1
Phase I: Establishing a New Analytical Procedure

1.1
Introduction

1.1.1
Objectives of Phase I

Before a new analytical procedure, especially one requiring calibration, can be used for routine analysis, the individual steps need not only be determined but, where necessary, they must also be optimized, and the entire fundamental analytical procedure must be verified for its performance. The performance characteristics obtained in this way are documented and/or published with the description of the analysis and they form the basis for later quality assurance in routine analysis.

1.1.2
When Are Characteristic Data Obtained?

The statistical methods described below find their application primarily during the establishment of a new analytical process. In addition, they are suitable for providing the analyst with information about recent analytical quality achieved during the testing and training phase of an analytical process that is in need of calibration (see also Section 2.3).

The process data (linear calibration function including precision measures) should be determined anew with each new calibration of the analytical process, for example after:

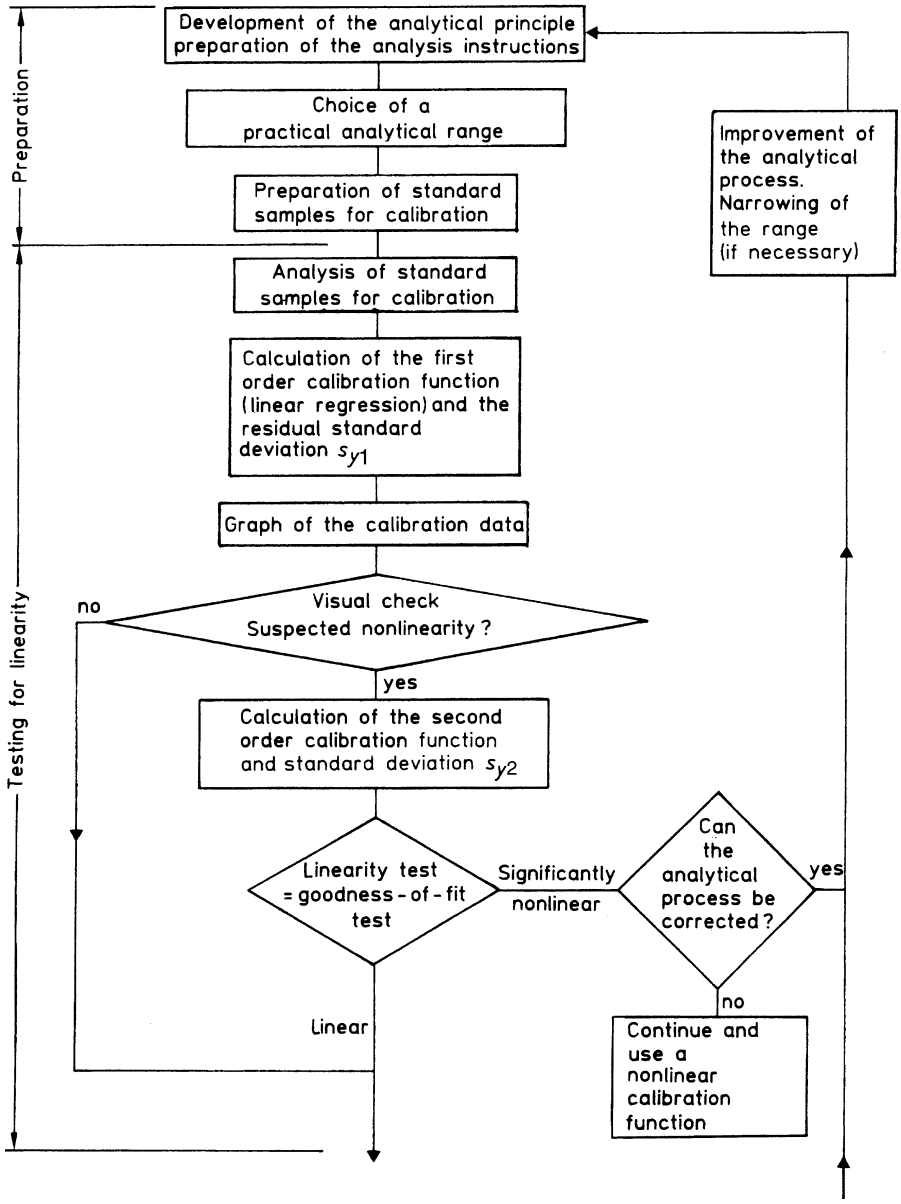
- changing reagents (new batches),
- technical intervention in analytical equipment (after technical modifications, maintenance and repair, for example bulb changes in photometry),
- changes in staff,
- etc.

