INTERNATIONAL STANDARD

ISO 11095

First edition
1996-02-01

Linear calibration using reference materials

Étalonnage linéaire utilisant des matériaux de référence

Reference number
ISO 11095:1996(E)

Contents

Page

1 Scope 1

2 Normative references 1

3 Definitions 1

4 General principles 2

5 Basic method 2

6 The steps of the basic method 4

7 Control method 10

8 Two alternatives to the basic method 13

9 Example 16

Annexes

A List of symbols and abbreviations 25

B Basic method when the number of replicates is not constant 27

C Bibliography 29

International Organization for Standardization
Case Postale 56 • CH-1211 Genève 20 • Switzerland

Printed in Switzerland

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.

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.

International Standard ISO 11095 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 6, *Measurement methods and results*.

Annexes A and B form an integral part of this International Standard. Annex C is for information only.

Introduction

Calibration is an essential part of most measurement procedures. It is a set of operations which establish, under specified conditions, the relationship between values indicated by a measurement system and the corresponding accepted values of some "standards". In this International Standard, the standards are reference materials.

A reference material (RM) is a substance or an artifact for which one or more properties are established sufficiently well to validate a measurement system. There exist several kinds of RMs:

a) an internal reference material is an RM developed by a user for his/her own internal use;

b) an external reference material is an RM provided by someone other than the user;

c) a certified reference material is an RM issued and certified by an organization recognized as competent to do so.

Linear calibration using reference materials

1 Scope

This International Standard:

a) outlines the general principles needed to calibrate a measurement system and to maintain that "calibrated" measurement system in a state of statistical control;

b) provides a basic method

— for estimating a linear calibration function under either one of two assumptions relating to the variability of the measurements,

— for checking the assumption of linearity of the calibration function and the assumptions on the variability of the measurements, and

— for estimating the value of a new unknown quantity by transforming the measured values obtained on that quantity with the calibration function;

c) provides a control method for extended use of a calibration function

— for detecting when the calibration function needs to be updated, and

— for estimating the uncertainty of the measured values after transformation with the calibration function;

d) provides two alternatives to the basic method under special conditions;

e) illustrates the basic method and the control method with an example.

This International Standard is applicable to measurement systems for which reference materials are available.

It is applicable to measurement systems with an assumed linear calibration function. It offers a method for examining the assumption of linearity. If it is known that the calibration function is nonlinear, then this International Standard is not applicable unless one uses the "bracketing technique" described in 8.3.

This International Standard does not make a distinction among the various types of RMs and considers that the accepted values of the RMs selected to calibrate the measurement system are without error.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this International Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this International Standard are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 3534-1:1993, *Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms.*

ISO 3534-2:1993, *Statistics — Vocabulary and symbols — Part 2: Statistical quality control.*

ISO Guide 30:1992, *Terms and definitions used in connection with reference materials.*

3 Definitions

For the purposes of this International Standard, the definitions given in ISO 3534-1 and ISO 3534-2 and the following definition apply.

3.1 reference material: A substance or an artifact for which one or more properties are established suf-

ficiently well to be used to validate a measurement system.

4 General principles

Calibration is a procedure that determines the systematic difference that may exist between a measurement system and a "reference" system represented by the reference materials and their accepted values. In this International Standard, the term system (measurement system or reference system) is used to represent not only a measuring instrument but also the set of procedures, operators and environment conditions associated with that instrument.

The output of a calibration procedure is a calibration function that is used to make transformations of future measurement results. In this International Standard, the term "transformation" refers to

— either a correction of the future measurements if both the accepted values of the reference materials (RMs) and the observed values have the same units,

— or a translation from the units of the observed measurements to the units of the RMs.

The validity of the calibration function depends on two conditions:

a) that the measurements from which the calibration function was calculated are representative of the normal conditions under which the measurement system operates; and

b) that the measurement system is in a state of control.

The calibration experiment must be designed to ensure that point a) is met. The control method determines, as soon as possible, when the system has to be considered out of control.

The procedure in this International Standard is only applicable to measurement systems which are linearly related to their reference systems. To check whether the assumption of linearity is valid, more than two RMs must be used during the calibration experiment. This is illustrated in the basic method. Using several RMs, the basic method provides a strategy and techniques to analyse the data collected during the calibration experiment. If linearity is not in question, then an alternative method, simpler than the basic method, can be used to estimate a linear calibration function based on one point. This "one-point calibration" method (following a zero-level transformation) does not allow for any test of assumptions, but it is a quick and easy method to "recalibrate" a system that has been studied more thoroughly during previous experiments. If linearity is in question, then a second alternative can be used, called "bracketing".

The basic method and the one-point method are based on the assumption that the effort invested in calibration will be valid over a period of stability of the process. To study the period over which the calibration is valid, a control method has to be in place. The control method is designed to detect whether changes have taken place in the system that justify an investigation and/or a recalibration. The control method also provides a simple way to determine the precision of the values that have been transformed with a given calibration function.

The bracketing method is labour intensive but may provide greater accuracy in the determination of the values of unknown quantities. This method consists of surrounding as tightly as possible (bracketing) each unknown quantity by two RMs and extracting a transformed value for the unknown quantity from measurements of both the unknown quantity and the values of the two RMs. Only short-term stability of the measurement process is assumed (stability during the measurement of the unknown quantity and of the two RMs). Linearity is assumed solely in the interval between the values of the two RMs.

5 Basic method

5.1 General

This clause describes how to estimate and use a linear calibration function when several (more than two) RMs are available. The availability of several RMs allows the linearity of the calibration function to be verified.

5.2 Assumptions

5.2.1 It is assumed that there is no error in the accepted values of the RMs (this assumption will not be checked in this International Standard). In practice, accepted values of RMs are quoted with their uncertainties. The assumption of no error in the accepted values of the RMs can be considered valid if the uncertainties are small compared to the magnitude of the errors in the measured values of these RMs (see ref. [1]).

NOTE 1 In situations where the RMs have been treated chemically or, in some instances physically, before instrument readings are taken, this International Standard may underestimate the uncertainty associated with the transformation of a new measurement result.

5.2.2 The calibration function is assumed to be linear (this assumption will be examined).

5.2.3 Repeated measurements of a given RM are assumed to be independent and normally distributed, with variance referred to as "residual variance" (the independence and normality assumptions will not be checked in this International Standard). The square root of the residual variance is referred to as the residual standard deviation.

5.2.4 The residual standard deviation is assumed to be either constant or proportional to the accepted value of the RM (this assumption will be examined).

5.3 Calibration experiment

5.3.1 Experimental conditions

Experimental conditions should be the same as the normal operating conditions of the measurement system; i.e. if, for example, more than one operator uses the measuring equipment then there should be more than one operator represented in the calibration experiment.

5.3.2 Choice of RMs

The range of values spanned by the selected RMs should include (as far as is possible) the range of values encountered during normal operating conditions of the measurement system.

The composition of the selected RMs should be as close as possible to the composition of the targeted material to be measured.

The values of the RMs should be distributed approximately equidistantly over the range of values encountered during normal operating conditions of the measurement system.

5.3.3 Number of RMs, N

The number of RMs used to assess the calibration function should be at least 3.

For an initial assessment of the calibration function, a number larger than 3 is recommended (at least 3 over any subinterval where there is a doubt about the linearity of the calibration function).

5.3.4 Number of replicates, K

Each RM should be measured at least twice (as many replicates as is possible in practice is recommended). The number of replicates should be the same for all RMs.

The time and conditions at which the replicates are taken should cover as wide a range as is necessary to ensure that all operating conditions are represented.

5.4 Strategy for analysing the data

5.4.1 Plot the data to check

a) the state of control of the measurement system during the calibration experiment,

b) the assumption of linearity, and

c) the variability of the measurements as a function of the accepted values of the RMs.

5.4.2 Estimate the linear calibration function under the assumption of constant residual standard deviation.

5.4.3 Plot the calibration function and the residuals. The residuals plot is a strong indicator of departure from either the assumption of linearity or from the assumption of constant residual standard deviation. If the assumption of constant residual standard deviation does hold, skip step 5.4.4 and continue with step 5.4.5. Otherwise, execute step 5.4.4.

5.4.4 Estimate the linear calibration function under the assumption of proportional residual standard deviation and plot the calibration function and the residuals.

5.4.5 Evaluate the lack of fit of the calibration function. If the variability due to lack of fit is large relative to the variability due to replication of measurements, investigate the procedures followed during the calibration experiment and re-examine the assumption of linearity of the calibration function. If the assumption of linearity does not hold, then an alternative is to use the bracketing technique described in 8.3.

NOTE 2 There exist other techniques, beyond the scope of this International Standard, that allow the fitting of a quadratic or polynomial curve to the data (see refs. [2] and [3]).

5.4.6 Transform future measured values with the calibration function.

The next clause describes the six steps of this strategy. Clause 9 illustrates the basic method with an example.

6 The steps of the basic method

6.1 Plot of the data collected during the calibration experiment

Figure 1 shows a plot of the measured values versus the corresponding accepted values of the RMs. Figure 1 as well as figures 2 to 5 are obtained from simulated data. The purpose of these five plots is to illustrate the type of information one can extract from such plots. A complete example is treated in clause 8 with data, plots and analysis.

The major purpose of the plot shown as figure 1 is to detect visually any unusual behaviour of the measurement system during the calibration experiment, and to identify potential outliers. If possible, label the order of the data points and look for obvious time trends. If some of the data are considered suspicious, or if a time trend is obvious, then an investigation shall take place to discover causes of irregularities. As soon as the causes of irregularities are removed, the calibration experiment should be repeated and new data should be collected to establish a calibration function.

If the causes for one or a very few outliers are found, and if these causes do not affect the remaining measurements, then the outliers can be eliminated. The calibration experiment then becomes unbalanced; i.e. there is an unequal number of measurements K_n instead of K for each RM. Estimation of the calibration function can still proceed with the formulae given in 6.2, 6.4 and 6.5 replaced by the ones in annex B.

