 10^{-3}(t - t_0)$ for gas B

The regression lines give the impression of a slightly decreasing detector sensitivity.

The next step is to investigate whether the two drift-control data series may be combined to obtain a common drift correction. For this purpose, the approach using recovery factors looks promising. According to the specified procedure, the quotients $c_{A,sm}(t)/c_{B,sm}(t)$ of smoothed drift-control data for gases A and B are examined for significant departure from the quotient $c_{A,ref}/c_{B,ref}$ of reference data for these gases. An equivalent test is to compare the smoothed recoveries $c_{A,sm}(t)/c_{A,ref}$ and $c_{B,sm}(t)/c_{B,ref}$ for significant differences. Using the regression equations obtained previously, the difference $d_{A,B}(t) = [(c_{A,sm}(t)/c_{A,ref}) - (c_{B,sm}(t)/c_{B,ref})]$ of recoveries at time t is given by

$$d_{A,B}(t) = 0{,}007\,97 - 8{,}183 \times 10^{-5}(t - t_0)$$

Table 2 records values of this difference at specified times t, and the standard uncertainty of these values.

Table 2 — Differences of smoothed recovery data

$t - t_0$	0 h	10 h	20 h	30 h	40 h	50 h	60 h	70 h	80 h
$d_{A,B}(t)$	0,008 0	0,007 2	0,006 3	0,005 5	0,004 7	0,003 9	0,003 1	0,002 2	0,001 4
$u[d_{A,B}(t)]$	0,012 7	0,011 5	0,010 5	0,009 8	0,009 5	0,009 7	0,010 4	0,011 3	0,012 6

The (standard) uncertainty of the differences $d_{A,B}(t)$ is obtained by uncertainty propagation from the uncertainties of the calculated data $c_{A,sm}(t)$, $c_{B,sm}(t)$ and the uncertainties of the reference data $c_{A,ref}$, $c_{B,ref}$. Due to the fact that the recoveries are close to unity, relative uncertainties may be used instead of absolute uncertainties, yielding

$$u^2[d_{A,B}(t)] \approx u_r^2[c_{A,sm}(t)] + u_r^2(c_{A,ref}) + u_r^2[c_{B,sm}(t)] + u_r^2(c_{B,ref})$$

Here $u_r[c_{A,sm}(t)]$ and $u_r[c_{B,sm}(t)]$ are obtained from the confidence interval for the regression line (where, depending on the software used, division by a default t-factor may be necessary), while $u_r(c_{A,ref})$ and $u_r(c_{B,ref})$ are obtained from the specification of the drift-control gases.

Examining the data in Table 2, no significant differences (i. e. $d > 2u[d]$) between smoothed recoveries for gases A and B are encountered. Therefore the experimental recovery data $Q_A(t_i) = c_A(t_i)/c_{A,ref}$ and $Q_B(t_i) = c_B(t_i)/c_{B,ref}$ are pooled, and the combined time series is smoothed by ordinary least-squares regression, yielding a linear equation for a mean recovery factor $Q(t)$ as follows.

$$Q(t) = 0{,}993\,2 - 4{,}037 \times 10^{-4}(t - t_0)$$

This equation would then be used to correct measurements on other samples carried out in the course of the measuring campaign.

The relative standard uncertainty of the corrected result is determined by

$$u_r^2\left(x_M\right) = u_r^2\left[x_M(t)\right] + u_r^2\left[Q(t)\right] \tag{8}$$

In this uncertainty budget, the first term is obtained from the uncertainty budget of the analytical procedure. The second term is estimated from the residual scattering of the pooled quotients used to determine the recovery factor $Q(t)$. In addition, the uncertainty contributions of the drift measurements and the composition of the drift-correction mixtures are included if significant.

If the recovery is approximately the same for different analytes, then a joint recovery factor may be derived for a group of such analytes from pooled time series.

5 Bias related to effects of sample composition

5.1 Principle

5.1.1 Basics

This clause specifies methods and protocols for investigation and handling of the bias of an analytical procedure due to effects of sample composition, within a specified measuring range. Throughout this clause, it is assumed that the measuring system under investigation is stable. Where appropriate, it should be confirmed that this is the case. The measuring range normally includes variations in the concentration of the analyte under consideration as well as variations in matrix composition. In this clause, the focus is on the investigation of bias as a function of analyte concentration. Bias investigation and handling may include effects due to within-specification variations in matrix composition. If matrix interferences are significant, corrections derived from the bias study data include averaged matrix contributions, and the uncertainty on the corrections includes contributions associated with the spread of biases due to these interferences. Principles for systematic investigation of matrix interferences are considered in Clause 6.

This clause is applicable for absolute methods, i.e. analytical methods where analyte concentration is determined directly, or where the relationship between measured response and analyte concentration is known. It is also applicable for comparison methods, i.e. analytical methods where the relationship between measured response and analyte concentration is determined empirically by calibration. If this calibration is performed according to ISO 6143 [7], and if the analytical procedure is used strictly within specification, then no bias investigation is necessary as long as the analytical system is stable. However, if the calibration is performed by a less rigorous procedure (e.g. single-level calibration), or if a procedure calibrated according to ISO 6143 [7] is used out of its original specification (range of analyte concentration, matrix composition), bias shall be investigated, and accounted for if significant.

Two different cases are considered:

a) Case A — Intra-laboratory assessment of a fully developed uncertainty budget

   An analytical procedure is investigated for bias as a final step after an exhaustive evaluation of measurement uncertainty. The aim of this investigation is to test the assumption that the uncertainty evaluation has properly accounted for all relevant random and systematic effects impacting the measurement result.

b) Case B — Estimation of measurement uncertainty from intra-laboratory validation data

   An analytical procedure is investigated for bias in addition to, or jointly with a precision study. The aim of these investigations is to obtain an estimate of measurement uncertainty by combination of bias and precision estimates.

5.1.2 Main steps involved in investigation and handling of bias

Potential bias is investigated by comparing results obtained on “known samples” with the corresponding reference values, and by examining whether any of the differences encountered are significant in comparison with the relevant uncertainty on that difference. If no significant bias (difference) is found, and provided that the samples are representative for the specified measuring range and the reference values are well established, then the analytical procedure has thereby been demonstrated to be unbiased. Depending on the type of bias study, no further action is required in case A (in addition to regular quality control), while in case B the observed bias and its associated uncertainty are included in the estimation of measurement uncertainty.

If significant bias is found, this requires corrective action. Depending on whether the bias is judged to be technically serious or acceptable, different actions are applicable.

For the purpose of this International Standard, actions on bias are recommended as shown in Table 3:

Table 3 — Recommended actions on bias

Bias	Case A	Case B
Serious bias	Examine and amend the analytical procedure to remove/reduce bias; or examine and amend the uncertainty budget for missing or underestimated uncertainty components.	Examine and amend the analytical procedure to remove/reduce bias.
Acceptable bias	Apply a correction for bias in the data evaluation procedure; or include an allowance for uncorrected bias in the uncertainty estimation.	Apply a correction for bias in the data evaluation procedure; or include an allowance for uncorrected bias in the uncertainty estimation.
Insignificant bias	(no action)	Include an allowance for uncorrected bias in the uncertainty estimation.

Criteria for the assessment of bias (significant/insignificant and serious/acceptable) and procedures for:

— correcting for bias and accounting for the uncertainty associated with a bias estimate,

— including an allowance for uncorrected bias in the uncertainty estimation,

are specified in subsequent clauses.

5.2 deals with local bias handling using a single reference sample, with separate subclauses for case A (5.2.1) and case B (5.2.2), or several matrix-varied reference samples (5.2.3). Bias handling for extended measuring ranges, using several reference samples, is addressed in 5.3.

By way of the procedures specified in this clause, the measurement uncertainty for an analytical procedure is made traceable to the reference values, including uncertainties, attributed to the reference samples (most often calibration gas mixtures). To this end it is important to ensure that

— the reference values and their uncertainties are well established,

— the reference samples are representative for the range of samples to be analysed,

— the variability of measuring conditions in the bias study covers the variability in the intended applications.

Normally bias investigations are repeated on a regular basis. If this is the case, then it is important to compare the data from the current investigation with those from previous ones. If the data are compatible, they may be pooled to improve the statistical basis of the estimates concerned (i.e. average deviations, average recoveries and related standard deviations). If they are not compatible, then useful information may be obtained from an investigation of the discrepancies. Corrective actions on observed bias should not be “piled”, as this would considerably complicate the data evaluation. Therefore, repeated bias investigations should always utilise the same measurement procedure, without any corrections or uncertainty allowances derived from previous investigations.

5.2 Local bias handling

5.2.1 Single reference sample — Case A

5.2.1.1 General

This subclause refers to the case where an analytical procedure is investigated for bias as a final step after an exhaustive evaluation of measurement uncertainty. The aim of this investigation is to test the assumption that

the uncertainty evaluation has properly accounted for all relevant random and systematic effects impacting the results.

For a local investigation of potential bias of a candidate analytical procedure, the procedure is applied to a reference sample, selected to represent the specified measuring range, and the results from (a minimum of n = 6) repeated measurements, obtained under appropriate "within-laboratory reproducibility conditions" (see ISO 5725-3 [8]) are compared with the reference value attributed to the reference sample. Alternatively, the candidate procedure and a reference procedure may be applied in parallel to an appropriate sample, and the results of the candidate procedure compared with those of the reference procedure.

The conclusions drawn from this comparison are only valid if it can be taken for granted that the results obtained on the reference sample are valid for the entire measuring range, that is, for the variety of samples to be analysed in the future. If this is not the case, there are two possibilities as follows: Either the measuring range has to be restricted accordingly, or a 2-level or a multi-level procedure has to be used (see 5.2.3 and 5.3).

In the description of the procedure, the following symbols, relating to the reference sample, are used:

x_{ref} the reference value of the measured quantity;

$u(x_{\text{ref}})$ the standard uncertainty of the reference value;

x_{obs} the result of a measurement using the candidate procedure (n replicates);

$u(x_{\text{obs}})$ the standard uncertainty of a measurement result x_{obs};

$\langle x_{\text{obs}} \rangle$ the mean value of a series of n replicates x_{obs};

s_{obs} the standard deviation of a series of n replicates x_{obs}.