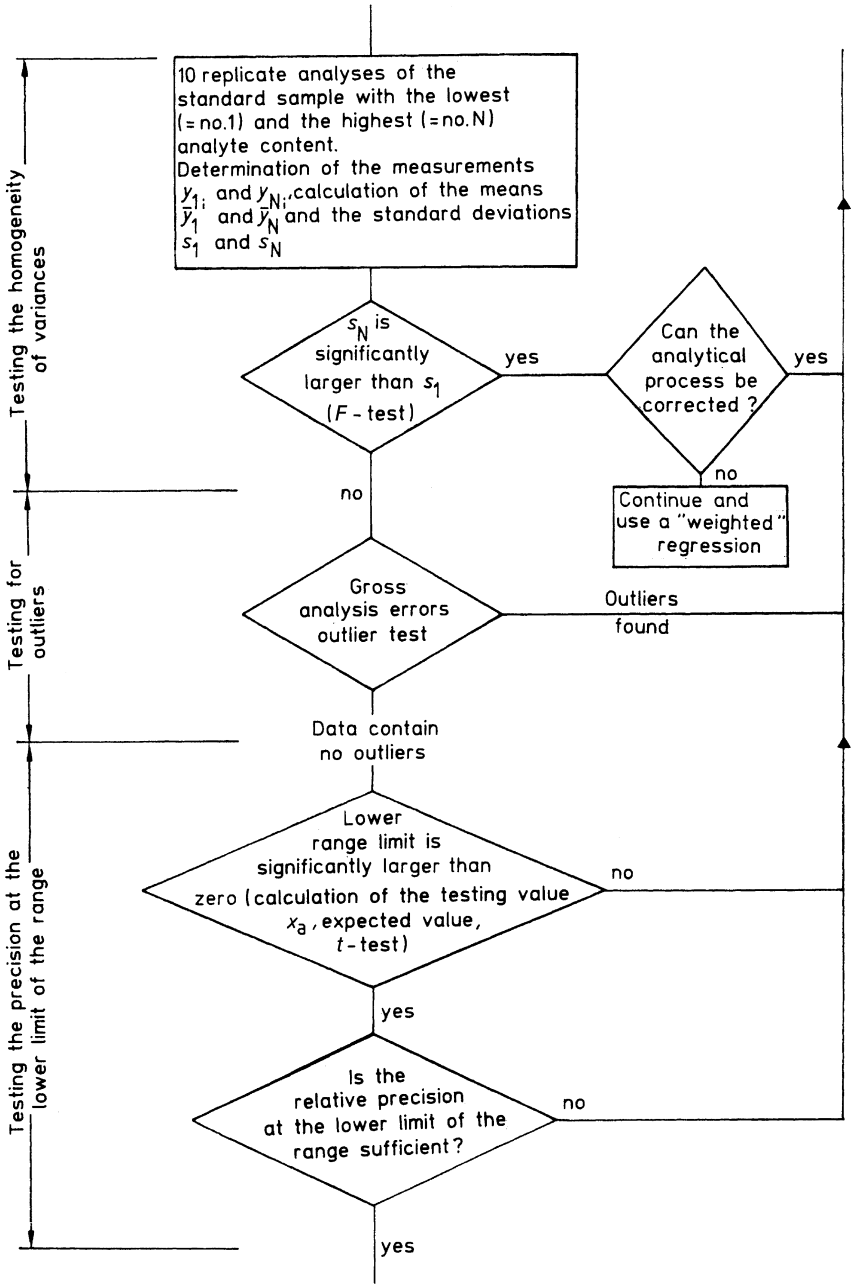
Quality Assurance in Analytical Chemistry: Applications in Environmental, Food, and Materials Analysis, Biotechnology, and Medical Engineering, Second Edition. W. Funk, V. Dammann, G. Donnevert