Figure 1 also allows an early diagnosis of the assumption of linearity of the calibration function, as well as a first look at the assumption of constant residual standard deviation. The linearity of the calibration function can be visually checked by visualizing a straight line through the data plotted in figure 1 (there seems to be some curvature in the data of figure 1). The assumption of constant residual standard deviation can be checked by looking at the spread of the points in figure 1 for a given RM. If it appears that this spread increases with the accepted values of the RMs, then the assumption of constant residual standard deviation is probably not correct (this does not seem to be the case for figure 1). A more sophisticated plot to check the assumptions of linearity and of constant residual standard deviation is presented in 6.3.

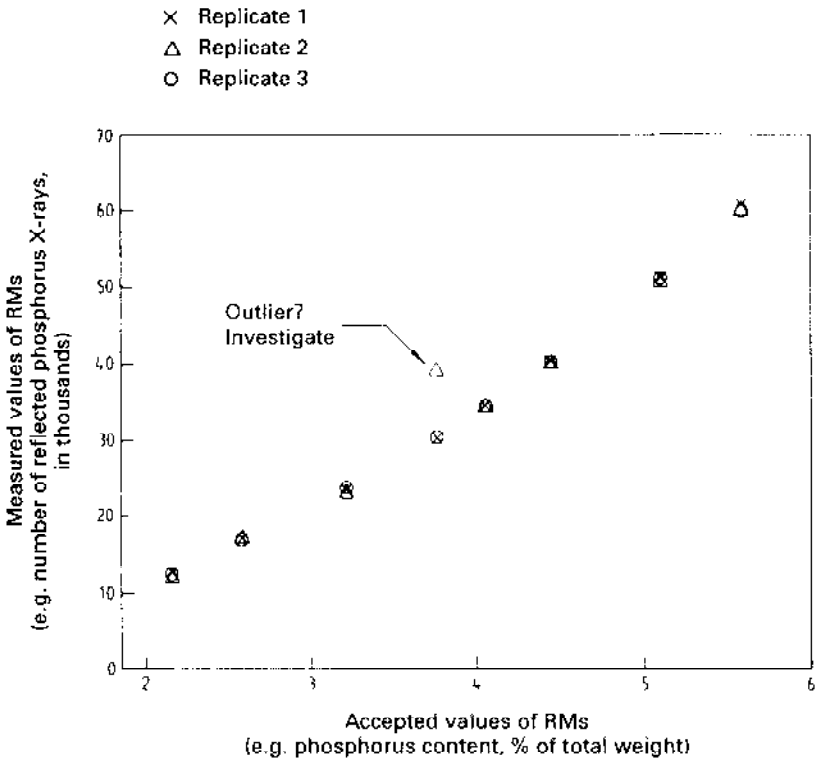


Figure 1 — Schematic diagram of data collected during the calibration experiment

6.2 Estimation of the linear calibration function under the assumption of constant residual standard deviation

6.2.1 Model

The assumptions of linearity of the calibration function and of constant residual standard deviation are captured by the model

$$y_{nk} = \beta_0 + \beta_1 x_n + \varepsilon_{nk}$$

where

x_n is the accepted value of the n^{th} RM ($n = 1, \ldots, N$);

y_{nk} is the k^{th} measurement of the n^{th} RM ($k = 1, \ldots, K$);

$\beta_0 + \beta_1 x_n$ represents the expected value of the measurements of the n^{th} RM;

ε_{nk} is the deviation between y_{nk} and the expected value of the measurement of the n^{th} RM (these deviations are assumed to be independent and normally distributed with mean 0 and with variance σ^2);

β_0, β_1 and σ^2 are three parameters to be estimated from the data collected during calibration:

- β_0 is the intercept of the calibration function,
- β_1 is its slope,
- σ^2 is a measure of the precision of the measurement system.

6.2.2 Estimates of the parameters

Estimates of the parameters β_0, β_1 and σ^2 can be obtained by using the formulae below or by running a linear regression software package with two columns of equal length as input, one for y and one for x.

NOTE 3 Estimates of parameters in this International Standard have a symbol ^ to differentiate them from the parameters themselves which are unknown.

$$\hat{\beta}_1 = \frac{\sum_{n=1}^{N} (x_n - \bar{x})(y_{n\cdot} - \bar{y})}{\sum_{n=1}^{N} (x_n - \bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\sigma}^2 = \frac{\mathrm{SSE}}{(NK - 2)}$$

where

$$\bar{x} = \frac{1}{N} \sum_{n=1}^{N} x_n$$

$$y_{n\cdot} = \frac{1}{K} \sum_{k=1}^{K} y_{nk}$$

$$\bar{y} = \frac{1}{N} \sum_{n=1}^{N} y_{n\cdot}$$

$$NK = N \times K$$

$$\hat{y}_n = \hat{\beta}_0 + \hat{\beta}_1 x_n$$

$$e_{nk} = y_{nk} - \hat{y}_n$$

$$\mathrm{SSE} = \sum_{n=1}^{N} \sum_{k=1}^{K} (e_{nk})^2$$

6.3 Plots of the calibration function and of the residuals

Figures 2 and 3 are recommended to test departures from the assumptions embedded in the model of 6.2.

6.3.1 Plot of the calibration function

In figure 2 the estimated calibration function is added to figure 1.

The plot shown as figure 2 primarily allows a check of the calculations given in 6.2.2. It also provides a visual check of the assumption of linearity of the calibration function.

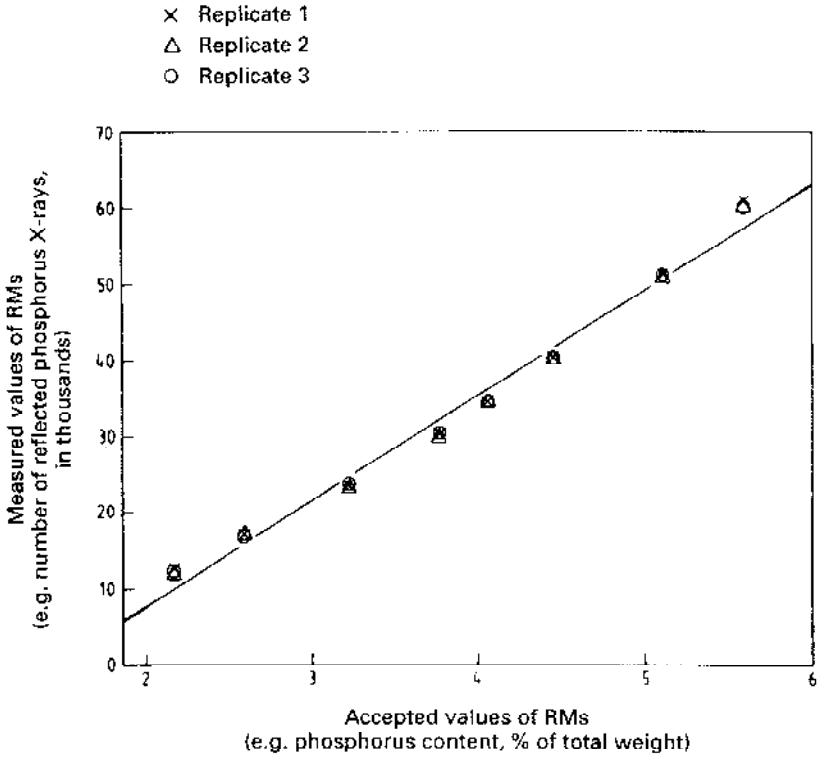


Figure 2 — Schematic diagram of a calibration curve

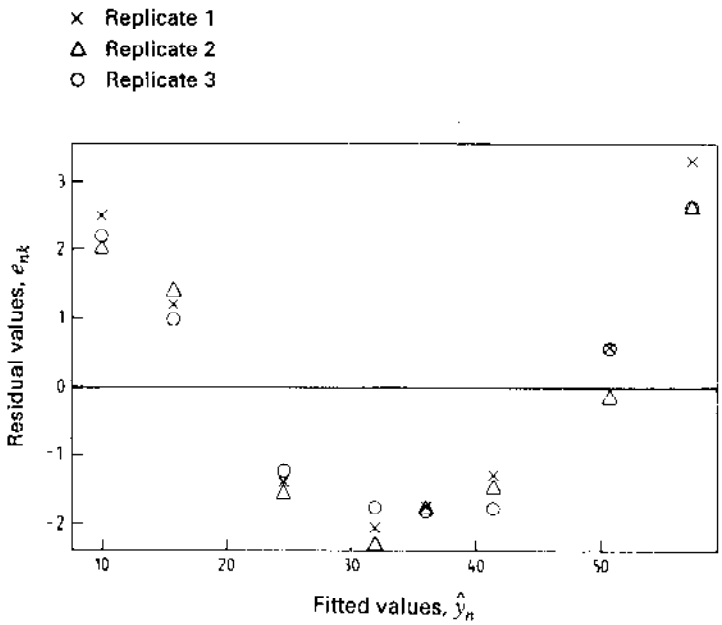


Figure 3 — Schematic diagram of a plot of residuals versus fitted values

6.3.2 Plot of the residuals versus the fitted values

The plot of the residuals e_{nk} versus the fitted values $\hat{y}_n$ (figure 3) is a powerful tool to detect departure from the two assumptions of linearity and of constant residual standard deviation. If these two assumptions hold, then figure 3 should display a plot of randomly distributed points centred around zero. Departure from the assumption of linearity is indicated by a systematic pattern between the residuals and the fitted values (as is the case in figure 3). Departure from the assumption of constant residual standard deviation is indicated by a dispersion in the data that increases or decreases with the fitted values. In figure 3, the dispersion of the residuals for any fitted value is almost constant throughout. Therefore, the assumption of constant residual standard deviation is tenable in this situation.

NOTE 4 Figure 8 illustrates the situation where the assumption of constant residual standard deviation is not tenable.

If the assumption of constant residual standard deviation does not hold, then the data collected during the calibration experiment must be re-analysed. A plot of the standard deviation of the replicated measurements of an RM versus the accepted value of that RM will indicate whether the assumption of proportional residual standard deviation is tenable. See figure 9 for such a plot.

a) If the assumption of proportional residual standard deviation seems to hold, then the data can be re-analysed according to step 6.4.

b) If the assumption of proportional residual standard deviation does not hold but there exists a model relating the residual standard deviation to the accepted values of the RMs (for example inverse proportionality), then an approach similar to the one presented in step 6.4 can be used.