5.2.1.2 Step 1 — Checking precision

As a first step in assessing the results obtained on the reference sample, the standard deviation s_{obs} is examined for compliance with the uncertainty estimate $u(x_{\text{obs}})$. Evidently s_{obs} only accounts for that part of the measurement uncertainty $u(x_{\text{obs}})$ which is due to variations in influence quantities taking place between replicate measurements. Assessment of s_{obs} therefore requires a decomposition of the measurement uncertainty for x_{obs} as follows:

$$u^2\left(x_{\text{obs}}\right) = u^2_{\text{c,var}}\left(x_{\text{obs}}\right) + u^2_{\text{c,inv}}\left(x_{\text{obs}}\right) \tag{9}$$

where

$u_{\text{c,var}}(x_{\text{obs}})$ is the combined standard uncertainty accounting for all influence quantities which are effectively varied between replications;

$u_{\text{c,inv}}(x_{\text{obs}})$ is the combined standard uncertainty accounting for all influence quantities which are effectively invariant under replication conditions.

Given the decomposition of $u(x_{\text{obs}})$, compliance requires that $s_{\text{obs}} \approx u_{\text{c,var}}(x_{\text{obs}})$, and an F-test [most often with infinite degrees of freedom for $u_{\text{c,var}}(x_{\text{obs}})$] may be made to formalise this.

To this end, the uncertainty component $u_{\text{c,var}}(x_{\text{obs}})$ may be either estimated from the uncertainty budget or taken directly from appropriate quality control data: the intermediate-precision standard deviation s_{IR} obtained from monitoring the within-laboratory reproducibility of the procedure (see ISO 5725-3 [8]), e.g. taken from a control chart.

NOTE Given a valid uncertainty budget, the measurement uncertainty estimate $u(x_{obs})$ encompasses the intermediate-precision standard deviation, i.e. $s_{IR} < u(x_{obs})$.

5.2.1.3 Step 2 — Testing significance of observed bias

Next, given a positive result for the check on s_{obs}, the deviations $\delta = x_{obs} - x_{ref}$ obtained by replicate measurements on the reference sample are examined. An individual deviation is rated significant if it exceeds the expanded uncertainty $U(\delta)$ of that difference. Utilizing $U = k \times u$ with $k = 2$ (for a confidence level of about 95 %) and calculating $u(\delta)$ according to $u^2(\delta) = u^2(x_{obs}) + u^2(x_{ref})$, an individual deviation is insignificant if $|\delta| \leqslant 2u(\delta)$, and significant if $|\delta| > 2u(\delta)$. Accordingly, if all individual deviations are insignificant, the overall bias is rated insignificant. If significant deviations are found, the overall bias is rated significant. In cases of weak significance, e.g. with only one deviation δ slightly exceeding the critical value $2u(\delta)$, the final decision concerning significance of bias may be made by examining the average deviation $\langle\delta\rangle = \langle x_{obs} - x_{ref}\rangle = \langle x_{obs}\rangle - x_{ref}$ for significant departure from zero, i.e. as to whether $|\langle\delta\rangle| > 2u(\langle\delta\rangle)$. The standard uncertainty $u(\langle\delta\rangle)$ required for this test is calculated according to $u^2(\langle\delta\rangle) = u^2(\langle x_{obs}\rangle) + u^2(x_{ref})$. Estimation of the standard uncertainty $u(\langle x_{obs}\rangle)$ needed for this purpose requires particular attention, due to correlations which have to be taken into account, see Annex B.

NOTE 1 The test criterion makes use of a coverage factor of $k = 2$, aiming at a significance level of approximately 95 %. For this purpose, use of Student´s factor t is common. However, replacement of the factor 2 by Student´s t, using the calculus of effective degrees of freedom specified in the GUM [3], would require a proper estimate of the degrees of freedom for individual deviations δ or average deviations $\langle\delta\rangle$ which are difficult to obtain.

NOTE 2 Instead of testing deviations $\delta = x_{obs} - x_{ref}$ for significant departure from zero, recoveries $Q = x_{obs}/x_{ref}$ can be tested for significant departure from unity. These two tests are essentially equivalent.

EXAMPLE Consider an analytical procedure for measuring carbon monoxide in nitrogen in a low 10^{-3} range, e.g. 0,5 mmol/mol to 5 mmol/mol. Potential bias is investigated using a single reference gas mixture with carbon monoxide content x_{ref} = 1,295 mmol/mol and standard uncertainty $u(x_{ref})$ = 0,006 mmol/mol. Ten measurements of the reference gas mixture are carried out under appropriate (intermediate) reproducibility conditions, yielding carbon monoxide contents (in mmol/mol) of 1,28; 1,30; 1,24; 1,28; 1,26; 1,24; 1,27; 1,27; 1,30; 1,26. The mean value of these data is $\langle x_{obs}\rangle$ = 1,27 mmol/mol, and the standard deviation is s_{obs} = 0,021 mmol/mol.

Precision monitoring using quality control samples in this range yielded standard deviations of about 2 % relative, and the value obtained for s_{obs} agrees well with that estimate.

An uncertainty budget for the procedure is available, including random effects such as variations in the amount of sample injected into the analytical system and variations in the instrumental response to the analyte, and systematic effects such as deviations of the analyte concentration in the gas mixtures used to calibrate the analyser. This gives a combined relative standard uncertainty of 2,4 %, with 2,1 % attributed to effects considered random, and another 1,2 % attributed to effects considered systematic, at the specified intermediate-reproducibility conditions. The value obtained for s_{obs} and the estimate of 2 % for the intermediate-precision standard deviation agree well with the estimate of 2,1 % for $u_{c,var}(x_{obs})$.

The largest deviation of the measured data from the reference value is δ_{max} = – 0,055 mmol/mol, while the expanded uncertainty associated with this difference is $2 \times [(0{,}024 \times 1{,}27)^2 + 0{,}36 \times 10^{-4}]^{1/2} = 2 \times 3{,}0 \times 10^{-2} = 0{,}06$ (unit: mmol/mol). Thus the observed difference is just below the critical value for significance.

The average deviation of the measured data from the reference value is $\langle\delta\rangle = \langle x_{obs}\rangle - x_{ref}$ = – 0,025 mmol/mol. Utilizing the uncertainty estimate for a mean value in Annex B, the expanded uncertainty associated with this deviation is obtained as $2 \times [(0{,}021 \times 1{,}27)^2/10 + (0{,}012 \times 1{,}27)^2 + 0{,}36 \times 10^{-4}]^{1/2} = 2 \times 1{,}8 \times 10^{-2} = 0{,}036$ mmol/mol. Thus the observed average deviation is insignificant, too.

In summary, no significant bias was observed, and this finding confirms the assumption that the analytical procedure and uncertainty budget provide correct results with a valid uncertainty. Therefore no further action is necessary in addition to regular quality control.

Examples for the application of a bias correction including the estimation of correction uncertainty, and for bias handling without correction, are given in 5.2.2.

5.2.1.4 Step 3 — Actions on bias

5.2.1.4.1 General

If no significant bias was observed, this finding confirms the assumption that the analytical procedure and uncertainty budget provide correct results with a valid uncertainty. Therefore no further action is necessary in addition to regular quality control.

If significant bias was observed, the next step is to decide whether the bias is technically serious or still acceptable.

a) **Serious deviations** indicate serious deficiencies of the analytical procedure or of the uncertainty budget, necessitating thorough review and amendment. The criterion for that is a matter of expert judgement, based on fitness-for-purpose considerations (e.g. concerning the target uncertainty) and previous experience with the method. In addition, the ratio $\langle\delta\rangle/u(x_{obs})$ should be in a reasonable range.

   Recommended actions are to

   — examine and amend the analytical procedure to remove/reduce bias, or

   — examine and amend the uncertainty budget for missing or underestimated uncertainty components.

b) **Acceptable deviations** fall within the range expected by expert judgement. As a consequence, amendment of the analytical procedure is not necessary, but the observed bias shall be accounted for in the data evaluation procedure.

   Recommended actions are to

   — apply a correction for bias in the data evaluation procedure, or

   — include an allowance for uncorrected bias in the uncertainty budget.

In case of doubt as to whether the available data and other relevant information enable a technically sound correction, an allowance for bias in the uncertainty estimation is recommended.

5.2.1.4.2 Step 3a — Correction for bias

The 1-level correction, as described in this clause, is only admissible if it can be taken for granted that either the absolute errors of measurement, or the relative errors of measurement, are constant over the entire measuring range under consideration.

For a clear distinction from the symbols for values referring to the reference sample, the symbol y is used for values referring to the subsequent analysis of test samples.

In the case of constant absolute errors, significant bias is corrected by subtracting the average deviation $\langle\delta\rangle$ determined in the bias study from the raw result y_{meas} of a measurement on a test sample according to

$$y_{corr} = y_{meas} - \langle\delta\rangle \tag{10}$$

In the case of constant relative errors, significant bias is corrected by way of dividing the raw result y_{meas} of a measurement on a test sample through the average recovery $\langle Q\rangle$ determined in the bias study according to

$$y_{corr} = \frac{y_{meas}}{\langle Q\rangle} \tag{11}$$

where the average recovery is given by $\langle Q\rangle = \langle x_{obs}\rangle/x_{ref}$.

The correction can be carried out either by adjustment of the measuring system, referring to the zero point or to the sensitivity, respectively, or by recalculation.

The standard uncertainty of the final measurement result after bias correction, y_{corr}, is calculated from the uncertainty of the uncorrected measurement result, y_{meas}, and the uncertainty of the correction, according to the rules of uncertainty propagation, as described in the GUM [3].

For correction using an average deviation, the standard uncertainty of corrected measurement results is determined by

$$u^2\left(y_{corr}\right) = u^2\left(y_{meas}\right) + u^2\left(\left\langle x_{obs}\right\rangle\right) - 2u\left(y_{meas}, \left\langle x_{obs}\right\rangle\right) + u^2\left(x_{ref}\right) \tag{12}$$

where

$u(y_{meas})$ standard uncertainty of the measurement result y_{meas} obtained on a test sample, as calculated from the uncertainty budget;

$u(\langle x_{obs}\rangle)$ standard uncertainty of the mean value x_{obs} of the results obtained on the reference sample;

$u(y_{meas}, \langle x_{obs}\rangle)$ covariance between y_{meas} and $\langle x_{obs}\rangle$;

$u(x_{ref})$ standard uncertainty of the reference value x_{ref} attributed to the reference sample.