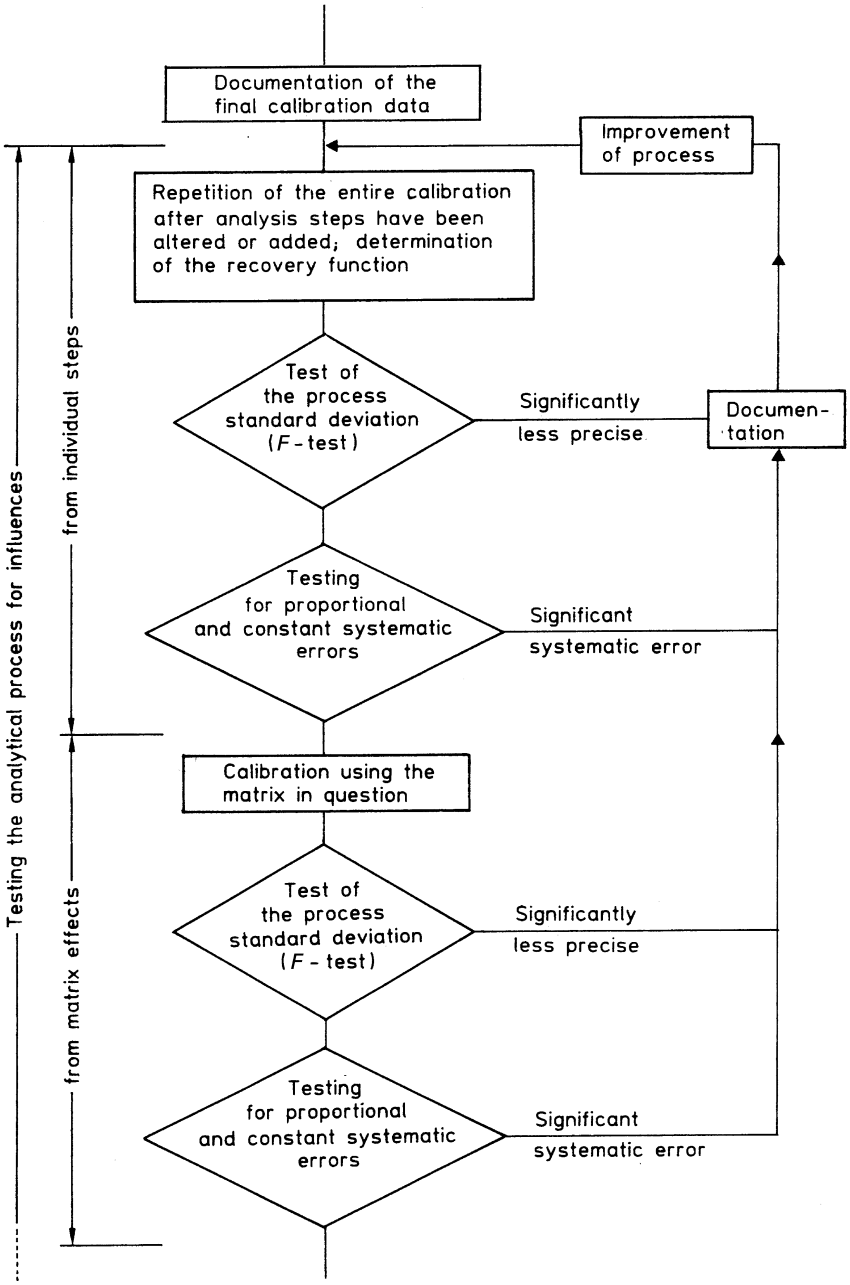
ISBN: 978-3-527-31114-9

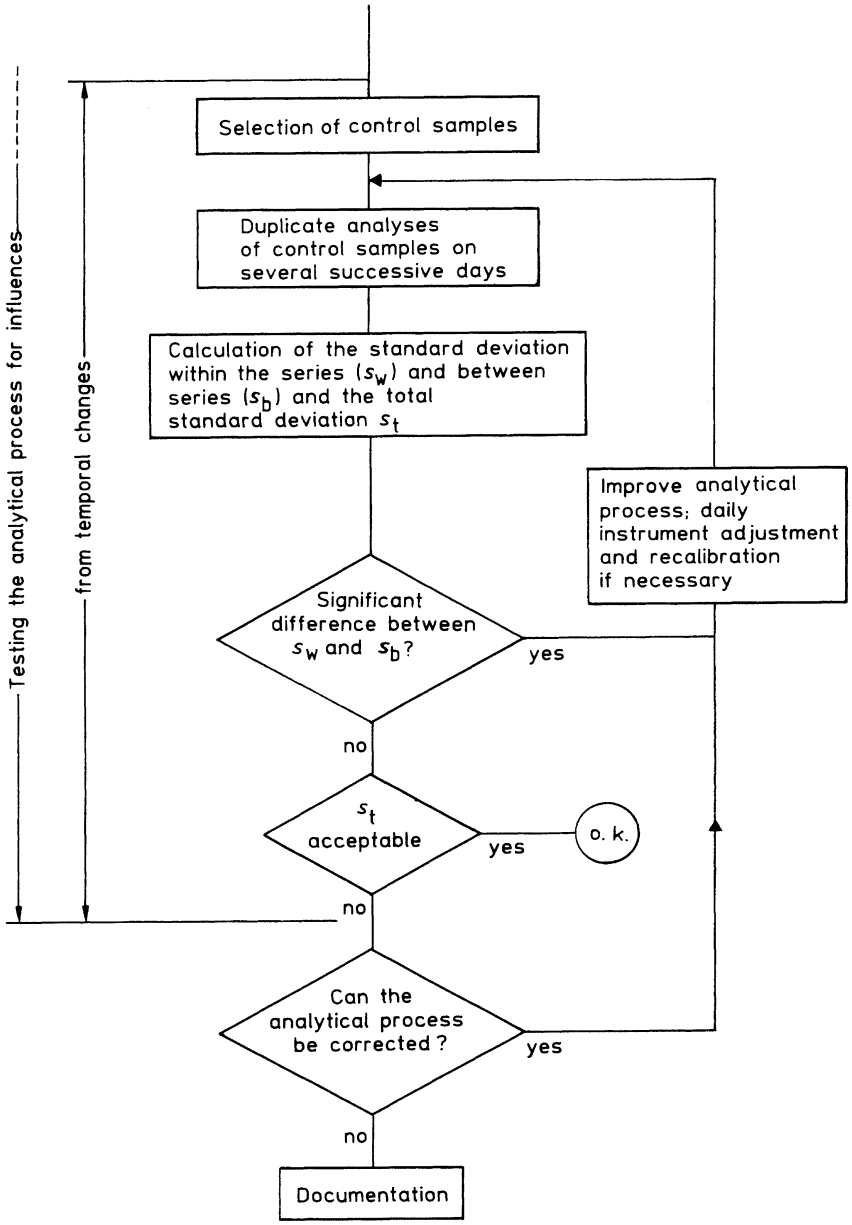
1.1.3 The Progression of Phase I

Scheme 1-1 Procedure of Phase I: flowchart.




Testing the homogeneity of variances
Testing for outliers
Testing the precision at the lower limit of the range
10 replicate analyses of the standard sample with the lowest (= no.1) and the highest (= no.N) analyte content.
Determination of the measurements y_{1i} and y_{Ni}, calculation of the means $\bar{y}_1$ and $\bar{y}_N$ and the standard deviations s_1 and s_N
s_N is significantly larger than s_1 (F-test)
yes
Can the analytical process be corrected?
yes
no
Continue and use a "weighted" regression
no
Gross analysis errors outlier test
Outliers found
Data contain no outliers
Lower range limit is significantly larger than zero (calculation of the testing value x_a, expected value, t-test)
no
yes
Is the relative precision at the lower limit of the range sufficient?
no
yes


Documentation of the final calibration data
Improvement of process
Repetition of the entire calibration after analysis steps have been altered or added; determination of the recovery function
Test of the process standard deviation (F-test)
Significantly less precise
Documen-tation
Testing for proportional and constant systematic errors
Significant systematic error
Calibration using the matrix in question
Test of the process standard deviation (F-test)
Significantly less precise
Testing for proportional and constant systematic errors
Significant systematic error
Testing the analytical process for influences
from individual steps
from matrix effects