If the assumption of linearity does not hold, then an alternative is to use the bracketing technique described in 8.3.

NOTE 5 There exist other techniques, beyond the scope of this International Standard, that allow the fitting of a quadratic or polynomial curve to these data (see refs. [2] and [3]).

Finally, testing the assumptions of independence and of normality of the ε_{nk} values is beyond the scope of this International Standard. These two assumptions are crucial to the validity of step 6.5 and can also be checked by studying the residuals. For example, a normal probability plot of the residuals allows a check of the normality assumption and a plot of the residuals against time allows a check of the assumption of independence of the measurements. Further information can be found in ref. [3].

6.4 Estimation of the calibration function under the assumption of proportional residual standard deviation and plot of the calibration function and the residuals

6.4.1 Model

An alternate model to the one given in step 6.2.1 is one where the calibration function is linear but the residual standard deviation increases with the accepted values of the RMs. This is captured in the model

$$y_{nk} = \gamma_0 + \gamma_1 x_n + \eta_{nk}$$

where

x_n is the accepted value of the n^{th} RM ($n = 1, \ldots, N$);

y_{nk} is the k^{th} measurement of the n^{th} RM ($k = 1, \ldots, K$);

$\gamma_0 + \gamma_1 x_n$ represents the expected value of the measurement of the n^{th} RM;

η_{nk} is the deviation between y_{nk} and the expected measurement of the n^{th} RM (these deviations are assumed to be independent and normally distributed with mean 0 and with a variance proportional to x_n^2); i.e.

$$\text{var}(\eta_{nk}) = \text{var}(y_{nk}) = x_n^2 \tau^2$$

γ_0, γ_1 and τ^2 are three parameters to be estimated from the data collected during calibration:

- γ_0 and γ_1 are, respectively, the intercept and the slope of the calibration function,
- τ^2 is a measure of the relative precision of the measurement system.

This model can be transformed into a model equivalent to the one given in 6.2.1; i.e. with errors having

constant variance. The transformation consists of dividing by x_n both sides of the equation

$$y_{nk} = \gamma_0 + \gamma_1 x_n + \eta_{nk}$$

This gives

$$\frac{y_{nk}}{x_n} = \frac{\gamma_0}{x_n} + \gamma_1 + \frac{\eta_{nk}}{x_n}$$

or, equivalently,

$$z_{nk} = \gamma_1 + \gamma_0 w_n + \varepsilon_{nk}$$

where

$$z_{nk} = y_{nk}/x_n$$

$$w_n = 1/x_n$$

$$\varepsilon_{nk} = \eta_{nk}/x_n$$

The new model can be analysed as in 6.2 after making the correct substitutions of terms.

6.4.2 Estimates of the parameters

The estimates of the parameters γ_0, γ_1 and τ^2 can be obtained by using the formula below or by running a weighted linear regression software package with three columns of equal length as input, one for y, one for x, and one for the weights ($= 1/x^2$). The same outputs can also be obtained by using a linear regression software package without weights but with the two input columns being z and w.

$$\hat{\gamma}_0 = \frac{\sum_{n=1}^{N} (w_n - \overline{w})(z_{n\cdot} - \overline{z})}{\sum_{n=1}^{N} (w_n - \overline{w})^2}$$

$$\hat{\gamma}_1 = \overline{z} - \hat{\gamma}_0 \overline{w}$$

$$\hat{\tau}^2 = \frac{\text{WSSE}}{(NK - 2)}$$

where

$$NK = N \times K$$

$$z_{nk} = \frac{y_{nk}}{x_n}$$

$$w_n = \frac{1}{x_n}$$

$$\overline{w} = \frac{1}{N} \sum_{n=1}^{N} w_n$$

$$z_{n\cdot} = \frac{1}{K} \sum_{k=1}^{K} z_{nk}$$

$$\overline{z} = \frac{1}{N} \sum_{n=1}^{N} z_{n\cdot}$$

$$\hat{z}_n = \hat{\gamma}_1 + \hat{\gamma}_0 w_n$$

$$u_{nk} = z_{nk} - \hat{z}_n$$

$$\text{WSSE} = \sum_{n=1}^{N} \sum_{k=1}^{K} (u_{nk})^2$$

6.4.3 Plot of the calibration function and residuals

As in 6.3, two plots are recommended:

a) a plot of the estimated calibration function $\hat{y} = \hat{\gamma}_0 + \hat{\gamma}_1 x$ with the data of figure 1;

b) a plot of the weighted residuals u_{nk} versus the weighted fitted values $\hat{z}_n$.

The interpretation of these plots is the same as that for figures 2 and 3.

6.5 Evaluation of the lack of fit of the calibration function

6.5.1 General

A comparison between

— the variability due to lack of fit of the model selected either in 6.2 or in 6.4 and

— the variability of the pure error representing the inability of the system to repeat measurements exactly

is carried out after constructing an ANOVA table. Such a comparison is possible because the measurements of each RM have been replicated.

The selection of the significance level α depends on particular applications and is left to the user of this International Standard.

6.5.2 Model with constant residual standard deviation (defined in 6.2)

6.5.2.1 The ANOVA table shown as table 1 can be obtained by using the formulae below or as an output of most linear regression software packages.

Table 1 — ANOVA table to compare lack of fit and pure error under the assumption of constant residual standard deviation

Source	Degrees of freedom, DF	Sum of squares, SS	SS/DF	F ratio
Calibration function	1	SSR = SST − SSE		
Residual	$NK-2$	SSE	$\hat{\sigma}^2 = \frac{\text{SSE}}{NK-2}$	
Lack of fit	$N-2$	SSE − SSP	$\hat{\sigma}_l^2 = \frac{\text{SSE}-\text{SSP}}{N-2}$	$\frac{\hat{\sigma}_l^2}{\hat{\sigma}_p^2}$
Pure error	$NK-N$	SSP	$\hat{\sigma}_p^2 = \frac{\text{SSP}}{NK-N}$	
Total	$NK-1$	SST		

$\text{SST} = \sum_{n=1}^{N}\sum_{k=1}^{K}(y_{nk}-\bar{y})^2$

$\text{SSP} = \sum_{n=1}^{N}\sum_{k=1}^{K}(y_{nk}-y_n)^2$

SSE is defined in 6.2.2.

6.5.2.2 The variability due to pure error is estimated by $\hat{\sigma}_p^2$. This variability is independent of the model ($y = \beta_0 + \beta_1 x$) fitted to the data. The variability due to lack of fit is estimated by $\hat{\sigma}_l^2$. A test of the validity of the model defined in 6.2.1 is carried out by comparing $\hat{\sigma}_l^2/\hat{\sigma}_p^2$ to $F_{(1-\alpha)}(N-2;\ NK-N)$ where $F_{(1-\alpha)}(N-2;\ NK-N)$ is the $(1-\alpha)$-quantile of the F-distribution with $N-2$ and $NK-N$ degrees of freedom.

a) If $\hat{\sigma}_l^2/\hat{\sigma}_p^2$ is not larger than $F_{(1-\alpha)}(N-2;\ NK-N)$, then there is no evidence to reject the linear model.

b) If $\hat{\sigma}_l^2/\hat{\sigma}_p^2$ is larger than $F_{(1-\alpha)}(N-2;\ NK-N)$, then potential causes for a large variability due to lack of fit relative to the pure error variability should be investigated. One common cause is the inadequacy of the linear assumption of the calibration function (see figures 2 and 3). Another possible cause may be the conditions under which the calibration experiment was performed (e.g. replications may not have been genuine repeats but just repetitions of the same reading).

6.5.3 Model with proportional residual standard deviation (defined in 6.4)

If the model with proportional residual standard deviation was used, then the ANOVA table is constructed as shown in table 2.

The same test, interpretation, conclusions and remarks apply to $\hat{\tau}_l^2/\hat{\tau}_p^2$ as to $\hat{\sigma}_l^2/\hat{\sigma}_p^2$ described in 6.5.2.2.

6.6 Transformation of future measured values with the calibration function

Once a calibration experiment has been carried out, measured values of new unknown quantities (in opposition to standards which have known or accepted values) will be transformed via the calibration function. Transforming these measured values will result in a single value x_0^* that estimates the true value of the unknown quantity. The transformation depends on the assumption made concerning the residual variance and is implemented as follows.

A new unknown quantity is measured p times, resulting in p measurements $y_{01}, y_{02}, \ldots, y_{0p}$. The mean $\bar{y}_0$ of these p measurements is obtained as

$$\bar{y}_0 = \frac{1}{p}\sum_{k=1}^{p} y_{0k}$$

If $p = 1$, then $y_0 = y_{01}$.

Table 2 — ANOVA table to compare lack of fit and pure error under the assumption of proportional residual standard deviation

Source	Degrees of freedom, DF	Sum of squares, SS	SS/DF	F ratio
Calibration function	1	WSSR = WSST − WSSE		
Residual	$NK - 2$	WSSE	$\hat{\tau}^2 = \frac{\text{WSSE}}{NK-2}$	
Lack of fit	$N - 2$	WSSE − WSSP	$\hat{\tau}_l^2 = \frac{\text{WSSE} - \text{WSSP}}{N-2}$	$\frac{\hat{\tau}_l^2}{\hat{\tau}_p^2}$
Pure error	$NK - N$	WSSP	$\hat{\tau}_p^2 = \frac{\text{WSSP}}{NK-N}$	
Total	$NK - 1$	WSST		

$$\text{WSST} = \sum_{n=1}^{N} \sum_{k=1}^{K} (z_{nk} - \bar{z})^2$$

$$\text{WSSP} = \sum_{n=1}^{N} \sum_{k=1}^{K} (z_{nk} - z_n)^2$$

WSSE is defined in 6.4.2.

a) If the model with constant residual standard deviation was selected, then

$$x_0^* = \frac{\bar{y}_0 - \hat{\beta}_0}{\hat{\beta}_1}$$

b) If the model with proportional residual standard deviation was selected, then

$$x_0^* = \frac{\bar{y}_0 - \hat{\gamma}_0}{\hat{\gamma}_1}$$

This International Standard does not provide confidence intervals, either one-at-a-time (see ref. [2]) or simultaneous (see refs. [4] and [5]), for the estimates of new unknown quantities based on the calibration experiment itself. Instead, this International Standard offers a control method which, among other benefits, allows for the derivation of confidence intervals based on the variability observed in monitoring a few RMs over a period of time.