NOTE 1 The covariance term arises because y_{meas} and $\langle x_{obs}\rangle$ are determined using the same measurement procedure and thus share common uncertainty sources.

Most often, neither the standard uncertainty $u(\langle x_{obs}\rangle)$ nor the covariance $u(y_{meas}, \langle x_{obs}\rangle)$ will be readily accessible from the uncertainty budget of the analytical procedure, and rigorous estimation will not be feasible. Therefore, in this International Standard, approximations using the intermediate-precision standard deviation s_{IR} are used, with details given in Annex B. Utilizing these approximations, the standard uncertainty of analytical results corrected by an average deviation is obtained as follows:

$$u^2\left(y_{corr}\right) = \frac{n+1}{n} s_{IR}^2 + u^2\left(x_{ref}\right) \tag{13}$$

For correction using an average recovery, an analogous equation applies, with the standard uncertainties $u(y_{corr})$, $u(x_{ref})$ and the intermediate-precision standard deviation s_{IR} replaced by relative standard uncertainties and the relative standard deviation.

NOTE 2 Remarkably, the uncertainty $u(y_{corr})$ of a measurement result corrected for significant bias, as given by Equation (13), will often be smaller than the original uncertainty $u(y_{meas})$, which is retained if no significant bias was observed. This gain is due to the compensation between uncertainty components common to $u(y_{meas})$ and $u(x_{obs})$, accounting for compensation of systematic effects shared by analyses of test sample and reference sample; an effect that may call for a second thought about whether to restrict bias correction to cases where significant bias was found. Re-considering this issue, bias correction could be conceived as a modification of the original procedure, whereby the measurements are made relative to a standard in order to eliminate bias and/or uncertainty components related to systematic effects. Adopting this view, testing for significance of bias is obsolete, and bias correction would be carried out in any case, significant or not. For this purpose, however, one would rather prefer a design involving replicate analyses of both, the test sample and the reference sample. The most simple design would be an alternating sequence yielding $y = \langle y_{meas} - x_{meas}\rangle + x_{ref}$. Assuming perfect cancellation of (constant additive) bias, the standard uncertainty of y could be estimated by the root sum of squares of the standard deviation of the differences $y_{meas} - x_{meas}$ and the standard uncertainty of x_{ref}. Uncertainty estimation based on a comprehensive uncertainty budget, properly addressing correlation of measurements, should give an equivalent result.

5.2.1.4.3 Step 3b — Bias handling without correction

In practice it may happen that significant bias is found on the reference sample, but the data are not sufficient for deriving a sound correction. For example, it may be doubtful whether a single-level correction, based on

measurements of a single reference sample, is applicable to the entire measuring range or, if so, whether correction utilising an average deviation or correction utilising an average recovery is more appropriate. Then additional reference samples should be investigated in order to characterise the bias to an appropriate degree. If this is not possible or not practical, no correction is performed. Instead, an allowance is made in the uncertainty estimation to account for the observed bias.

Currently a range of different approaches to account for uncorrected bias has been proposed, and a generally accepted procedure has not yet emerged; see Annex B for an overview. In this International Standard, the approach by Lira and Woeger [9] is adopted:

If, in case of significant bias, another estimate y_{est} is used for the measurand instead of the corrected measurement result y_{corr}, then the standard uncertainty associated with y_{est} is determined by

$$u^2(y_{\text{est}}) = u^2(y_{\text{corr}}) + (y_{\text{corr}} - y_{\text{est}})^2 \tag{14}$$

In the deviation mode, if significant bias was found, but no correction (by $\langle\delta\rangle$) is applied to the raw measurement result y_{meas}, then y_{est} is obtained from y_{meas} by subtracting zero.

$$y_{\text{est}} = y_{\text{meas}} \tag{15}$$

The standard uncertainty associated with this estimate is determined by

$$u^2(y_{\text{est}}) = u^2(y_{\text{corr}}) + \langle\delta\rangle^2 \tag{16}$$

In the recovery mode, an analogous equation applies, involving relative instead of absolute quantities.

Often the use of a constant absolute uncertainty allowance $\langle\delta\rangle^2$ or a constant relative uncertainty allowance $\langle Q\rangle^2$, obtained on a single reference sample, over extended concentration ranges will be doubtful. Instead, for wider concentration ranges, estimation of bias-related uncertainty should rather be based on the empirical fact that concentration dependence of analytical uncertainty is between u = constant and $u \sim$ concentration. Therefore a safe approximation is to use the uncertainty estimate obtained on the reference sample as a constant absolute value for $y_{\text{meas}} < \langle x_{\text{obs}}\rangle$ and as a constant relative value for $y_{\text{meas}} \geqslant \langle x_{\text{obs}}\rangle$. This gives

$$u^2(y_{\text{est}}) = u^2(y_{\text{corr}}) + \langle\delta\rangle^2 \text{ for } y_{\text{meas}} < \langle x_{\text{obs}}\rangle \tag{17}$$

and

$$u^2(y_{\text{est}}) = \left(\frac{y_{\text{meas}}}{\langle x_{\text{obs}}\rangle}\right)^2 \left[u^2(y_{\text{corr}}) + \langle\delta\rangle^2\right] \text{ for } y_{\text{meas}} \geqslant \langle x_{\text{obs}}\rangle \tag{18}$$

5.2.2 Single reference sample — Case B

5.2.2.1 General

This subclause refers to the case where an analytical procedure is investigated for bias in addition to, or jointly with a precision study. The aim of these investigations is to obtain an estimate of measurement uncertainty by combination of bias and precision estimates.

For a local investigation of potential bias of a candidate analytical procedure, the procedure is applied to a reference sample, selected to represent the specified measuring range, and the results from (a minimum of n = 6) repeated measurements, obtained under appropriate "within-laboratory reproducibility conditions" (see ISO 5725-3 [8]) are compared with the reference value attributed to the reference sample. Alternatively, the candidate procedure and a reference procedure may be applied in parallel to an appropriate sample, and the results of the candidate procedure compared with those of the reference procedure.

The conclusions drawn from this comparison are only valid if it can be taken for granted that the results obtained on the reference sample are valid for the entire measuring range, that is, for the variety of samples to be analysed in the future. If this is not the case, there are two possibilities, as follows: either the measuring range shall be restricted accordingly, or a 2-level or a multi-level procedure shall be used (see 5.2.3 and 5.3).

In the description of the procedure, the following symbols, relating to the reference sample, are used:

x_{ref} the reference value of the measured quantity;

$u(x_{ref})$ the standard uncertainty of the reference value;

x_{obs} the result of a measurement using the candidate procedure (n replicates);

$\langle x_{obs} \rangle$ the mean value of a series of n replicates x_{obs};

s_{obs} the standard deviation of a series of n replicates x_{obs}.

5.2.2.2 Step 1 — Checking precision

In constructing an estimate of measurement uncertainty using data from precision and bias investigations carried out within a laboratory, the first step should be an estimation of the intermediate-precision standard deviation s_{IR}, referring to appropriate within-laboratory reproducibility conditions, see ISO 5725-3 [8]. This will be obtained from replicate analyses of selected test samples as a standard deviation taken from a control chart, or as a pooled standard deviation if several test samples are utilised. Given that, s_{obs} is examined for compliance with s_{IR}, i. e. whether $s_{obs} \leqslant s_{IR}$ (at least s_{obs} should not be significantly larger than s_{IR}).

5.2.2.3 Step 2 — Assessing observed bias

Given a positive result in the check on s_{obs}, the average deviation $\langle \delta \rangle = \langle x_{obs} - x_{ref} \rangle = \langle x_{obs} \rangle - x_{ref}$ is examined as to whether this deviation is technically serious or acceptable, and (in case of an acceptable deviation) whether this deviation is statistically significant or not.

a) **Serious deviations** indicate serious deficiencies of the analytical procedure, necessitating thorough review and amendment of the protocol and equipment. The criterion for that is a matter of expert judgement, based on fitness-for-purpose considerations (e.g. concerning the target uncertainty) and previous experience with the method. In addition, the ratio $\langle \delta \rangle / s_{IR}$ should be in a reasonable range.

b) **Acceptable deviations** fall within the range expected by expert judgement. As a consequence, amendment of the measurement procedure is not necessary, but a correction for bias or an allowance for bias in the uncertainty estimation has to be made. In case of doubt as to whether the available data and other relevant information enable a technically sound correction, an allowance for bias in the uncertainty estimation is recommended.

   As a subclass of acceptable deviations, **insignificant deviations** are not significantly different from zero, the criterion being that $|\langle \delta \rangle| \leqslant 2u(\langle \delta \rangle)$. The standard uncertainty required for this test is calculated according to $u^2(\langle \delta \rangle) = s_{obs}^2/n + u^2(x_{ref})$, where n denotes the number of replicate measurements of the reference sample. For insignificant deviations bias correction does not make much sense. Instead, an allowance for bias in the uncertainty estimation is recommended.

NOTE 1 The test criterion makes use of a coverage factor of $k = 2$, aiming at a significance level of approximately 95 %. For this purpose, use of Student´s factor t is common. However, replacement of the factor 2 by Student´s t, using the calculus of effective degrees of freedom specified in the GUM, would require a proper estimate of the degrees of freedom for an average deviation $\langle \delta \rangle$ which is difficult to obtain.

NOTE 2 Instead of testing deviations $\delta = x_{obs} - x_{ref}$ for significant departure from zero, recoveries $Q = x_{obs}/x_{ref}$ could be tested for significant departure from unity. These two tests are essentially equivalent.

EXAMPLE Consider an analytical procedure for measuring carbon monoxide (CO) in nitrogen in a low 10^{-3} range, e.g. 0,5 mmol/mol to 5 mmol/mol. Potential bias is investigated using a single reference gas mixture with CO content x_{ref} = 1,295 mmol/mol and standard uncertainty $u(x_{ref})$ = 0,006 mmol/mol. Ten measurements of the reference gas mixture are carried out under appropriate (intermediate) reproducibility conditions, yielding CO contents (in mmol/mol) of 1,28; 1,30; 1,24; 1,28; 1,26; 1,24; 1,27; 1,27; 1,30; 1,26. The mean value of these data is $\langle x_{obs} \rangle$ = 1,27 mmol/mol, and the standard deviation is s_{obs} = 0,021 mmol/mol.