Testing the analytical process for influences
from temporal changes
Selection of control samples
Duplicate analyses of control samples on several successive days
Calculation of the standard deviation within the series (s_w) and between series (s_b) and the total standard deviation s_t
Improve analytical process; daily instrument adjustment and recalibration if necessary
Significant difference between s_w and s_b?
yes
no
s_t acceptable
yes
o. k.
no
Can the analytical process be corrected?
yes
no
Documentation

Phase I consists primarily of the following five main steps (see Scheme 1-1):

- preparation of the calibration experiment,
- testing for
  - linearity of the calibration function
  - precision (variance homogeneity, absence of outliers, securing the lower bounds of the range)
- final determination of the calibration data,
- testing for the influence of the individual process steps or for matrix effects on the calibration data,
- testing for the influence of time on the analytical process.

Of special importance is the carefully conducted correction of the analytical process during the testing phase the moment the test results of a part of the system indicate an unacceptable analytical quality. If all possibilities for improving quality within the framework of the analytical process are exhausted and have not led to an improvement, in most cases a restriction of the range (reduction of the highest substance concentration chosen) can lead to an acceptable precision and to linearity of the calibration function as well as to the required variance homogeneity.

1.1.4 Results of Phase I; Statistical Data

Phase I of analytical quality assurance lays the foundation for later routine analysis by providing quality data. Above all, this includes:

- the range tested,
- the coefficients of the calibration function:
  - in the case of a first-order calibration function ($y = a + bx$): axis intercept a and slope b (characteristic of the sensitivity of the analytical procedure),
  - in the case of a second-order calibration function ($y = a + bx + cx^2$): axis intercept a, coefficient b of the linear term, as well as the coefficient c of the quadratic term; the sensitivity E of the analytical process determined from the function,
- the standard deviation of the procedure, s_{xo}, as an absolute measure of precision for the calibration, and
- the process *variation coefficient*, V_{xo}, as a *relative* measure of precision.

In addition, the general evaluation of the analytical process also documents the following:

- *decision limit*, x_{DL}, as the substance content that produces a measurement larger than the blank value with a probability of error α (e.g., $\alpha = 5\%$),
- *minimum detectable value*, x_{MDV}, as the substance content that is larger than the blank value with a probability of error β (e.g., $\beta = 5\%$),
- *limit of quantification*, x_{LQ}, as the substance content that can be determined with a maximum allowable relative result uncertainty,

- *auxiliary test value*, x_a, for the validation of a suitable range,
- recognized constant and/or proportional systematic deviations with certain matrices, as well as an indication of the time dependency of the analytical results, the time required, and other special remarks.

1.2 Calibration of the Fundamental Analytical Procedure (Fundamental Calibration)

For analytical processes in need of or capable of being calibrated, the application of physical measurement principles does not lead directly to an analytical result; the observations only represent the result of a physical measurement, which must be converted into the analytical result using data obtained empirically using a calibration experiment [101].

The use of an *analysis function*

Analytical result = function of the observation

or in mathematical notation:

$$\hat{x} = f(\hat{y})$$

(with $\hat{y}$ as the observation and $\hat{x}$ as the substance content/analytical result)

is based on the application of the *calibration function* obtained from the calibration experiment

observation = function of substance content

or

$$y = f(x)$$

(with x = substance content of the standard sample and y = accompanying observation)

and the precision data therefrom for the determination of an unknown substance content in a sample. After solving for x, the *calibration function* then becomes the *analytical function*, which, after inserting the observation of the tested sample, $\hat{y}$, gives the analytical result, $\hat{x}$ [101].

A *fundamental calibration* represents the calibration of the fundamental analytical procedure, or, in other words, no sample preparation steps such as extraction or work-up are performed, only standards in pure solvents (e. g., distilled water) are analyzed.