7 Control method

7.1 General

When the calibration function is to be used for an extended period of time, it is desirable to implement a control method to check the validity of the calibration curve, as well as to identify, and subsequently eliminate, sources of undesired variation. The control method monitors on a regular basis the measurement system in order to detect quickly when the system behaves erratically or shifts, thus potentially making the calibration function useless if not harmful.

Detection is achieved by monitoring the measured values (after transformations by the calibration function) of a set of m RMs with a control chart technique.

NOTE 6 This approach in an extension of traditional control charts described in ISO 7870[6] and ISO 8258[7].

The control chart is first established from the data collected during the calibration experiment. The control chart is then used to decide if the calibration function needs to be re-estimated. The same control chart is also used to estimate the uncertainty in the measurements after they have been transformed with the calibration function.

7.2 Calculation of upper and lower control limits

7.2.1 Model with constant residual standard deviation

a) Calculate the upper control limit U_d and the lower control limit L_d as

$$U_d = \frac{\hat{\sigma}}{\hat{\beta}_1} t_{(1-\zeta/2)}(NK-2)$$

$$L_d = -\frac{\hat{\sigma}}{\hat{\beta}_1} t_{(1-\zeta/2)}(NK-2)$$

where

$\hat{\sigma}$ is the square root of the estimate $\hat{\sigma}^2$ obtained from the calibration experiment (see 6.2.2);

$NK-2$ is the number of degrees of freedom associated with the estimation of σ^2 (see 6.5);

$\hat{\beta}_1$ is the estimate of β_1 obtained from the calibration experiment (see 6.2.2);

α is the significance level selected for the control chart;

$t_{(1-\zeta/2)}(NK-2)$ is the $(1-\zeta/2)$-quantile of the t-distribution with $NK-2$ degrees of freedom; i.e.

$$P[t > t_{(1-\zeta/2)}(NK-2)] = \zeta/2$$

ζ is the significance level associated with each individual RM and with the limits U_d and L_d such that the overall significance of α is obtained for all the m RMs simultaneously; ζ is obtained (for small values of α) as

$$\zeta = 1 - \exp\left(\frac{\ln(1-\alpha)}{m}\right)$$

$$\approx \frac{\alpha}{m}$$

b) Plot the limits U_d and L_d on the control chart.

7.2.2 Model with proportional residual standard deviation

a) Calculate the upper control limit U_c and the lower control limit L_c as

$$U_c = \frac{\hat{\tau}}{\hat{\gamma}_1} t_{(1-\zeta/2)}(NK-2)$$

$$L_c = -\frac{\hat{\tau}}{\hat{\gamma}_1} t_{(1-\zeta/2)}(NK-2)$$

where

$\hat{\tau}$ is the square root of the estimate $\hat{\tau}^2$ obtained from the calibration experiment (see 6.4.2);

$NK-2$ is the number of degrees of freedom associated with the estimation of τ^2 (see 6.5);

$\hat{\gamma}_1$ is the estimate of γ_1 obtained from the calibration experiment (see 6.4.2).

NOTE 7 α, $t_{(1-\zeta/2)}(NK-2)$ and ζ are as defined in 7.2.1 a).

b) Plot the limits U_c and L_c on the control chart.

7.3 Collection and plotting of the data

7.3.1 Select m RMs such that their accepted values cover the range of values encountered under normal operating conditions of the measurement system. A minimum of two RMs is needed. Three RMs are recommended. It is preferable but not mandatory to use RMs that are different from the ones used during the calibration experiment.

7.3.2 On a regular basis (e.g. once a day or once every shift), make one measurement on each of these m RMs.

7.3.3 Obtain the transformed values of each one of these m RMs (see 6.6). These transformed values are referred to as x_i^* for $i = 1, \ldots, m$.

7.3.4 Calculate the differences d_i between the transformed values x_i^* and the accepted values of these RMs, x_i, as

$$d_i = x_i^* - x_i$$

7.3.5 If the calibration model assumes constant residual standard deviation, let the differences d_i be referred to as the control values.

If the calibration model assumes proportional residual standard deviation, normalize the differences d_i by dividing them by x_i. Let the resulting values c_i be referred to as the control values where

$$c_i = \frac{x_i^* - x_i}{x_i}$$

7.3.6 Plot the appropriate control values (d_i or c_i) versus the time at which the m RMs are measured on the control chart. Figure 4 illustrates a control chart for the constant residual standard deviation model. A similar control chart can be drawn for the proportional residual standard deviation model (see figure 12).

7.4 Decision about the state of the system

If one or more values of d_i falls outside the control limits U_d and L_d for the model with constant residual standard deviation, the system is declared out-of-control at that time. The m RMs should be remeasured. If at least one of the new measurements of the m RMs is still outside the limits, an investigation shall take place at this point to determine the cause of the problem. Depending on the nature of the problem, the calibration function may need to be re-estimated from a new calibration experiment.

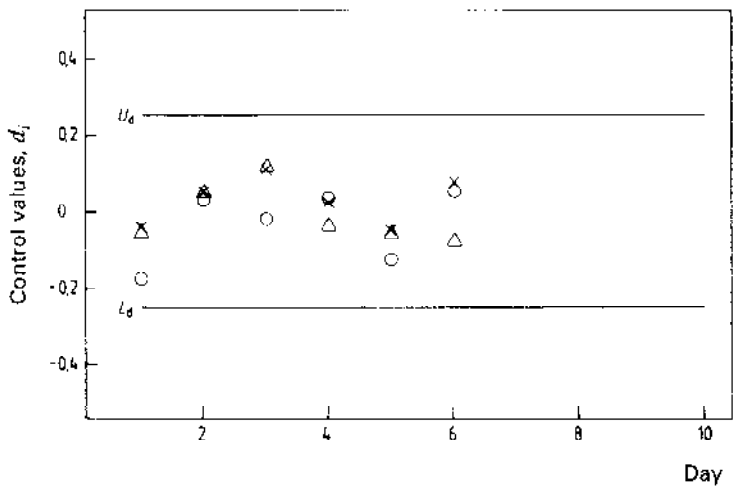
The same conclusions are reached for the model with proportional residual standard deviation by comparing the c_i values to the limits U_c and L_c.

7.5 Estimation of the uncertainty of the transformed values

7.5.1 Estimation during the validity period of a given calibration function

For the calibration function subject to the control method, the uncertainty of the transformed values is approximated by the pooled variance of the control values of two RMs (out of the m RMs selected for the control method): the RMs with smallest and largest values. This is explained by the fact that the transformed values at the end of the range of values encountered during the calibration experiment are expected to have a larger variance than the ones in the middle of that range. Thus, the confidence interval for a transformed value derived from the variability of the two extreme RMs is approximately correct for the values at the end of the range of applications and conservative for the values in the middle of that range.

To calculate such a confidence interval, carry out the procedure given in 7.5.1.1 for the appropriate model (constant or proportional residual standard deviation).



Key: x = RM with low value, Δ = RM with middle value, o = RM with upper value

Figure 4 — Schematic diagram of a control chart to validate the calibration curve under the assumption of constant residual standard deviation

7.5.1.1 Model with constant residual standard deviation

Let d_{lj} and d_{mj} be the control values of the smallest and largest RMs, where j represents the time at which the measurements were made. Then, over a period of J times when the measurement system is in a state of statistical control, the standard deviation of a transformed value is approximated by

$$\hat{\sigma}_{cal} = \sqrt{\frac{\sum_{j=1}^{J}(d_{lj}^2 + d_{mj}^2)}{2J}}$$

with $2J$ degrees of freedom.

An approximate confidence interval for the unknown true value of a quantity estimated by the transformed value x_0^* (derived from p measurements made over a short period of time) with a confidence level of $(1 - \alpha)$ is obtained as

$$x_0^* \pm \hat{\sigma}_{cal}\, t_{(1-\alpha/2)}(2J)$$

where $t_{(1-\alpha/2)}(2J)$ is the $(1 - \alpha/2)$-quantile of the t-distribution with $2J$ degrees of freedom.

7.5.1.2 Model with proportional residual standard deviation

Let c_{lj} and c_{mj} be the control values of the smallest and largest RMs, where j represents the time at which the measurements were made. Then, over a period of J times when the measurement system has been in a state of statistical control, the coefficient of variation of a transformed value is approximated by

$$\hat{\tau}_{cal} = \sqrt{\frac{\sum_{j=1}^{J}(c_{lj}^2 + c_{mj}^2)}{2J}}$$

with $2J$ degrees of freedom.

An approximate confidence interval for the unknown true value of a quantity estimated by the transformed value x_0^* (derived from p measurements made over a short period of time) with a confidence level of $(1 - \alpha)$ is obtained as

$$x_0^* \pm \hat{\tau}_{cal}\, t_{(1-\alpha/2)}(2J)\, x_0^*$$

where $t_{(1-\alpha/2)}(2J)$ is the $(1 - \alpha/2)$-quantile of the t-distribution with $2J$ degrees of freedom.

7.5.2 Estimation over a period that includes recalibration

To insure that the variability due to the calibration procedure is included in the uncertainty statement pick one set of control values (d_{lj}, d_{mj}) or (c_{lj}, c_{mj}) from each calibration interval and use the same formula for $\hat{\sigma}_{cal}$ or $\hat{c}_{cal}$, where j is now the number of recalibrations.

8 Two alternatives to the basic method

8.1 General

Under special conditions, two alternative methods can be used to calibrate a measurement process. These two methods are actually special cases of the basic method, where only one or two RMs are used. The one-point calibration method is a fast technique that allows one to "recalibrate" a measurement system when there is no doubt about the linearity of the calibration function. The bracketing method is a labour-intensive technique that allows the determination of the value of an unknown quantity with great precision and with a minimal set of assumptions.

8.2 One-point calibration method

8.2.1 General

This method is useful for a quick recalibration when there is no doubt about the linearity of the function over a given range $[0, ..., M]$. The "zero-point" is obtained by adjusting some dials to ensure that an unknown quantity with true value 0 is measured as 0. Only a blank (quantity with true value 0) and one RM are used in this method.