Precision monitoring using quality control samples in this range yielded standard deviations of about 2 % relative, and the value obtained for s_{obs} agrees well with that estimate.

The average deviation of the measured data from the reference value is $\langle \delta \rangle = \langle x_{obs} \rangle - x_{ref}$ = – 0,025 mmol/mol, while the expanded uncertainty associated with this difference is $2 \times [(0{,}45 \times 10^{-4}) + (0{,}36 \times 10^{-4})]^{1/2} = 2 \times 0{,}9 \times 10^{-2}$ = 0,018 (unit: mmol/mol). Thus the observed difference is significant, i.e. a significant, but certainly still acceptable, bias is diagnosed.

5.2.2.4 Step 3 — Actions on bias

5.2.2.4.1 Step 3a — Correction for bias

The 1-level correction, as described in this clause, is only admissible if it can be taken for granted that either the absolute errors of measurement, or the relative errors of measurement, are constant over the entire measuring range under consideration.

For a clear distinction from the symbols for values referring to the reference sample, the symbol y is used for values referring to the subsequent analysis of test samples.

In the case of constant absolute errors, significant bias is corrected by subtracting the average deviation $\langle \delta \rangle$ determined in the bias study from the raw result y_{meas} of a measurement on a test sample according to

$$y_{corr} = y_{meas} - \langle \delta \rangle \tag{19}$$

In the case of constant relative errors, significant bias is corrected by way of dividing the raw result y_{meas} of a measurement on a test sample through the average recovery $\langle Q \rangle$ determined in the bias study according to

$$y_{corr} = \frac{y_{meas}}{\langle Q \rangle} \tag{20}$$

where the average recovery is given by $\langle Q \rangle = \langle x_{obs} \rangle / x_{ref}$.

The correction can be carried out either by adjustment of the measuring system, referring to the zero point or to the sensitivity respectively, or by re-calculation.

The standard uncertainty of the final measurement result after bias correction, y_{corr}, is calculated from the uncertainty of the uncorrected measurement result, y_{meas}, and the uncertainty of the correction, according to the rules of uncertainty propagation, as described in the GUM [3].

For correction using an average deviation, the standard uncertainty of corrected measurement results is determined by

$$u^2(y_{corr}) = s^2(y_{meas}) + s^2(\langle x_{obs} \rangle) + u^2(x_{ref}) = s_{IR}^2 + \frac{s_{obs}^2}{n} + u^2(x_{ref}) \tag{21}$$

In Equation (21), the intermediate-precision standard deviation is used to estimate the precision of y_{meas}. If y_{meas} is the mean value of m replicate measurement results, the term s_{IR}^2 has to be replaced by s_{IR}^2/m. Moreover, s_{obs} may be replaced by s_{IR} to obtain a more robust estimate.

NOTE 1 In 5.2.1 the estimation of correction uncertainty includes a covariance term $u(y_{meas}, \langle x_{obs} \rangle)$ to account for correlations due to common systematic effects on measurements of a test sample and measurements of the reference sample. This is not the case here, because such effects are not included in the precision estimates used in Equation (21).

NOTE 2 Remarkably, the uncertainty estimate according to Equation (21) coincides with the case-A estimate for the uncertainty of corrected measurement results that was derived starting from a comprehensive uncertainty budget, see 5.2.1, Equation (13).

For correction using an average recovery, an analogous equation applies, with the standard uncertainties $u(y_{corr})$, $u(x_{ref})$ and the intermediate-precision standard deviation s_{IR} replaced by relative standard uncertainties and the relative standard deviation.

EXAMPLE Continuing the example in Step 2, the decision then is taken whether the available data, obtained with a single reference gas mixture, is sufficient for deriving a correction, and if so, whether the deviation mode or the recovery mode is appropriate.

If the decision is in favour of the deviation mode, then the correction is given by $y_{corr} = y_{meas} + 0{,}025$ (mmol/mol). If the decision is in favour of the recovery mode, then the correction is given by $y_{corr} = y_{meas}/0{,}98$ (mmol/mol).

The uncertainty budget for measurement results including a bias correction in the deviation mode is given by $u(y_{corr}) = [s^2(y_{meas}) + u^2(x_{ref}) + s_{obs}{}^2/n]^{1/2} = [s^2(y_{meas}) + 0{,}36 \times 10^{-4} + 0{,}45 \times 10^{-4}]^{1/2} = [s^2(y_{meas}) + 0{,}81 \times 10^{-4}]^{1/2}$. In this expression, $s(y_{meas})$ is the standard uncertainty of the uncorrected measurement result, expressed in mmol/mol. Utilizing the value of s_{IR} = 2 % relative, $u(y_{corr}) = [(0{,}02 \times y_{meas})^2 + 0{,}81 \times 10^{-4}]^{1/2}$.

If bias correction is carried out in the recovery mode, then the uncertainty budget for measurement results after correction is given in terms of relative standard deviations and uncertainties as:

$u_r(y_{corr}) = [s_r{}^2(y_{meas}) + u_r{}^2(x_{ref}) + s_{r\,obs}{}^2/n]^{1/2} = [s_r{}^2(y_{meas}) + 0{,}21 \times 10^{-4} + 0{,}28 \times 10^{-4}]^{1/2} = [s_r{}^2(y_{meas}) + 0{,}49 \times 10^{-4}]^{1/2}$.

In this expression, $s_r(y_{meas})$ is the relative standard uncertainty of the uncorrected measurement result. Utilizing the value of s_{IR} = 2 % relative, $u_r(y_{corr}) = [4{,}0 \times 10^{-4} + 0{,}49 \times 10^{-4}]^{1/2}$ = 2,1 %.

NOTE 3 If the same correction is applied to independent measurement results, the final results including bias correction are correlated. For correction using an average deviation $\langle\delta\rangle$ the corresponding covariance is given by $u^2(\langle\delta\rangle)$. Analogously, for correction using an average recovery $\langle Q\rangle$ the corresponding covariance term is given by the relative standard uncertainty $u_r{}^2(\langle Q\rangle)$.

5.2.2.4.2 Step 3b — Bias handling without correction

As an alternative, applicable to acceptable bias (significant or not), no correction is applied to the raw measurement result, but an allowance for bias is made in the uncertainty estimation.

In case of significant bias found on the reference sample, this action is preferred over bias correction if the available data are not sufficient for deriving a sound correction and further investigations to supply additional data are impossible or impractical. For example, it may be doubtful whether a single-level correction, based on measurements of a single reference sample, is applicable to the entire analytical range, or if so, whether correction utilising an average deviation or correction utilising an average recovery is more appropriate. Then, in principle, additional reference samples should be investigated in order to characterise the bias to an appropriate degree. If this is not possible or not practical, no correction is performed. Instead, an allowance is made in the uncertainty estimation to account for the observed bias.

An allowance for bias in the uncertainty estimation is also made in case of insignificant bias found on the reference sample, to account for the empirical fact that within-laboratory estimates of precision almost always fail to cover the variability of analytical results obtained in collaborative studies, and therefore laboratory bias has to be included.

For this purpose, the procedure proposed by Lira and Woeger [9] introduced in 5.2.1 is recommended. Putting this approach into equations gives

$$y_{est} = y_{meas} \tag{22}$$

$$u^2(y_{est}) = s_{IR}^2 + \frac{s_{obs}^2}{n} + u^2(x_{ref}) + \langle\delta\rangle^2 \tag{23}$$

where the symbol y_{est} is used for an estimate of the measurand (the concentration of the target analyte), taking into account results of bias investigation by other means than bias correction.

According to this procedure, instead of applying a correction for the observed bias, $\langle\delta\rangle$, the bias is added in quadrature to the standard uncertainty that would be obtained when applying the correction. In the uncertainty budget, according to Equation (23), s_{obs} may be replaced by s_{IR} to obtain a more robust estimate.

Equation (23) is valid only for analyte concentrations near to those of the reference sample, where it can be taken for granted that the absolute error of analysis on the test sample is about the same as on the reference sample. Otherwise an analogue, designed for a concentration range with constant relative errors of analysis, could be used where appropriate.

However, most often a constant absolute uncertainty allowance $\langle\delta\rangle^2$ or a constant relative absolute uncertainty allowance $\langle Q\rangle^2$, obtained on a single reference sample, over extended concentration ranges, will be doubtful. Instead, for wider concentration ranges, estimation of bias-related uncertainty should rather be based on the empirical fact that concentration dependence of analytical uncertainty is between u = constant and $u \sim$ concentration. Therefore a safe approximation is to use the uncertainty estimate obtained on the reference sample as a constant absolute value for $y_{meas} < \langle x_{obs}\rangle$ and as a constant relative value for $y_{meas} \geqslant \langle x_{obs}\rangle$. This gives

$$u^2\left(y_{est}\right) = s_{IR}^2 + \frac{s_{obs}^2}{n} + u^2\left(x_{ref}\right) + \left\langle\delta\right\rangle^2 \text{ for } y_{meas} < \langle x_{obs}\rangle \tag{24}$$

and

$$u^2\left(y_{est}\right) = \left(\frac{y_{meas}}{\langle x_{obs}\rangle}\right)^2\left[s_{IR}^2 + \frac{s_{obs}^2}{n} + u^2\left(x_{ref}\right) + \left\langle\delta\right\rangle^2\right] \text{ for } y_{meas} \geqslant \langle x_{obs}\rangle \tag{25}$$

In Equations (24) and (25), s_{obs} may be replaced by s_{IR} to obtain a more robust estimate.

EXAMPLE Continuing the example in Step 2 and Step 3 a, consider the case where the available data, obtained with a single reference gas mixture, are not found sufficient for deriving a correction. Then additional reference samples should be investigated, but if this is not possible or not practical, an additional allowance is made in the uncertainty estimation to account for the observed (significant) bias instead of a correction.