1.2.1 Establishment of an Analytical Range

Every calibration begins with the choice of a preliminary range as determined by:

a) The objective of the application as related to practical conditions. This should cover a large range of applications. In addition, the mean of the range should be roughly equal to the sample concentration most often expected, as far as this is possible in individual cases.

b) The technical possibilities.
   b1) The measured values at the lower end of the range must be significantly different from the process blank values. A lower range limit is only useful when it is at least the same or larger than the *minimum detectable value* (see Section 1.3.2) of the procedure. In addition, dilution and concentration steps must be easily and flawlessly realized.
   b2) The required analytical precision must be achievable throughout the entire range (see also "limit of quantification", Section 1.3.3). Since the imprecision of an analysis increases absolutely with increasing substance content, the range in question must not be chosen too large. If another range is necessary for routine analysis, then this should be divided into overlapping segments.

The applicability of the simple linear regression equation also requires that:

   b3) The analytical precision must be constant over the entire range [151] (homogeneity of variances). Inhomogeneity of variances that is ignored can result in a large increase in the measurement uncertainty of analytical results obtained using the calibration function (see Section 1.2.4.2.1).
   b4) There must be a linear relationship between substance content and measured value (linearity of the calibration function; see Section 1.2.4.1).

In the case of inhomogeneity of variances or nonlinearity, the chosen range must be reduced so as to fulfill these conditions, or more complicated calibration methods must be chosen, for example higher order regression functions (see Section 1.2.3.2) [72, 79] or weighted regression equations [16, 159].

1.2.2 Preparation of Standard Samples

Requirements of a standard sample are:

- purity, either lack of a matrix or a defined matrix,
- homogeneity,
- representativeness: the substance to be analyzed must be present in the standard sample in ways comparable to those expected in later analysis samples; in other words, compounds must be
  – chemically similar,
  – have the same oxidation state, etc.

- stability, ability to be preserved,
- storage: samples must not be influenced by containers or outside conditions.

Preparation of a standard sample:

- During preparation of a standard sample, the precision of the balance and volumetric equipment must be taken into account. Of the two, weighing is the more exact form of sample preparation and is therefore preferable to volumetric measurement [107]. To ensure precision, these instruments must be regularly tested and calibrated.
- There should be no successive dilutions since this entails the risk of error propagation.

After establishing the preliminary range, $N = 5...10$ standard samples are prepared so that their concentrations are distributed as equidistantly as possible over the entire chosen range.

1.2.3 Determination of the Calibration Function and Process Data

The preliminary first- and second-order calibration functions are calculated from the measured values obtained from these standard samples. The process data are necessary for further statistical tests.

Notes:

1. For reasons of clarity, the physical units of the measurement signals, concentrations, and statistical data below have been omitted during the calculations and have been added only at the final result. The plausibility of the physical units of a resulting value can, if necessary, be checked by means of a *dimensional analysis*.

   Example:
   slope of the linear calibration function:
   substance content x: mg/l
   observation: peak height y: mm
   slope of the calibration function:
   [peak height per substance content] = $\frac{\text{mm}}{\text{mg/l}}$

   Verification by calculation formulae:

$$\text{slope} = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\sum \left\{ \frac{\text{mg}}{\text{l}} \cdot \text{mm} \right\}}{\left\{ \left(\frac{\text{mg}}{\text{l}} \right)^2 \right\}}$$

$$= \frac{\left\{\frac{\text{mg}}{\text{l}} \cdot \text{mm}\right\}}{\left\{\frac{\text{mg}}{\text{l}} \cdot \frac{\text{mg}}{\text{l}}\right\}}$$

$$= \left\{\frac{\text{mm}}{\text{mg/l}}\right\} \qquad \text{(q.e.d.)}$$

2. Generally, the running index for summation signs ($\sum$) has been omitted. In these cases, the "i" forms the index and it runs from 1 to N.

   Example:

   $\sum (x_i - \bar{x})^2$ is equivalent to $\sum_{i=1}^{N} (x_i - \bar{x})^2$

1.2.3.1 Process Data for the Linear Calibration Function

Regression analysis provides the calibration function (see Figure 1-1) with the characteristic data.

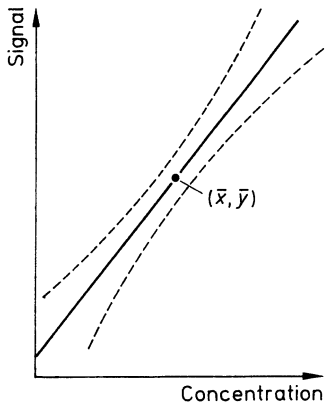


Fig. 1-1 Linear calibration function ($y = a + bx$) with prognosis interval.