One can note that, historically, this method is called a one-point calibration but in reality this is a two-point calibration carried out with one blank and one RM. This so-called "one-point calibration" is a weak and uncertain method because of the doubtfulness of the zero point. It should not be recommended for calibration purposes, but primarily for checking an existing linear calibration function.

8.2.2 Assumptions

It is assumed that:

a) there are no errors in the accepted value of the only RM and of the blank used with this method (assumption not tested);

b) the calibration function is linear over the range $[0, ..., M]$ (assumption not tested);

c) the residual standard deviation is constant (assumption not tested).

8.2.3 One-point calibration experiment

a) Experimental conditions: the experimental conditions should be the same as the normal operating conditions of the measurement system.

b) Choice of RM: the only RM used in this experiment should have an accepted value that is greater (as far as is possible) than the values encountered during normal conditions of the measurement system.

c) Number of replicates: the RM should be measured at least twice.

8.2.4 Estimation of the calibration function

8.2.4.1 Model

This model is similar to the one of the basic method with constant residual variance defined in 6.2.1, but without an intercept. This model is

$$y_k = \beta x + \varepsilon_k$$

where

x is the accepted value of the only RM used;

y_k is the k^{th} measurement of that RM ($k = 1, ..., K$);

ε_k is the deviation between y_k and the expected value of the measurement of the RM (these deviations are assumed to be independent and normally distributed with mean 0 and variance σ^2);

β and σ^2 are two parameters to be estimated from the data collected during the experiment.

8.2.4.2 Estimates of β and σ^2

These are obtained from the following formulae:

$$\hat{\beta} = \frac{\bar{y}}{x}$$

$$\hat{\sigma}^2 = \frac{1}{K-1}\sum_{k=1}^{K}(y_k - \bar{y})^2$$

where

$$\bar{y} = \frac{1}{K}\sum_{k=1}^{K} y_k$$

8.2.4.3 Plot of the data

Plot the data collected during the experiment, as shown in figure 5.

The plot shown as figure 5 allows visual identification of potential outliers for investigation. It also displays the linear calibration function constrained to go through the origin.

8.2.5 Transformation of future measurements with the calibration function

An unknown quantity is measured p times, resulting

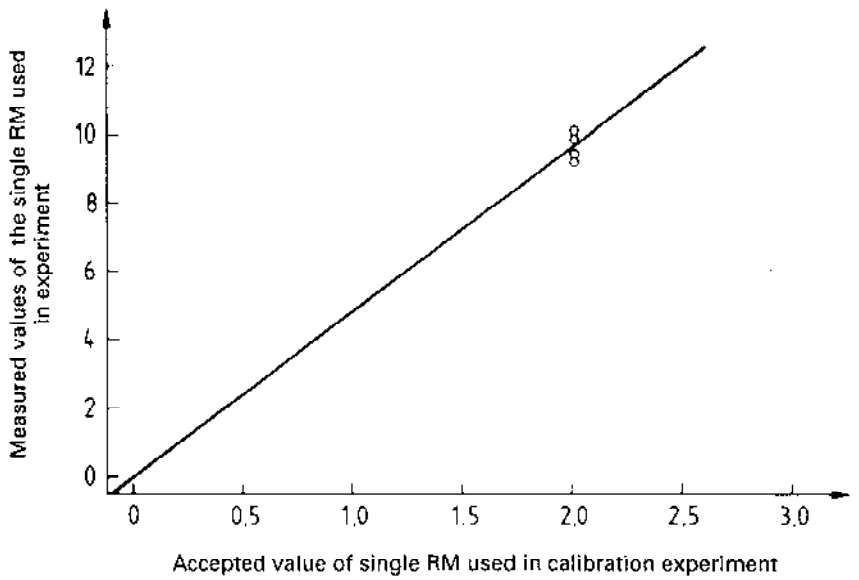


Figure 5 — Schematic diagram of the data in a one-point calibration experiment

in p measurements y_{01}, y_{02}, ..., y_{0p}. The mean $\bar{y}_0$ of these p measurements is obtained as

$$\bar{y}_0 = \frac{1}{p} \sum_{k=1}^{p} y_{0k}$$

If $p = 1$, then $\bar{y}_0 = y_{01}$. Transforming these measurements will result in a single value being reported as $x_0^* = \bar{y}_0/\hat{\beta}$.

NOTE 8 In principle, the blank does not always have a true value of 0 but instead has an accepted value of x_b, known to have a measurement of y_b. If x_b is not negligible, the one-point calibration method described in 8.2.3 can be used with the following adaptations.

a) Measure the blank and adjust the dials of the measuring instrument to read y_b.

b) Measure the only RM used, as in the case of a blank with value 0.

c) The model becomes

$$y_k - y_b = \beta(x - x_b) + \varepsilon_k$$

d) The estimate of β becomes

$$\hat{\beta} = (y - y_b)/(x - x_b)$$

e) The estimate of σ^2 is unchanged.

f) The estimate of the true value of an unknown quantity measured p times (y_{01}, y_{02}, ..., y_{0p}) is

$$x_0^* = x_b + (\bar{y}_0 - y_b)/\hat{\beta}$$

8.3 Bracketing technique

8.3.1 General

This method is useful when there is some doubt about the linearity of the calibration function over the full range of values encountered during normal operations of the measurement system. This method is also useful when there is some concern about the stability of the measurement process. The principle of the method consists in reducing as much as possible the interval over which the linearity of the calibration function is assumed. This leads to surrounding as tightly as possible (or bracketing) the value of the unknown quantity by two values of reference materials (RMs). Because of the tight surrounding of each unknown quantity by two RMs, and because of the short period of time needed for this procedure (time to measure the unknown quantity and the two RMs), the bracketing technique usually yields greater accuracy in determining the transformed value of an unknown quantity.

The unknown quantity and the two RMs are measured together. The value of the unknown quantity is estimated directly, based on a linear interpolation between the values of the two RMs.

8.3.2 Assumptions

Because only two RMs are used, the bracketing technique does not allow for checking of any of the following assumptions:

a) that there is no error in the accepted values of the RMs;

b) that the calibration function between the two RMs is linear;

c) that the residual standard deviation is constant.

8.3.3 Bracketing experiment

a) Experimental conditions: the experimental conditions should be such that the variability between measurements of the same RMs is as small as possible.

b) Choice of RMs: the range of values spanned by the two RMs should be as small as possible and shall include the value of the unknown quantity to be measured.

c) Number of RMs: two RMs are used for each unknown quantity.

d) Number of replicates: both RMs and the unknown quantity should be measured at least twice.

8.3.4 Estimation of the unknown quantity

8.3.4.1 Model

The model is the same as that for the basic method with constant residual standard deviation (see 6.2), i.e.

$$y_{ik} = \beta_0 + \beta_1 x_i + \varepsilon_{ik}$$

where

i is an index that refers to both RMs ($i = 1, 2$) as well as to the unknown quantity ($i = 0$);

x_1 and x_2 are the accepted values of the RMs;

x_0 is the unknown true value of the unknown quantity;

y_{1k}, y_{2k} and y_{0k} are the measurements of the two RMs and of the unknown quantity, respectively ($k = 1, ..., K$);

ε_{ik} is the deviation between y_{ik} and the expected value of the measurement of either an RM or the unknown quantity (depending on the value of i) (these deviations are assumed to be randomly distributed with mean 0 and variance σ^2);

β_0, β_1, x_0 and σ^2 are the four parameters to be estimated from the data collected during the bracketing experiment (there is no interest in β_0 and β_1 except for the fact that they impact the parameter x_0).

8.3.4.2 Estimates of x_0 and of the residual variance σ^2

These are obtained from the following formulae:

$$\hat{x}_0 = \frac{x_2(\bar{y}_0 - \bar{y}_1) - x_1(\bar{y}_0 - \bar{y}_2)}{\bar{y}_2 - \bar{y}_1}$$

$$\hat{\sigma}^2 = \frac{\sum_{k=1}^{K}(y_{1k} - \bar{y}_1)^2 + \sum_{k=1}^{K}(y_{2k} - \bar{y}_2)^2 + \sum_{k=1}^{K}(y_{0k} - y_0)^2}{3(K-1)}$$

where

$$\bar{y}_i = \frac{1}{K}\sum_{k=1}^{K} y_{ik}$$

when $i = 0, 1, 2$

9 Example

9.1 General

This example illustrates the basic method for estimating a linear calibration function for a measurement system and the control method for monitoring the same measurement system. The example is based on ref. [8].

9.2 Basic method

9.2.1 Background and data

Measurements of line-spacing in integrated circuits photomasks in the 0,5 µm to 12 µm range can be made with an optical-imaging system (an optical microscope fitted with a measurement attachment). Such a system can be calibrated using the standard reference material SRM-474 issued by the National Institute of Standards and Technology (NIST). SRM-474 contains (among other things) a row of ten randomly arranged spacings in the range of 0,5 µm to 12 µm.

This example describes a calibration experiment conducted on an optical-imaging system. Each one of the ten line-spacings of the standard was measured four times. These repetitions were spaced over a 2-week interval to ensure independence among the measurements. The data displayed in table 3 consist of four ($K = 4$) replicates of measurements on ten ($N = 10$) line-spacings for which NIST provides accepted values.

9.2.2 Plot of the data

The plot of the data collected during the experiment, as shown in figure 6, does not identify obvious outliers or unusual behaviour of the system during the calibration experiment. It supports the assumption of linearity of the calibration function and raises questions as to the assumption of constant residual standard deviation, since the spread of the data for a given NIST value seems to increase slightly with that NIST value.