Using the deviation mode, where y_{est} is obtained from y_{meas} by subtracting zero, the standard uncertainty of the estimate y_{est} is given by $u(y_{est}) = [s^2(y_{meas}) + u^2(\langle\delta\rangle) + \langle\delta\rangle^2]^{1/2} = [s^2(y_{meas}) + 0{,}81 \times 10^{-4} + 6{,}25 \times 10^{-4}]^{1/2} = [s^2(x_{meas}) + 7{,}1 \times 10^{-4}]^{1/2}$. In this expression, $s(x_{meas})$ is the standard uncertainty of the uncorrected measurement result, expressed in mmol/mol. Utilizing the value of s_{IR} = 2 % relative, $u(y_{est}) = [(0{,}02 \times y_{meas})^2 + 7{,}1 \times 10^{-4}]^{1/2}$.

For a safe estimate of uncertainty over the entire measuring range, the uncertainty estimate obtained using the deviation mode may be extrapolated as follows: $u(y_{est}) = [(0{,}02 \times y_{meas})^2 + 7{,}1 \times 10^{-4}]^{1/2}$ for $y_{meas} < 1{,}27$ mmol/mol and $u(y_{est}) = (y_{meas}/1{,}27) \times [(0{,}02 \times y_{meas})^2 + 7{,}1 \times 10^{-4}]^{1/2}$ for $y_{meas} \geqslant 1{,}27$ mmol/mol.

NOTE As a notable difference concerning the handling of insignificant bias, in this procedure an allowance for insignificant bias is made, while in 5.2.1 no such action is recommended. This is due to the difference in starting positions and objectives. In the case considered in 5.2.1, a measurement procedure is investigated for potential residual bias as a final step after an exhaustive evaluation of measurement uncertainty. If this investigation does not reveal any significant bias, this confirms the expectation that method development and uncertainty evaluation have properly addressed all relevant systematic effects impacting the measurement result, and no further action is required. In the case considered in 5.2.2, the objective is to construct an estimate of measurement uncertainty using data from precision and bias investigations carried out within a laboratory. Based on experiences from collaborative studies, which have demonstrated the importance to account for laboratory bias, here the expectation is that bias estimates provide a major contribution irrespective of whether the observed deviations are significant or not. Also it should be noted that here significance of deviations is defined differently from 5.2.1, accounting only for the random variability in the bias investigation. For these reasons, an allowance for bias in the uncertainty estimation appears to be reasonable also in cases of insignificant deviations.

5.2.3 Matrix-varied reference samples

5.2.3.1 General

For investigation and handling of bias within a narrow range of analyte concentration, often a 1-level study will be sufficient in principle. However, if matrix interferences are relevant, two or more reference gas mixtures varying in matrix composition shall be used. This clause specifies procedures for deriving an average bias correction, for estimating the uncertainty associated with this correction, and for making an allowance for uncorrected bias in the uncertainty budget, using two reference samples of similar analyte concentration but different matrix composition. The procedures are designed for a case-B investigation (see 5.1). However, analogous considerations apply to case-A investigations.

The symbols and units used in the description of the 2-level procedure are the same as in 5.2.2, with additional indices 1 and 2 denoting quantities referring to two reference samples R_1 and R_2, respectively.

For a clear distinction from the symbols for values referring to the reference samples, the symbol y is used for values referring to the subsequent analysis of test samples.

5.2.3.2 Procedure a) — Correction for bias

In the case of constant absolute errors significant bias is corrected by subtracting the mean value of the average deviations $\langle\delta_1\rangle$ and $\langle\delta_2\rangle$ determined in the bias study from the raw result y_{meas} of a measurement on a test sample according to

$$y_{\text{corr}} = y_{\text{meas}} - \frac{\langle\delta_1\rangle + \langle\delta_2\rangle}{2} \tag{26}$$

In the case of constant relative errors significant bias is corrected by way of dividing the raw result y_{meas} of a measurement on a test sample through the mean value of the average recoveries $\langle Q_1\rangle$ and $\langle Q_2\rangle$ determined in the bias study according to

$$y_{\text{corr}} = \frac{2y_{\text{meas}}}{\langle Q_1\rangle + \langle Q_2\rangle} \tag{27}$$

For correction using deviations, the standard uncertainty of corrected measurement results is estimated by

$$u^2(y_{\text{corr}}) = s^2(y_{\text{meas}}) + \left(\frac{\langle\delta_1\rangle - \langle\delta_2\rangle}{2}\right)^2 + \frac{u^2\langle\delta_1\rangle + u^2\langle\delta_2\rangle}{2} \tag{28}$$

The uncertainty budget according to Equation (28) comprises three components, accounting for the precision of the raw measurement result y_{meas}, the range of observed deviations $\langle\delta_1\rangle$, $\langle\delta_2\rangle$, and the uncertainty on these deviations.

The quantities $\langle\delta_1\rangle$, $\langle\delta_2\rangle$, $u^2(\langle\delta_1\rangle)$, $u^2(\langle\delta_2\rangle)$ and $s^2(y_{\text{meas}})$ are determined according to the procedures in 5.2.2.

For correction using recoveries, an analogous equation applies, with (absolute) standard uncertainties and deviations replaced by relative ones, and average deviations replaced by average recoveries.

NOTE As an alternative to the uncertainty estimate according to Equation (28), the uncertainty associated with an average correction could be estimated using one-way Analysis of Variance (ANOVA) on the combined data of deviations or recoveries obtained on the reference samples.

EXAMPLE Continuing the examples in 5.2.2, a second reference gas mixture with similar analyte content in a different matrix is measured to obtain a bias correction which includes matrix effects on an average basis. While the first reference gas mixture (R_1) treated in previous examples is CO in pure nitrogen, the second mixture (R_2) contains in addition CO_2 and propane. The reference data are x_1^{ref} = 1,295 mmol/mol, $u(x_1^{\text{ref}})$ = 0,006 mmol/mol and x_2^{ref} = 4,76 mmol/mol, $u(x_2^{\text{ref}})$ = 0,017 mmol/mol. In addition to ten repeated measurements for mixture R_1 (see 5.2.1) with

mean value $\langle x_1^{obs} \rangle$ = 1,27 mmol/mol and standard deviation s_1^{obs} = 0,021 mmol/mol, another series of ten measurements under similar conditions were made on mixture R_2, giving a mean value $\langle x_2^{obs} \rangle$ = 4,65 mmol/mol and standard deviation s_2^{obs} = 0,10 mmol/mol. Considering previous experience with the analytical method in use, the measuring range 0,5 mmol/mol – 5 mmol/mol under consideration is sufficiently narrow as to justify the use of single-level bias correction.

For both these mixtures, the observed bias is significant, and the decision is taken to utilize an average bias correction. Since the differences $\langle x_1^{obs} \rangle - x_1^{ref}$ = – 0,025 mmol/mol = – 1,9 % (relative) and $\langle x_2^{obs} \rangle - x_2^{ref}$ = – 0,11 mmol/mol = –2,4 % (relative) are rather far apart on an absolute scale but rather similar on a relative scale, the correction will be based on an average recovery according to $y_{corr} = y_{meas}/\langle Q \rangle$. The average recovery is calculated according to $\langle Q \rangle$ = ½ ($\langle x_1^{obs} \rangle / x_1^{ref} + \langle x_2^{obs} \rangle / x_2^{ref}$) = ½ (1,27/1,295 + 4,65/4,76) = 0,979.

The uncertainty budget for measurement results after correction is given in terms of relative standard uncertainties as $u_r(y_{corr}) = [s_r^2(y_{meas}) + u_r^2(\langle Q \rangle)]^{1/2}$. Copying the recipe for the estimation of uncertainty of an average deviation, the uncertainty of $\langle Q \rangle$ is estimated as follows:

$$
\begin{aligned}
u_r^2(\langle Q \rangle) &= ¼ (\langle Q_1 \rangle - \langle Q_2 \rangle)^2 + ½[u_r^2(\langle Q_1 \rangle) + u_r^2(\langle Q_2 \rangle)] \\
&= ¼ (\langle x_1^{obs} \rangle / x_1^{ref} - \langle x_2^{obs} \rangle / x_2^{ref})^2 + ½ [u_r^2(\langle x_1^{obs} \rangle) + u_r^2(x_1^{ref}) + u_r^2(\langle x_2^{obs} \rangle) + u_r^2(x_2^{ref})] \\
&= ¼ (0{,}38 \times 10^{-2})^2 + ½ (0{,}27 \times 10^{-4} + 0{,}21 \times 10^{-4} + 0{,}46 \times 10^{-4} + 0{,}13 \times 10^{-4}) \\
&= 0{,}57 \times 10^{-4}.
\end{aligned}
$$

This gives a (relative) standard uncertainty for corrected measurements of $u_r(y_{corr}) = [s_r^2(y_{meas}) + 0{,}57 \times 10^{-4}]^{1/2}$.

Due to the fact that the individual recovery factors are almost the same, these results are very close to the results obtained previously with the single reference mixture R_1 (see the examples in 5.2.2). Apparently in this case matrix interferences are no major source of bias and imprecision.

5.2.3.3 Procedure b) — Bias handling without correction

As an alternative, applicable to acceptable bias (significant or not), no correction is applied to the raw measurement result, but an allowance for bias is made in the uncertainty estimation according to

$$y_{est} = y_{meas} \tag{29}$$

$$u^2(y_{est}) = u^2(y_{corr}) + (y_{corr} - y_{meas})^2 = s^2(y_{meas}) + \frac{\langle \delta_1 \rangle^2 + \langle \delta_2 \rangle^2}{2} + \frac{u^2\langle \delta_1 \rangle + u^2\langle \delta_2 \rangle}{2} \tag{30}$$

where the symbol y_{est} is used for an estimate of the measurand (the concentration of the target analyte), taking into account results of bias investigation by other means than bias correction.

The quantities $\langle \delta_1 \rangle$, $\langle \delta_2 \rangle$, $u^2(\langle \delta_1 \rangle)$, $u^2(\langle \delta_2 \rangle)$ and $s^2(y_{meas})$ are determined according to the procedures in 5.2.2.

Equation (30) is valid only for analyte concentrations near to those of the reference samples, where it can be taken for granted that the absolute error of analysis on the test sample is about the same as on the reference samples. Otherwise an analogue, designed for a concentration range with constant relative errors of analysis, could be used where appropriate.