Table 3 — Calibration experiment for line-spacing

Values in micrometres

NIST value	Measured value			
	Replicate 1	Replicate 2	Replicate 3	Replicate 4
x_n	y_{n1}	y_{n2}	y_{n3}	y_{n4}
6,19	6,31	6,27	6,31	6,28
9,17	9,27	9,21	9,34	9,23
1,99	2,21	2,19	2,22	2,20
7,77	8,00	7,81	7,95	7,84
4,00	4,27	4,15	4,15	4,15
10,77	10,93	10,73	10,92	10,89
4,78	4,95	4,87	5,00	5,00
2,99	3,24	3,17	3,21	3,21
6,98	7,14	7,07	7,18	7,20
9,98	10,23	10,02	10,07	10,17

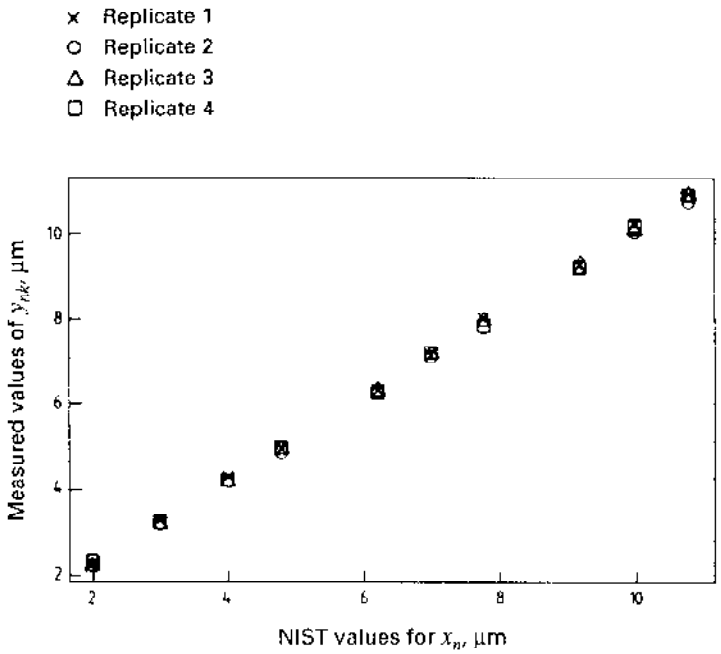


Figure 6 — Data collected during the calibration experiment for line-spacing

9.2.3 Estimation of the linear calibration function under the assumption of constant residual standard deviation

The formulae given in 6.2.2 lead to:

a) $N = 10,\ K = 4$

b) $\bar{x} = 6{,}462$

c) $y_{i\cdot}$ is as given in table 4

d) $\bar{y} = 6{,}614$

e) SSE = 0,146 2

f) $\hat{\beta}_1 = 0{,}987\ 0$

g) $\hat{\beta}_0 = 0{,}235\ 8$

h) $\hat{\sigma}^2 = 0{,}003\ 8$

The calibration function is

$$\hat{y} = 0{,}235\ 8 + 0{,}987x$$

The fitted values $\hat{y}_n$ are obtained by replacing x in this formula with the NIST values x_n listed in table 3.

The residuals are obtained as

$$e_{nk} = y_{nk} - \hat{y}_n$$

These residuals are listed in table 5.

Table 4 — Values of $y_{i\cdot}$

i	1	2	3	4	5	6	7	8	9	10
$y_{i\cdot}$	6,292	9,263	2,205	7,900	4,180	10,868	4,955	3,208	7,148	10,12

Table 5 — Linear calibration under the assumption of constant residual standard deviation

Values in micrometres

NIST value	Fitted value	Residual value			
x_n	$\hat{y}_n$	e_{n1}	e_{n2}	e_{n3}	e_{n4}
6,19	6,345 5	−0,035 5	−0,075 5	0,035 5	−0,065 5
9,17	9,286 9	−0,016 9	−0,076 9	0,053 1	−0,056 9
1,99	2,200 0	0,010 0	−0,010 0	0,020 0	0,000 0
7,77	7,905 0	0,095 0	−0,095 0	0,045 0	−0,065 0
4,00	4,183 9	0,086 1	−0,033 9	−0,033 9	−0,033 9
10,77	10,866 2	0,063 8	−0,136 2	0,053 8	0,023 8
4,78	4,953 8	−0,003 8	−0,083 8	0,046 2	0,046 2
2,99	3,187 0	0,053 0	−0,017 0	0,023 0	0,023 0
6,98	7,125 3	0,014 7	−0,055 3	0,054 7	0,074 7
9,98	10,086 4	0,143 6	−0,066 4	−0,016 4	0,083 6

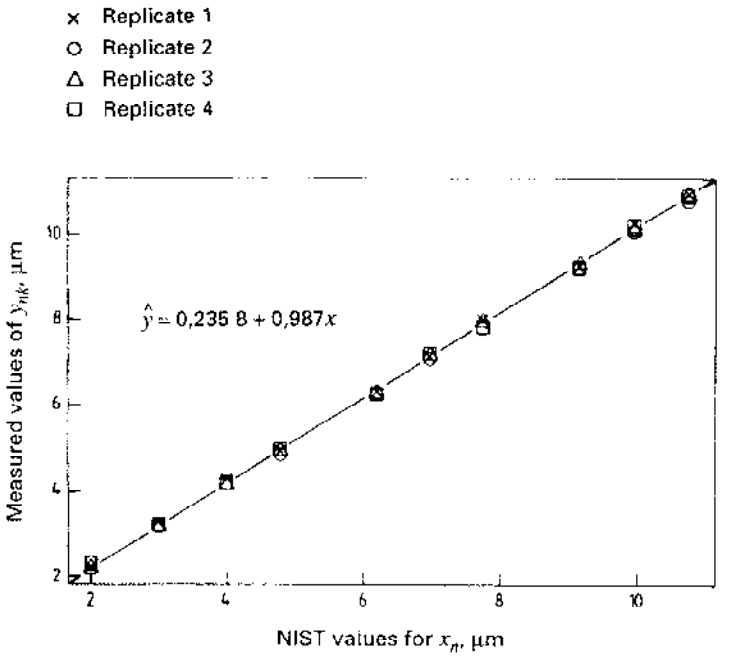


Figure 7 — The calibration curve for line-spacing under the assumption of constant residual standard deviation

9.2.4 Plots of the calibration function and of the residuals

The plot of the calibration function (figure 7) confirms that a linear calibration function seems to be appropriate.

The plot of the residuals (figure 8) shows that replicate 2 has consistently lower residual values than the other replicates. These low residual values can be traced to the original data in table 3 which are consistently lower for replicate 2 than for the other replicates. No definite explanation was found for this phenomenon and the data from replicate 2 were retained as representative of the behaviour of the measurement system under normal operating conditions.

A more sophisticated model than the two models proposed respectively in 6.2.1 and 6.4.1 could be used to analyse these data in order to take into consideration the systematic differences between replicates. For the sake of simplicity and for an illustration of the basic and control methods, this effect will be ignored and the present strategy and associated models will be pursued.

Figure 8 also indicates that the assumption of constant residual standard deviation does not seem to hold. This suggestion is confirmed with figure 9, which shows a plot of the standard deviation of the replicated measurements of a RM versus the accepted values of that RM.

9.2.5 Estimation of the calibration function under the assumption of proportional residual standard deviation

Estimate the calibration function under the assumption of proportional residual standard deviation and plot the calibration function and the residuals.

The formulae given in 6.4.2 lead to:

a) $N = 10$, $K = 4$

b) $\dot{w} = 0{,}203$

c) z_i, is as given in table 6

d) $\bar{z} = 1{,}035$

e) WSSE = 0,003 4

f) $\hat{\gamma}_1 = 0{,}985\ 1$

g) $\hat{\gamma}_0 = 0{,}246\ 9$

h) $\hat{\tau}^2 = 0{,}889 \times 10^{-4}$

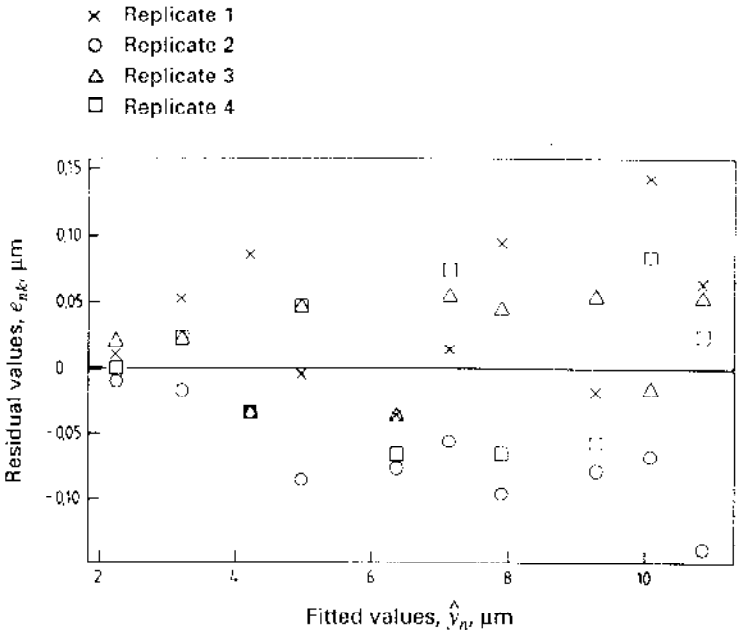


Figure 8 — Residuals versus fitted values for line-spacing under the assumption of constant residual standard deviation

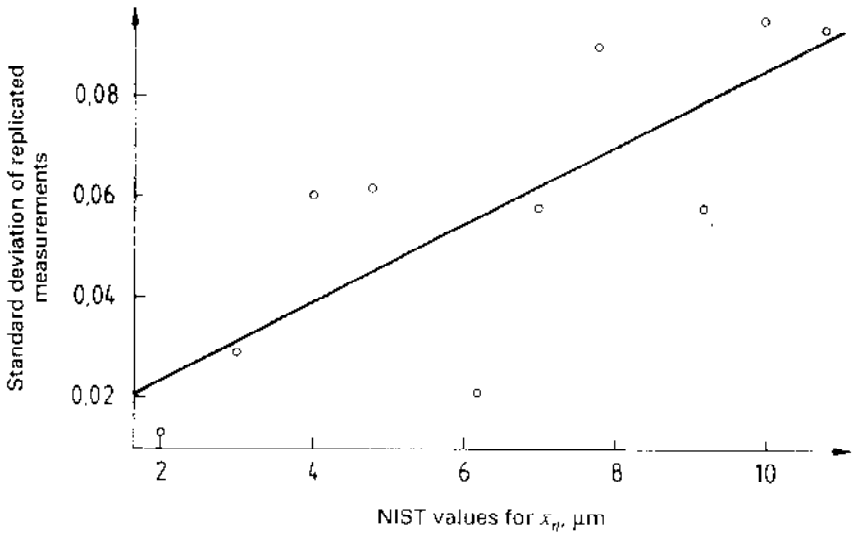


Figure 9 — Standard deviations of replicated measurements for line-spacing versus NIST values

Table 6 — Values of z_i

i	1	2	3	4	5	6	7	8	9	10
z_i	1,017	1,010	1,108	1,017	1,045	1,009	1,037	1,073	1,024	1,014

The calibration function is

$$\hat{y} = 0{,}246\ 9 + 0{,}9851x$$

The fitted values, $\hat{y}_n$, are obtained by replacing x in this formula with the NIST values x_n. These fitted values are listed in table 7.