However, often a constant absolute uncertainty allowance or a constant relative absolute uncertainty allowance, based on two reference samples with almost the same analyte concentration, over extended concentration ranges, will be doubtful. Instead, for wider concentration ranges, estimation of bias-related uncertainty should rather be based on the empirical fact that concentration dependence of analytical uncertainty is between u = constant and $u \sim$ concentration. Therefore a safe approximation is to use the mean of the uncertainty estimates obtained on the two reference samples as a constant absolute value for $y_{meas} < \langle x_{obs} \rangle$ and as a constant relative value for $y_{meas} \geqslant \langle x_{obs} \rangle$, where $\langle x_{obs} \rangle$ denotes the mean value of the average results obtained on the reference samples, i.e. $\langle x_{obs} \rangle = (\langle x_1^{obs} \rangle + \langle x_2^{obs} \rangle)/2$. This gives

$$u^2(y_{est}) = s^2(y_{meas}) + \frac{\langle \delta_1 \rangle^2 + \langle \delta_2 \rangle^2}{2} + \frac{u^2\langle \delta_1 \rangle + u^2\langle \delta_2 \rangle}{2} \text{ for } y_{meas} < \langle x_{obs} \rangle \tag{31}$$

and

$$u^2(y_{\text{est}}) = \left(\frac{y_{\text{meas}}}{x_{\text{obs}}}\right) s^2(y_{\text{meas}}) + \frac{\langle\delta_1\rangle^2 + \langle\delta_2\rangle^2}{2} + \frac{u^2\langle\delta_1\rangle + u^2\langle\delta_2\rangle}{2} \quad \text{for } y_{\text{meas}} \geqslant \langle x_{\text{obs}}\rangle \tag{32}$$

The procedures described above may be extended to include more than two reference samples, and could be based on recoveries instead of deviations.

5.3 Bias handling for an extended measuring range

5.3.1 Design of investigations

For investigating potential bias of a candidate analytical procedure in an extended measuring range, the procedure is applied to appropriate reference samples R_1, R_2, ..., R_r and the results are compared with the reference values attributed to the reference samples. Alternatively, the candidate procedure and a reference procedure are applied in parallel to appropriate samples and the results of the candidate procedure are compared with those of the reference procedure.

The conclusions drawn from these comparisons are only valid if the samples R_1, R_2, ..., R_r are representative for the specified measuring range of the candidate procedure. If this is not the case, there are two possibilities as follows: either the measuring range shall be restricted accordingly, or the number of levels (samples) shall be increased.

The procedures specified in this clause are designed for a case-B investigation (see 5.1). However, analogous considerations apply to case-A investigations.

For each reference sample R_i the following data are required.

— Reference data x_i^{ref}, $u(x_i^{\text{ref}})$: In the case of reference samples taken from reference gas mixtures the reference value and its standard uncertainty are taken from the certificate of mixture composition. In the case of other samples analysed in parallel with a reference procedure the reference value is determined according to the specification of the reference procedure, and the standard uncertainty of the reference value is determined from the uncertainty budget of the reference procedure.

— Measured data $\langle x_i^{\text{obs}}\rangle$, $s(\langle x_i^{\text{obs}}\rangle)$: The result of the candidate procedure and its standard deviation are determined as the mean value of an appropriate number of repeated measurements, and the standard deviation of this mean, as established so far. Most often this standard deviation will be an intermediate-precision standard deviation, referring to specified within-laboratory reproducibility conditions (see ISO 5725-3 [8]). It can either be determined directly from the standard deviation of the measurement series, or taken from available precision monitoring data, e.g. a control chart. For a proper statistical basis, a minimum of 6 repeated measurements per reference sample are required.

For each reference sample R_i, the standard deviation s_i^{obs} of the actual measuring series used to determine the mean value should be checked for compliance with previously established precision data.

5.3.2 Assessing observed bias

For each reference sample R_i, the average deviation $\langle\delta_i\rangle = \langle x_i^{\text{obs}} - x_i^{\text{ref}}\rangle = \langle x_i^{\text{obs}}\rangle - x_i^{\text{ref}}$ is examined as to whether this deviation is technically serious or acceptable, and in case of an acceptable deviation, whether this deviation is statistically significant or not.

a) **Serious deviations** indicate serious deficiencies of the analytical procedure, necessitating thorough review and amendment of the protocol and equipment. The criterion for that is a matter of expert judgement, based on fitness-for-purpose considerations (e. g. concerning the target uncertainty) and previous experience with the method. In addition, the ratios $\langle\delta_i\rangle/s_i^{\text{obs}}$ should be in a reasonable range.

b) **Acceptable deviations** fall within the range expected by expert judgement. As a consequence, an amendment of the measurement procedure is not necessary, but a correction for bias or an allowance for bias in the uncertainty estimation has to be made. In case of doubt as to whether the available data and

other relevant information enable a technically sound correction, an allowance for bias in the uncertainty estimation is recommended.

As a subclass of acceptable deviations, **insignificant deviations** are not significantly different from zero, the criterion being that $|\langle\delta_i\rangle| \leqslant 2u(\langle\delta_i\rangle)$. The standard uncertainty required for this test is calculated according to $u^2(\langle\delta_i\rangle) = s^2(\langle x_i^{obs}\rangle) + u^2(x_i^{ref})$. In this expression the standard deviation of the mean value $\langle x_i^{obs}\rangle$ is given by $s(\langle x_i^{obs}\rangle) = s_i^{obs}/\sqrt{n_i}$, where s_i^{obs} is the standard deviation of the replicates used to determine the mean value and n_i is the number of replicates.

NOTE 1 The test criterion makes use of a coverage factor of $k = 2$, aiming at a significance level of approximately 95 %. For this purpose, use of Student´s factor t is common. However, replacement of the factor 2 by Student´s t, using the calculus of effective degrees of freedom specified in the GUM [3], would require a proper estimate of the degrees of freedom for an average deviation $\langle\delta_i\rangle$ which is difficult to obtain.

NOTE 2 Instead of testing deviations $\langle\delta_i\rangle = \langle x_i^{obs}\rangle - x_i^{ref}$ for significant departure from zero, recoveries $\langle Q_i\rangle = \langle x_i^{obs}\rangle / x_i^{ref}$ could be tested for significant departure from unity. These two tests are essentially equivalent.

If significant bias is found with any of the reference samples R_1, R_2, ..., R_r, a bias correction is determined. In addition, the correction uncertainty is estimated and added to the precision of the candidate procedure to obtain an estimate of the measurement uncertainty. Alternatively, instead of applying a correction, the bias may the taken into account by an additional allowance in the uncertainty estimation.

An allowance for bias in the uncertainty estimation is also made in case of insignificant bias found on the reference samples, to account for the empirical fact that within-laboratory estimates of precision almost always fail to cover the variability of analytical results obtained in collaborative studies, and therefore laboratory bias has to be included.

5.3.3 Correction for bias

5.3.3.1 General

If significant bias is found with any of the samples R_1, R_2, ..., R_r investigated according to 5.3.2, a bias correction is derived, using the data collected so far, i.e. the reference values x_i^{ref} and their standard uncertainties $u(x_i^{ref})$ as well as the results $\langle x_i^{obs}\rangle$ of the candidate procedure and their standard deviations $s(\langle x_i^{obs}\rangle)$ for all the reference samples R_i. Whether these data are sufficient for deriving a meaningful bias correction depends on several factors:

— the prior information on the concentration dependence of potential bias;

— the observed pattern of deviations $\langle x_i^{obs}\rangle - x_i^{ref}$;

— the number p of parameters of the correction model selected.

If the number r of reference samples should turn out insufficient, additional samples should be investigated. If this is not possible or not practical, no bias correction is performed. Instead, the bias is taken into account by an additional allowance in the uncertainty estimation according to the procedure specified in 5.3.4.

For a clear distinction from the symbols for values referring to the reference samples, the symbol y is used for values referring to the subsequent analysis of test samples.

Bias correction is performed in Steps A to D as follows.

5.3.3.2 Step A — Selection of the correction model

The correction model is chosen according to the pattern of observed deviations and the prior information on the concentration dependence of potential bias. For example, if from prior information at most a constant offset is expected, and if the observed deviations are approximately constant, then correction by a constant additive term according to the equation $y_{corr} = y_{meas} + b_1$ is appropriate. Similarly, if from prior information bias is expected to grow proportionally with analyte concentration, and if the observed deviations agree with this

expectation, then correction by a constant factor according to the equation $y_{corr} = b_2\, y_{meas}$ is appropriate. If no prior information is available, then the correction model can only be chosen on the basis of the observed deviations.

If a single-parameter correction is not sufficient to account for the observed deviations, corrections involving two or more parameters have to be considered, using e.g. a linear correction function $y_{corr} = b_2\, y_{meas} + b_1$ or a second-order polynomial $y_{corr} = b_3\, y_{meas}^2 + b_2\, y_{meas} + b_1$.

NOTE By tradition, non-linear curves are most often modelled using polynomials, being linear in the parameters to be determined. However, with the capacity of today's personal computers and standard computational software, there is no need to refrain from the use of non-linear functions. A compilation of function types for regression modeling, addressing various practical items as well as the statistical behaviour of these models, is given in Reference [10].

5.3.3.3 Step B — Selection of the number of reference samples and the data evaluation procedure

For determining the parameters of the selected correction model, two basically different techniques are available:

— interpolation, using a minimum correction data design;

— least-squares regression, using a redundant calibration design.

In the case of a **minimum correction data design,** the number r of reference sample data sets x_i^{ref}, $u(x_i^{ref})$, $\langle x_i^{obs} \rangle$, $s(\langle x_i^{obs} \rangle)$ equals the number p of correction parameters. Accordingly, the correction parameters are determined by solving the equations resulting if the terms obtained by applying the correction to the $\langle x_i^{obs} \rangle$ are equated to the corresponding reference values x_i^{ref}.

In the case of a **redundant correction data design,** the number r of sample data sets x_i^{ref}, $u(x_i^{ref})$, $\langle x_i^{obs} \rangle$, $s(\langle x_i^{obs} \rangle)$ exceeds the number p of correction parameters. Then the correction parameters are determined from the reference sample data using an appropriate least squares procedure.

The general recommendation of this International Standard is to use a redundant calibration design. If the bias study is restricted to effects of varying analyte concentration, the number r of reference sample data should be

— $r \geqslant 2$ for one-parameter corrections;

— $r \geqslant 3$ for two-parameter corrections;

— $r \geqslant 5$ for three-parameter corrections.

If the bias study includes matrix interferences, a redundant design is mandatory, and a larger number of reference samples will be required than the minimum numbers given above. Clause 6 provides guidance for the design of such studies.

5.3.3.4 Step C — Determination of the correction parameters

In the case of a minimum calibration design, the correction parameters are determined by solving the equations resulting if the terms obtained by applying the correction to the $\langle x_i^{obs} \rangle$ are equated to the corresponding reference values x_i^{ref} (see 5.2.2 for the case of a single parameter).

Using this procedure, the parameters for a linear correction according to $y_{corr} = b_2\, y_{meas} + b_1$ are determined from the data of two samples R_1 and R_2 as follows.

$$b_2 = \frac{x_1^{ref} - x_2^{ref}}{\left\langle x_1^{obs} \right\rangle - \left\langle x_2^{obs} \right\rangle} \tag{33}$$

$$b_1 = \frac{1}{2}\left[\left(x_1^{\text{ref}} + x_2^{\text{ref}}\right) - b_2\left(\left\langle x_1^{\text{obs}}\right\rangle + \left\langle x_2^{\text{obs}}\right\rangle\right)\right] \tag{34}$$

In more complex cases, e.g. for $p > 2$, or if non-linear functions are involved, the equations are solved numerically by using a computer program (an "equation solver"), as contained in common mathematical software.

In the case of a redundant-correction data design, the correction parameters are determined by least-squares adjustment to the reference sample data set. For this purpose, the computer program for ISO 6143 [7] could be used, but often ordinary (unweighted) least-squares regression will be sufficient.

For an additive correction according to $y_{\text{corr}} = y_{\text{meas}} + b_1$, the ordinary least-squares estimate of the correction term b_1 is given by the mean value of the individual differences $x_i^{\text{ref}} - \langle x_i^{\text{obs}}\rangle$. Analogously, for correction by means of a factor according to $y_{\text{corr}} = b_2\, y_{\text{meas}}$, the correction factor b_2 may be estimated by the mean value of the individual quotients $x_i^{\text{ref}}/\langle x_i^{\text{obs}}\rangle$.

For the parameters of a linear correction function and low-order polynomials, the ordinary least-squares solution is readily obtained using common mathematical software.

5.3.3.5 Step D — Estimation of correction uncertainty

The uncertainty of corrected measurements is obtained by combination of contributions from three different sources:

a) uncertainty of the uncorrected measurement,

b) uncertainty of the correction parameters,

c) uncertainty related to the correction model.

$$u^2\left(y_{\text{corr}}\right) = u^2\left(y_{\text{corr}}/\text{meas}\right) + u^2\left(y_{\text{corr}}/\text{para}\right) + u^2\left(y_{\text{corr}}/\text{model}\right) \tag{35}$$

In this uncertainty budget, the first term is essentially given by the intermediate-precision variance $s^2(y_{\text{meas}})$ of the uncorrected measurement result (multiplied with a sensitivity factor near to unity). The second term is obtained from the uncertainties $u(x_i^{\text{ref}})$, $s(\langle x_i^{\text{obs}}\rangle)$ of the reference sample data. For a redundant correction data design, the third term may be estimated from the lack of fit, e.g. the residual variance. For a minimum-correction data design, the reference sample data do not provide any information about the quality of the fit. Therefore expert judgement has to be used for an estimate of the third term.

5.3.4 Bias handling without correction

If significant bias is found on one or several reference samples, but a correction based on these findings is doubtful and investigation of additional reference samples is not possible or not practical, no correction is applied, and an additional allowance in the uncertainty estimation is made to account for the observed bias according to

$$y_{\text{est}} = y_{\text{meas}} \tag{36}$$

$$u^2\left(y_{\text{est}}\right) = u^2\left(y_{\text{corr}}\right) + \left(y_{\text{corr}} - y_{\text{meas}}\right)^2 \tag{37}$$

This procedure is also used when no significant bias is observed on any of the reference samples. For this purpose, it is not necessary to first calculate a correction function and estimate correction uncertainty. Instead, the uncertainty for measurement of a test sample may be estimated using the data for the nearest reference sample, R_*, according to the single-level procedure specified in 5.2.2. This gives

$$u^2\left(y_{\text{est}}\right) = s^2\left(y_{\text{meas}}\right) + \frac{s^2_{\text{obs}}}{n_*} + u^2\left(x_*^{\text{ref}}\right) + \left(\left\langle x_*^{\text{obs}}\right\rangle - x_*^{\text{ref}}\right)^2 \quad \text{for } y_{\text{meas}} < \langle x_*^{\text{obs}}\rangle \tag{38}$$

and

$$u^2\left(y_{\text{est}}\right) = \left(\frac{y_{\text{meas}}}{\left\langle x_*^{\text{obs}}\right\rangle}\right)^2 s^2\left(y_{\text{meas}}\right) + \frac{s^2_{\text{obs}}}{n_*} + u^2\left(x_*^{\text{ref}}\right) + \left(\left\langle x_*^{\text{obs}}\right\rangle - x_*^{\text{ref}}\right)^2 \quad \text{for } y_{\text{meas}} \geqslant \langle x_*^{\text{obs}}\rangle \tag{39}$$

The standard deviation of the raw measurement result, $s(y_{\text{meas}})$, may either be determined by replicate measurements of the respective test sample or, preferably, be taken from available precision-monitoring data. If y_{meas} is a mean value of several replicates, this shall be accounted for in the standard deviation $s(y_{\text{meas}})$.

6 Treatment of matrix interferences

6.1 General

Treatment of matrix interferences for extended ranges of analyte concentration and matrix composition depends on various factors, including

— the range of analytes and concentration levels,

— the range of matrix composition,

— the analytical system,

— the intended purpose of the investigation.

6.2 Case A

For the specified analytical range (i.e. the range of sample composition under consideration) and the analytical system to be used, there is no information about critical combinations of analytes X and on interferents I available. Then carry out an interference study using Analysis-of-Variance techniques (see GUM:1993, Annex H.5 for an introduction). In principle, such study requires a set of 2^N ($N = N_X + N_I$) calibration gases, where all analytes X (in total N_X) and all relevant matrix constituents I (in total N_I) are varied independently between maximum and minimum levels. For evaluating combined effects of groups of interferents, fractional designs with a smaller number of calibration gases may be used. As the main result of the study, for each analyte the critical interferents are identified. Use these results for a follow-up study according to case B.

6.3 Case B

For the specified analytical range (i.e. the range of sample composition under consideration) and the analytical system to be used, the critical combinations of an analyte X and an interferent I are known in advance. Then estimate the maximum effect of I on measurement of X , preferably by direct measurement using calibration gases with extreme levels of interferent I, and investigate whether the effect is significant. If significant interferences were found, the uncertainty for measuring analyte X may be increased as to account for the effect of all relevant interferents within specified concentration ranges, see Reference [11]. Alternatively, a multivariate calibration may be carried out (see case C), giving better uncertainty at the expense of higher efforts.

6.4 Case C

Instead of increasing the uncertainty for measuring a specified analyte, the effect of relevant interferents may be accounted for by multivariate calibration instead of separate univariate calibrations for each analyte. To this end, multiple responses, e.g. intensities at several wavelengths, are measured on reference gases with varying concentrations of analytes and relevant interferents. The multidimensional calibration data obtained in this manner are modelled by appropriate response surfaces (taking the place of response curves in univariate calibration), using techniques of multivariate data analysis. See Reference [12] for an introduction. Using these response surfaces, or rather corresponding multi-parameter functions, multiple responses obtained on an unknown mixture are converted into estimates of composition data.

Annex A
(normative)

Critical values for the trend test

Table A.1 — Critical values for the statistical trend test according to 4.2.3

N	99 %	95 %	N	99 %	95 %
4	0,625 6	0,780 5	33	1,228 3	1,443 4
5	0,537 9	0,820 4	34	1,238 6	1,451 1
6	0,561 5	0,890 2	35	1,248 5	1,458 5
7	0,614 0	0,935 9	36	1,258 1	1,465 6
8	0,662 8	0,982 5	37	1,267 3	1,472 6
9	0,708 8	1,024 4	38	1,276 3	1,479 3
10	0,751 8	1,062 3	39	1,285 0	1,485 8
11	0,791 5	1,096 5	40	1,293 4	1,492 1
12	0,828 0	1,127 6	41	1,301 7	1,498 2
13	0,861 8	1,155 8	42	1,309 6	1,504 1
14	0,893 1	1,181 6	43	1,317 2	1,509 8
15	0,922 1	1,205 3	44	1,324 6	1,515 4
16	0,949 1	1,227 2	45	1,331 7	1,520 6
17	0,974 3	1,247 3	46	1,338 7	1,525 7
18	0,997 9	1,266 0	47	1,345 3	1,530 5
19	1,019 9	1,283 4	48	1,351 5	1,535 1
20	1,040 6	1,299 6	49	1,357 3	1,539 5
21	1,060 1	1,314 8	50	1,362 9	1,543 7
22	1,078 5	1,329 0	51	1,368 3	1,547 7
23	1,095 8	1,342 5	52	1,373 8	1,551 8
24	1,112 2	1,355 2	53	1,379 2	1,555 7
25	1,127 8	1,367 1	54	1,384 6	1,559 6
26	1,142 6	1,378 5	55	1,389 9	1,563 4
27	1,156 7	1,389 2	56	1,394 9	1,567 0
28	1,170 2	1,399 4	57	1,399 9	1,570 7
29	1,183 0	1,409 1	58	1,404 8	1,574 3
30	1,195 1	1,418 3	59	1,409 6	1,577 9
31	1,206 7	1,427 0	60	1,414 4	1,581 4
32	1,217 7	1,435 4	∞	2,000 0	2,000 0

Critical values taken from Sachs [13].