The weighted fitted values are obtained by replacing x in the formula

$$\hat{z} = 0{,}985\ 1 + 0{,}246\ 9/x$$

by the NIST values x_n.

The weighted residuals are obtained as

$$u_{nk} = z_{nk} - \hat{z}_n$$

These weighted residuals are listed in table 7.

Figure 10 shows the original data and the calibration function under the assumption of proportional residual standard deviation.

Table 7 — Linear calibration under the assumption of proportional residual standard deviation

NIST value x_n µm	Fitted value $\hat{y}_n$ µm	Weighted fitted value $\hat{z}_n$	Weighted residual			
			u_{n1}	u_{n2}	u_{n3}	u_{n4}
6,19	6,344 9	1,025 0	−0,005 6	−0,012 1	−0,005 6	−0,010 5
9,17	9,280 7	1,012 1	−0,001 2	−0,007 7	0,006 5	−0,005 5
1,99	2,207 4	1,109 2	0,001 3	−0,008 7	0,006 4	−0,003 7
7,77	7,901 5	1,016 9	0,012 7	−0,011 8	0,006 2	−0,007 9
4,00	4,187 5	1,046 9	0,020 6	−0,009 4	−0,009 4	−0,009 4
10,77	10,856 9	1,008 1	0,006 8	−0,011 8	0,005 9	0,003 1
4,78	4,955 9	1,036 8	−0,001 2	−0,018 0	0,009 2	0,009 2
2,99	3,192 5	1,067 7	0,015 9	−0,007 5	0,005 9	0,005 9
6,98	7,123 2	1,020 5	0,002 4	−0,007 6	0,008 1	0,011 0
9,98	10,078 6	1,009 9	0,015 2	−0,005 9	−0,000 9	0,009 2

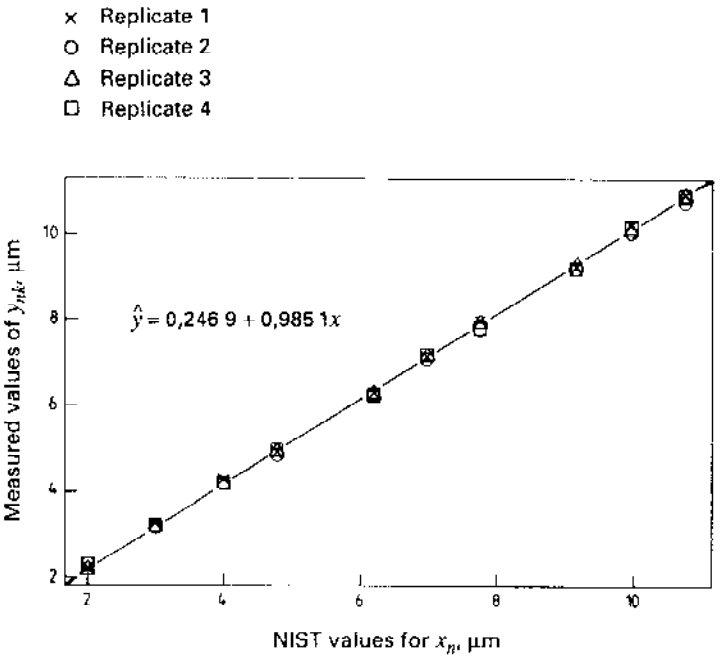


Figure 10 — Calibration curve for line-spacing under the assumption of proportional residual standard deviation

Figure 10, similarly to figure 7, supports the assumption of linearity. The coefficients of the linear calibration function have slightly changed compared to the ones of figure 7. This change is the result of assigning less weight to the measured values for large line-spacings than to the measured values for small line-spacings (assumption of proportional residual standard deviation).

Figure 11 shows a plot of the weighted residuals.

The weighted residuals shown in figure 11 appear to be randomly distributed. The increasing spread of the residuals of figure 8 has disappeared, lending more credence to the assumption of proportional residual standard deviation. As in figure 8, figure 11 shows lower weighted residual values for replicate 2.

9.2.6 Evaluation of the lack of fit of the calibration function

Table 8 shows the ANOVA table under the model of proportional residual standard deviation given in 6.5.3.

The ANOVA table reveals that the variability in the residuals due to lack of fit ($\hat{\tau}_l^2$) is smaller than the variability in the data due to pure error ($\hat{\tau}_p^2$). The ratio $\hat{\tau}_l^2/\hat{\tau}_p^2$ is smaller than the $F_{0,95}(8,30)$ value equal to 2,27. This confirms that the assumption of linearity is appropriate for the calibration experiment described in this example.

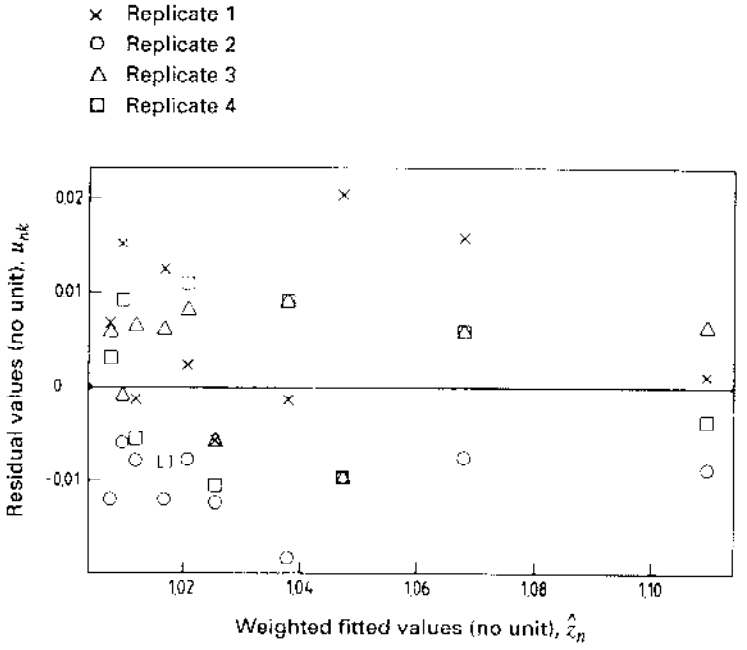


Figure 11 — Weighted residuals versus weighted fitted values for line-spacing under the assumption of proportional residual standard deviation

Table 8 — ANOVA table to compare lack of fit and pure error for line-spacing under the assumption of proportional residual standard deviation

Source	Degrees of freedom, DF	Sum of squares, SS	SS/DF	F ratio
Calibration function	1	WSSR = 0,0369	MSR = 0,0369	
Residual	38	WSSE = 0,0034	$\hat{\tau}^2 = 0{,}89 \times 10^{-4}$	
Lack of fit	8	WSSE − WSSP = 0,00055	$\hat{\tau}_l^2 = 0{,}69 \times 10^{-4}$	$\hat{\tau}_l^2/\hat{\tau}_p^2 = 0{,}73$
Pure error	30	WSSP = 0,0028	$\hat{\tau}_p^2 = 0{,}94 \times 10^{-4}$	
Total	39	WSST = 0,0403		

9.2.7 Tranformation of future measurements

Based on the calibration function obtained in 6.4, measurement or measurements on new unknown line-spacings will be transformed as follows:

a) a single measurement y_0 of an unknown line-spacing will lead to a reported line-spacing value of

$$x_0^* = \frac{y_0 - 0{,}246\ 9}{0{,}985\ 1}$$

b) several measurements of the same unknown line-spacing $y_{01}, y_{02}, \ldots, y_{0p}$ will lead to a single reported line-spacing value of

$$x_0^* = \frac{\bar{y}_0 - 0{,}246\ 9}{0{,}985\ 1}$$

9.3 Control method

9.3.1 Background and data

Two line-spacings were selected for the control method ($m = 2$). These line-spacings were selected in such a way that they cover as large a range as possible of values encountered during normal operating conditions. Each line-spacing was measured every day. Table 9 shows measurements obtained during the first 7 days, together with the NIST value x_i.

9.3.2 Calculation of upper and lower control limits

A value of 0,05 was selected for α. From 6.4.2 one has

$\hat{\tau}^2 = 0{,}889 \times 10^{-4}$

$\hat{\gamma}_1 = 0{,}985\ 1$

$NK - 2 = 38$ and

$\zeta = 0{,}025$

These values lead to:

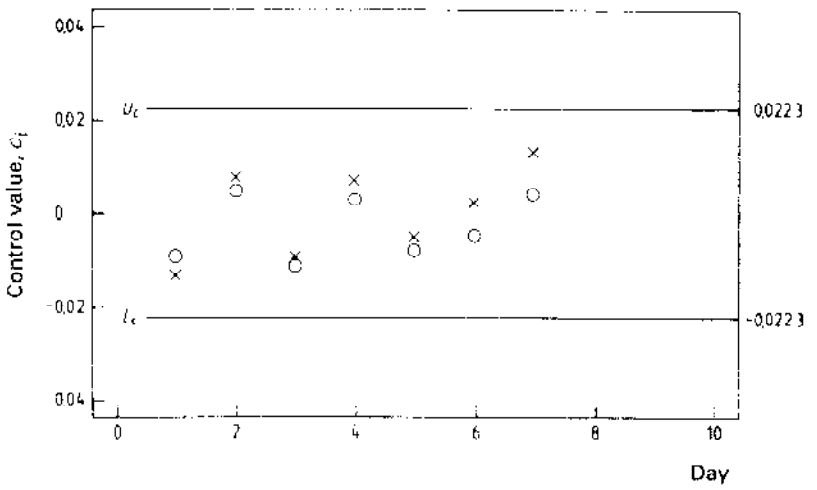
$U_c = 0{,}009\ 4 \times 2{,}334\ 2/0{,}985\ 1 = 0{,}022\ 3$

$L_c = -0{,}022\ 3$

These limits are plotted in figure 12.