Annex B
(informative)

Uncertainty issues

B.1 Accounting for correlation between repeated measurements

Let y and y' be measurement results obtained using the same measurement procedure. Then y and y' may be viewed as depending on the same input variables, taking partly the same and partly different values. For example, if the ambient temperature in the laboratory (monitored on a regular basis) is an input variable, then y and y' could either depend on the same temperature measurement or on different temperature measurements. Assuming that all input quantities were determined independently, the covariance between y and y' is given by

$$u(y,y') = \sum c_y(z)\,c_{y'}(z)\,u^2(z) \tag{B.1}$$

where the sum is over all input quantities z shared by y and y', and the c's are the sensitivity coefficients concerned. Let $u_{\text{com}}(y)$ and $u_{\text{com}}(y')$ be that part of the combined standard uncertainty of y and y', respectively, which is due to the shared (common) input values, i.e.

$$u_{\text{com}}^2(y) = \sum \left[c_y(z)\,u(z)\right]^2 \tag{B.2}$$

Then $u(y, y') \leqslant u_{\text{com}}(y)\,u_{\text{com}}(y')$, and the product on the right-hand side of this inequality is in fact a reasonable approximation if $c_y(z) \approx c_{y'}(z)$ for all z, or if the ratio $c_y(z)/c_{y'}(z)$ is approximately constant. This may be assumed if y and y' are obtained using the same measurement procedure on similar samples. An estimate of the covariance between y and y' is required for proper evaluation of the uncertainty for combinations of measurement results such as sums and differences, where

$$u^2(y \pm y') = u^2(y) + u^2(y') \pm 2u(y,y') \tag{B.3}$$

and for products and quotients, where analogous expressions, then involving relative uncertainties and covariances, apply.

The approximation $u(y,y') \approx u_{\text{com}}(y)u_{\text{com}}(y')$ is used in 5.2.1 to evaluate the covariance term arising in the uncertainty budget according to Equation (12) using

$$u\left(y_{\text{meas}}, \langle x_{\text{obs}} \rangle\right) = u\left(y_{\text{meas}}, x_{\text{obs}}\right) \approx u_{\text{com}}\left(y_{\text{meas}}\right) u_{\text{com}}\left(x_{\text{obs}}\right) \tag{B.4}$$

The estimate $u(y,y') = u_{\text{com}}(y)u_{\text{com}}(y')$ is *a fortiori* applicable to replicate measurements on the same object. This may be utilised in the estimation of the standard uncertainty of a mean value as follows. Let $x_1, x_2, \ldots, x_n$ be the results of replicate measurements on the same object or sample, and let $\langle x \rangle$ denote the mean value of the x_i. The standard uncertainty of this mean is determined by

$$u^2\left(\langle x \rangle\right) = \frac{1}{n^2}\left[\sum_i u^2(x_i) + \sum_{j \neq k} u\left(x_j, x_k\right)\right] \tag{B.5}$$

Since we are dealing with replicates, we may assume that all the $u(x_i)$ are the same and replace them by a common estimate $u(x)$. Now consider a decomposition of the combined standard uncertainty $u(x)$ of a single measurement according to

$$u^2(x) = u_{\text{var}}^2(x) + u_{\text{inv}}^2(x) \tag{B.6}$$

where $u_{\text{var}}(x)$ is the combined standard uncertainty accounting for all influence quantities which are effectively varied between replications, while $u_{\text{inv}}(x)$ is the combined standard uncertainty accounting for all influence quantities which are effectively invariant under replication conditions and are thus shared by all replicates. Then the covariance between any two replicates is given by $u(x,x') = u_{\text{inv}}(x)u_{\text{inv}}(x') = u_{\text{inv}}^2(x)$. Using this estimate and Equation (B.6), the standard uncertainty of a mean value $\langle x \rangle$ is obtained as follows:

$$u^2\left(\langle x \rangle\right) = \frac{1}{n^2}\left[nu^2(x) + n(n-1)u_{\text{inv}}^2(x)\right] = \frac{u_{\text{var}}^2(x)}{n} + u_{\text{inv}}^2(x) \tag{B.7}$$

Based on Equations (B.4) and (B.7), the uncertainty budget according to Equation (12) in 5.2.1 is derived as follows. Utilising decompositions of $u(y_{\text{meas}})$ and $u(x_{\text{obs}})$ into two parts, accounting for individual and common influences, according to $u^2(\ldots) = u_{\text{ind}}^2(\ldots) + u_{\text{com}}^2(\ldots)$, the standard uncertainty of corrected measurement results is obtained as

$$u^2(y_{\text{corr}}) = u_{\text{ind}}^2(y_{\text{meas}}) + u_{\text{ind}}^2(x_{\text{obs}})/n + \left[u_{\text{com}}(y_{\text{meas}}) - u_{\text{com}}(x_{\text{obs}})\right]^2 + u^2(x_{\text{ref}}) \tag{B.8}$$

Assuming that $u_{\text{com}}(y_{\text{meas}}) \approx u_{\text{com}}(x_{\text{obs}})$, and using the intra-laboratory reproducibility standard deviation s_{IR} as a common estimate of $u_{\text{ind}}(y_{\text{meas}})$ and $u_{\text{ind}}(x_{\text{obs}})$, the final result is

$$u^2(y_{\text{corr}}) = \frac{n+1}{n}s_{\text{IR}}^2 + u^2(x_{\text{ref}}) \tag{B.9}$$

B.2 Accounting for uncorrected bias

In practice it may happen that significant bias is found, but the data is not sufficient for deriving a sound correction. For example, it may be doubtful whether a single-level correction, based on measurements of a single standard, is applicable to the entire measuring range, or if so, whether correction utilising an average deviation or correction utilising an average recovery is more appropriate. Then additional measurements, e.g. including another reference sample, should be made in order to characterise the bias to an appropriate degree. If this is not possible or not practical, a pragmatic alternative is to increase the uncertainty to account for the observed bias instead of a bias correction.

The GUM:1993 [3], in the Note to 6.3.1, appears to rather discourage such procedure, stating “Occasionally one may find that a known correction for a systematic effect has not been applied to the reported result of a measurement, but instead an attempt is made to take the effect into account by enlarging the “uncertainty” assigned to the result. This should be avoided; only in very special circumstances should corrections for known systematic effects not be applied to the results of a measurement. ... Evaluating the uncertainty of a measurement result should not be confused with assigning a safety limit to some quantity.”

In appreciating this guidance, a key phrase to recognise is that of a “known correction”. Certainly systematic effects (i.e. bias) that have been characterised to a degree that the applicable corrections can be considered as known, should be corrected, unless this entails unacceptable expense. In practice, however, it will often be the expense for deriving rather than for applying a “known correction” that is prohibitive. Then increasing measurement uncertainty to account for significant bias is most certainly better than applying a doubtful correction or, even worse, ignoring the bias.

Currently, a range of different approaches to account for uncorrected bias has been proposed, and a generally accepted procedure has not yet emerged. Concerning the basic philosophy of current approaches, two major lines of thought, responding to different objectives, may be differentiated:

a) aiming at expanded uncertainty to provide 95 % coverage;

b) aiming at an extension of standard uncertainty for biased estimators.

Throughout this clause, U denotes expanded uncertainty and $\delta = \langle x_{obs} - x_{ref} \rangle$ is a bias estimate, assumed here to be positive.

The majority of proposals belong to the first category, including (among others):

— replace an expanded uncertainty U by $U + \delta$, where δ is the bias correction that is not applied and U is obtained under the assumption $\delta = 0$; source: GUM:1993 [3], F.2.4.5,

— use an asymmetric confidence interval according to $y - U_{-} \leqslant Y \leqslant y + U_{+}$ with $U_{-} = U - \delta$ and $U_{+} = U + \delta$; source: Philips *et al.* [14],

— replace an expanded uncertainty U by $\sqrt{U^2 + \delta^2}$; source: Philips *et al.* [14],

— replace the standard uncertainty u by $\sqrt{u^2 + (\delta/k)^2}$ where k is the coverage factor applied in the calculation of expanded uncertainty; source: Ellison and Williams [15].

The rationale for the second category of proposals is the need to extend the methodology of the GUM, which is restricted to unbiased measurements, to accommodate measurement bias. To this end, the standard deviation associated with an unbiased estimate has to be generalised to a performance measure applicable to biased estimators, and an obvious candidate for that is the mean squared error. Proposals from this category include:

— replace the standard uncertainty u by $\sqrt{u^2 + \delta^2}$; source: Philips *et al.* [14].

— replace the standard uncertainty u by $\sqrt{u^2 + \delta^2 + u^2(\delta)}$ where $u(\delta)$ is the standard uncertainty associated with the bias estimate; source: Lira and Woeger [9].

In this International Standard the approach proposed by Lira and Woeger [9] is adopted.

Bibliography

[1] ISO 14111, *Natural gas — Guidelines to traceability in analysis*

[2] ISO/TS 14167, *Gas analysis — General quality assurance aspects in the use of calibration gas mixtures — Guidelines*

[3] *Guide to the expression of uncertainty in measurement (GUM)*. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, 1st edn., 1993, corrected and reprinted in 1995

[4] ISO 7504:2001, *Gas analysis — Vocabulary*

[5] *International vocabulary of basic and general terms in metrology (VIM)*. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, 2nd edn., 1993

[6] ISO 8258, *Shewhart control charts*

[7] ISO 6143, *Gas analysis — Comparison methods for determining and checking the composition of calibration gas mixtures*

[8] ISO 5725-3, *Accuracy (trueness and precision) of measurement methods and results — Part 3: Intermediate measures of the precision of a standard measurement method*

[9] LIRA I.H. and WOEGER W. Evaluation of the uncertainty associated with a measurement result not corrected for systematic effects. *Meas. Sci. Technol.*, **9**, 1998, pp. 1010-1011

[10] RATKOWSKY D.A. *Handbook of nonlinear regression models*. Marcel Dekker Inc., 1990

[11] ISO 14956, *Air quality — Evaluation of the suitability of a measurement procedure by comparison with a required measurement uncertainty*

[12] BRERETON R.D. Introduction to multivariate calibration in analytical chemistry. *Analyst*, **125**, 2000, pp. 2125-2154

[13] SACHS L. *Angewandte Statistik*. 7. Auflage, Springer-Verlag, 1992

[14] PHILLIPS S.D., EBERHARDT K.R., PARRY B. Guidelines for Expressing the Uncertainty of Measurement Results Containing Uncorrected Bias. *J. Res. Natl. Inst. Stand. Technol.,* **102**, 1997, pp. 577-585

[15] ELLISON S.L.R. ISO uncertainty and collaborative trial data. *Accred. Qual. Assur.,* **3**, 1998, pp. 95-100

ICS 71.040.40

Price based on 33 pages