Table 9 — Data collected for the control method

Day	NIST value x_i µm	Measured value y_i µm	Transformed value x_i^* µm	Control value c_i
1	2,99	3,154	2,951	−0,013
	10,77	10,760	10,673	−0,009
2	2,99	3,215	3,013	0,008
	10,77	10,909	10,823	0,005
3	2,99	3,165	2,962	−0,009
	10,77	10,740	10,652	−0,011
4	2,99	3,213	3,011	0,007
	10,77	10,892	10,806	0,003
5	2,99	3,179	2,976	−0,005
	10,77	10,772	10,685	−0,008
6	2,99	3,198	2,996	0,002
	10,77	10,807	10,720	−0,005
7	2,99	3,230	3,028	0,013
	10,77	10,897	10,811	0,004



Key: x = RM with low value, o = RM with high value

Figure 12 — Control chart to validate the calibration curve for line-spacings under the assumption of proportional residual standard deviation

9.3.3 Transformation and plot of the data

a) The values y_i are transformed to x_i^* using the calibration function and the control values

$$c_i = \frac{x_i^* - x_i}{x_i}$$

are obtained. A model with proportional residual standard deviation was adopted to derive the calibration function, the control method uses the normalized differences as control values rather than the regular differences $(d_i = x_i^* - x_i)$. The control values are listed in table 9.

b) The control values are plotted in the control chart (figure 12).

9.3.4 Decision about the state of the system

The system seems to be under control and the calibration function does not need to be updated as of day 7.

9.3.5 Estimation of the uncertainties of the transformed values during the validity period of the calibration function

Since only two RMs are used in the control chart, all control values c_i are included in the calculation of an estimate of the coefficient of variation of a transformed value. This estimate is equal to

$$\hat{\tau}_{cal} = \sqrt{\frac{\sum_{j=1}^{J}(c_{lj}^2 + c_{mj}^2)}{2J}} = 0{,}007\ 9$$

with $2J = 14$ degrees of freedom.

An approximate confidence interval for the unknown true value of a quantity estimated by the transformed value x_0^* with a confidence level of 0,95 is obtained as

$$x_0^* \pm \hat{\tau}_{cal}\, t_{(1-\alpha/2)}(2J)x_0^* = x_0^* \pm 0{,}007\ 9 \times 2{,}145 x_0^*$$

Annex A
(normative)

List of symbols and abbreviations

N	Number of reference materials
K or K_n	Number of replicate measurements per reference material
NK	Total number of measurements on all reference materials
x	Accepted value of a reference material
x_b	Accepted value of a blank
$\bar{x}$	Average of all accepted values
w	Inverse of an accepted value of a reference material ($1/x$)
$\bar{w}$	Average of all inverse accepted values
y	Measurement of a reference material
y_b	Measurement of a blank
$\bar{y}$	Average of all measurements
$\bar{y}_i$	Average of measurements of a specific reference material
z	Ratio of a measurement of a specific RM over the accepted value of the same RM (y/x)
β_0	Intercept of the calibration function under the assumption of constant residual standard deviation
β_1	Slope of the calibration function under the assumption of constant residual standard deviation
γ_0	Intercept of the calibration function under the assumption of proportional residual standard deviation
γ_1	Slope of the calibration function under the assumption of proportional residual standard deviation
ε	Deviation between a measurement and its expected value under the assumptions of linearity and of constant residual standard deviation
η	Deviation between a measurement and its expected value under the assumptions of linearity and of proportional residual standard deviation
e	Residual under the assumptions of linearity and of constant residual standard deviation
u	Weighted residual under the assumptions of linearity and of proportional residual standard deviation
σ^2	Constant residual variance (variance of ε)
σ_p^2	Variance associated with pure error under assumption of constant residual standard deviation
σ_l^2	Variance associated with lack of fit under assumption of constant residual standard deviation
τ^2	Proportional residual variance (variance of η/x)
τ_p^2	Variance associated with pure error under assumption of proportional residual standard deviation
τ_l^2	Variance associated with lack of fit under assumption of proportional residual standard deviation
SSE	Sum of squared residuals e
WSSE	Sum of squared weighted residuals u
SST, WSST	Total sum of squared deviations under, respectively, the assumption of constant or proportional residual standard deviation
SSP, WSSP	Sum of squared deviations due to pure error under, respectively, the assump-

	tion of constant or proportional residual standard deviation
SSR, WSSR	Sum of squared deviations explained by the calibration function under the assumption, respectively, of constant or proportional residual standard deviation
α	Significance level
$1 - \alpha$	Confidence level
$F_{(1-\alpha)}(n_1; n_2)$	$(1 - \alpha)$-quantile of the F-distribution with n_1 and n_2 degrees of freedom
$t_{(1-\zeta)}(n_1)$	$(1 - \zeta)$-quantile of the t-distribution with n_1 degrees of freedom
U_d	Upper control limit under assumption of constant residual standard deviation
L_d	Lower control limit under assumption of constant residual standard deviation
U_c	Upper control limit under assumption of proportional residual standard deviation
L_c	Lower control limit under assumption of proportional residual standard deviation
d	Control value under assumption of constant residual standard deviation
c	Control value under assumption of proportional residual standard deviation

Annex B
(normative)

Basic method when the number of replicates is not constant

When the number of replicates for each RM, K_n, is not constant, the calibration function can still be estimated by using adjusted formulae in 6.2.2, 6.4.2 and 6.5.

B.1 Estimates of β_0, β_1 and σ^2 are calculated as follows:

$$\hat{\beta}_1 = \frac{\sum_{n=1}^{N}(x_n - \bar{x})\sum_{k=1}^{K_n}(y_{nk} - \bar{y})}{\sum_{n=1}^{N} K_n(x_n - \bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$$

$$\hat{\sigma}^2 = \frac{\text{SSE}}{(NK - 2)}$$

where

$$NK = \sum_{n=1}^{N} K_n$$

$$\bar{x} = \frac{1}{NK}\sum_{n=1}^{N} K_n x_n$$

$$\bar{y} = \frac{1}{NK}\sum_{n=1}^{N}\sum_{k=1}^{K_n} y_{nk}$$

$$\hat{y}_n = \hat{\beta}_0 + \hat{\beta}_1 x_n$$

$$e_{nk} = y_{nk} - \hat{y}_n$$

$$\text{SSE} = \sum_{n=1}^{N}\sum_{k=1}^{K_n}(e_{nk})^2$$

B.2 Estimates of γ_0, γ_1 and τ^2 are calculated as follows:

$$\hat{\gamma}_0 = \frac{\sum_{n=1}^{N}(w_n - \bar{w})\sum_{k=1}^{K_n}(z_{nk} - \bar{z})}{\sum_{n=1}^{N} K_n(w_n - \bar{w})^2}$$

$$\hat{\gamma}_1 = \bar{z} - \hat{\gamma}_0\bar{w}$$

$$\hat{\tau}^2 = \frac{\text{WSSE}}{(NK - 2)}$$

where

$$NK = \sum_{n=1}^{N} K_n$$

$$z_{nk} = \frac{y_{nk}}{x_n}$$

$$w_n = \frac{1}{x_n}$$

$$\bar{w} = \frac{1}{NK}\sum_{n=1}^{N} K_n w_n$$

$$\bar{z} = \frac{1}{NK}\sum_{n=1}^{N}\sum_{k=1}^{K_n} z_{nk}$$

$$\hat{z}_n = \hat{\gamma}_1 + \hat{\gamma}_0 w_n$$

$$u_{nk} = z_{nk} - \hat{z}_n$$

$$\text{WSSE} = \sum_{n=1}^{N}\sum_{k=1}^{K_n}(u_{nk})^2$$

B.3 The lack of fit is evaluated as follows. Tables 1 and 2 still apply where

$$y_{n\cdot} = \frac{1}{K_n}\sum_{k=1}^{K_n} y_{nk}$$

$$\text{SST} = \sum_{n=1}^{N}\sum_{k=1}^{K_n}(y_{nk} - \bar{y})^2$$

$$\text{SSP} = \sum_{n=1}^{N}\sum_{k=1}^{K_n}(y_{nk} - y_{n\cdot})^2$$

SSE is as defined in B.1

and

$$z_{n\cdot} = \frac{1}{K_n} \sum_{k=1}^{K_n} z_{nk}$$

$$\text{WSST} = \sum_{n=1}^{N} \sum_{k=1}^{K_n} (z_{nk} - \bar{z})^2$$

$$\text{WSSP} = \sum_{n=1}^{N} \sum_{k=1}^{K_n} (z_{nk} - z_{n\cdot})^2$$

WSSE is as defined in B.2.

Annex C
(informative)

Bibliography

[1] MANDEL, J. Fitting straight lines when both variables are in error. *Journal of Quality Technology*, **16** (No. 1), 1984, pp. 1-14.

[2] GRAYBILL, F. *Theory and Applications of the Linear Model*. Duxbury Press, North Scituate, MA, 1976.

[3] DRAPER, N. and SMITH, H. *Applied Regression Analysis*. 2nd edn., Wiley, New York, 1981.

[4] CARROLL, R., SPIEGELMAN, C. and SACKS, J. A quick and easy multiple-use calibration curve procedure. *Technometrics*, **30**, 1988, pp. 137-141.

[5] MEE, R., EBERHARDT, K. and REEVE, C. Calibration and simultaneous tolerance intervals for regression. *Technometrics*, **23**, 1991, pp. 211-219.

[6] ISO 7870:1993, *Control charts — General guide and introduction*.

[7] ISO 8258:1991, *Shewhart control charts*.

[8] CROARKIN, C. and VARNER, R. Measurement assurance for dimensional measurements on integrated-circuits photomasks. *NBS Technical Note 1164*, 1982.

ICS 03.120.30

Descriptors: statistical analysis, measurement, measuring instruments, calibration, reference materials.

Price based on 29 pages