Slope (measure of sensitivity):

$$b = \frac{\sum \left[(x_i - \bar{x}) \cdot (y_i - \bar{y})\right]}{\sum (x_i - \bar{x})^2} \tag{1}$$

Axis intercept:

$$a = \bar{y} - b\bar{x} \qquad \text{with} \qquad \bar{x} = \frac{1}{N} \cdot \sum x_i$$

$$\bar{y} = \frac{1}{N} \cdot \sum y_i \tag{2}$$

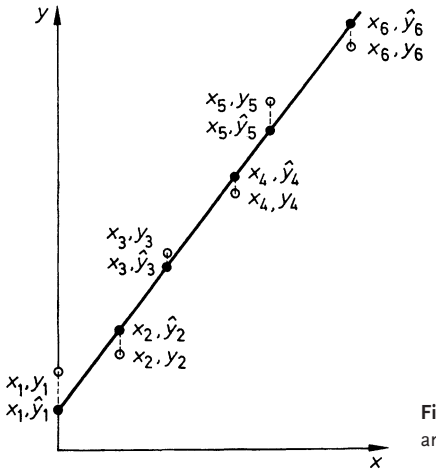


Fig. 1-2 Distribution of measured values (○) around the regression line.

Residual standard deviation (scatter of measured values around the regression line; see Figure 1-2):

$$s_y = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{N-2}} \qquad \text{with} \qquad \hat{y}_i = a + bx_i \tag{3}$$

Process standard deviation:

$$s_{xo} = \frac{s_y}{b} \tag{4}$$

Process variation coefficient = relative process standard deviation:

$$V_{xo} = \frac{s_{xo}}{\bar{x}} \cdot 100\,(\%) \tag{5}$$

1.2.3.2 Process Data for the Second-Order Calibration Function

Here, regression analysis yields the second-order calibration function (see Figure 1-3) with its characteristic process data [72].

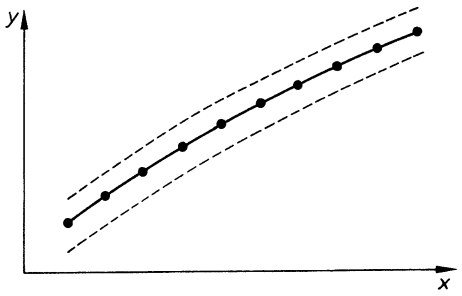


Fig. 1-3 Second-order calibration function ($y = a + bx + cx^2$).

Function coefficients a, b, c:

$$a = \frac{1}{N}\left(\sum y_i - b \cdot \sum x_i - c \cdot \sum x_i^2\right) \tag{6}$$

$$b = \frac{Q_{xy} - c \cdot Q_{x^3}}{Q_{xx}} \tag{7}$$

$$c = \frac{Q_{xy} \cdot Q_{x^3} - Q_{x^2y} \cdot Q_{xx}}{\left(Q_{x^3}\right)^2 - Q_{xx} \cdot Q_{x^4}} \tag{8}$$

$$Q_{xx} = \sum x_i^2 - \left(\left(\sum x_i\right)^2 / N\right) \tag{9}$$

$$Q_{xy} = \sum (x_i \cdot y_i) - \left(\left(\sum x_i\right) \cdot \left(\sum y_i\right) / N\right) \tag{10}$$

$$Q_{x^3} = \sum x_i^3 - \left(\left(\sum x_i\right) \cdot \left(\sum x_i^2\right) / N\right) \tag{11}$$

$$Q_{x^4} = \sum x_i^4 - \left(\left(\sum x_i^2\right)^2\right) / N \tag{12}$$

$$Q_{x^2y} = \sum (x_i^2 \cdot y_i) - \left(\left(\sum y_i\right) \cdot \left(\sum x_i^2\right) / N\right) \tag{13}$$

Residual standard deviation:

$$s_y = \sqrt{\frac{\sum \left(y_i - \hat{y}_i\right)^2}{N - 3}} \quad \text{where} \quad \hat{y}_i = a + bx_i + cx_i^2 \tag{14}$$

Sensitivity E:

The measure of sensitivity results from the change in the measured value caused by a change in the concentration values. If the calibration function for an analytical procedure is linear, then the sensitivity is constant over the entire range and is equivalent to the regression coefficient *b* [72]. In the case of a curved calibration function, the sensitivity is still dependent on the given concentration value and is equivalent to the first derivative of the calibration function:

$$E(x) = b + 2\,c \cdot x \tag{15}$$

As a characteristic process quantity, it is recommended that the sensitivity is expressed in the middle of the working range:

$$E(\bar{x}) = b + 2\,c \cdot \bar{x} \tag{16}$$

From this, one can derive the process standard deviation and the relative process standard deviation:

Process standard deviation: *Relative process standard deviation:*

$$s_{xo} = \frac{s_y}{E(\bar{x})} \qquad V_{xo} = \frac{s_{xo}}{\bar{x}} \cdot 100\,(\%) \tag{17}$$

1.2.3.3 Calculating Analytical Results with the Aid of the Calibration Function

1.2.3.3.1 Results Obtained Using the Linear Calibration Function

Analytical results and confidence intervals can be calculated through the use of the linear calibration function as follows:

$$\text{Result} = \hat{x} \pm CI(\hat{x})$$

$$= \frac{\hat{\bar{y}} - a}{b} \pm s_{xo} \cdot t(P = 95\%, f = N_c - 2) \cdot \sqrt{\frac{1}{N_c} + \frac{1}{N_a} + \frac{\left(\hat{\bar{y}} - \bar{y}\right)^2}{b^2 \cdot Q_{xx}}} \tag{18}$$

with $\hat{\bar{y}}$ = mean value from N_a multiple analyses (N_a can also be equal to 1) and the data from the linear calibration:

a = axis intercept
b = slope
s_{xo} = standard deviation of the method
N_c = number of calibration standards
Q_{xx} = the sum of $(x_i - \bar{x})^2$

Example:

1. Calibration data of nitrite determination

i	x_i in mg/l	y_i in abs.
1	0.05	0.140
2	0.10	0.281
3	0.15	0.405
4	0.20	0.535
5	0.25	0.662
6	0.30	0.789
7	0.35	0.916
8	0.40	1.058
9	0.45	1.173
10	0.50	1.303

$\bar{x}$ = 0.275 mg/l
$\bar{y}$ = 0.726 abs.
a = 0.018 abs.
b = 2.575 abs./(mg/l)
s_{xo} = 0.0020 mg/l
N_c = 10
Q_{xx} = 0.20625 $(\text{mg/l})^2$

2. Analytical result

The measurement obtained from an unknown sample is $\hat{y} = 0.641$ abs.

According to Eq. (18), the analytical result can be calculated with a 95 % prognosis interval [$t(f = 8, P = 95\,\%) = 2.31$]:

$$\hat{x}_{1,2} = \frac{0.641 - 0.018}{2.575} \pm 0.0020 \cdot 2.31 \sqrt{\frac{1}{10} + \frac{1}{1} + \frac{(0.641 - 0.7262)^2}{2.575^2 \cdot 0.20625}}$$

$$\hat{x}_{1,2} = 0.24 \pm 0.005 \text{ mg/l}$$

1.2.3.3.2 Results Obtained Using the Linear Second-Order Calibration Function

Analytical results and confidence intervals can be calculated through the use of the linear second-order calibration function as follows:

Result for negative curvature: $$\hat{x} = -\frac{b}{2c} - \sqrt{\left(\frac{b}{2c}\right)^2 - \frac{a - \hat{y}}{c}}$$

Result for positive curvature: $$\hat{x} = -\frac{b}{2c} + \sqrt{\left(\frac{b}{2c}\right)^2 - \frac{a - \hat{y}}{c}}$$

with

$$CI(\hat{x}) = \frac{s_y \cdot t}{(b + 2c\hat{x})} \cdot \sqrt{\frac{1}{N_c} + \frac{1}{N_a} + \frac{1}{Q_{x^4} \cdot Q_{xx} - \left(Q_{x^3}\right)^2}} \cdot$$

$$\cdot \left\{ (\hat{x} - \bar{x})^2 Q_{x^4} + \left(\hat{x}^2 - \frac{\sum x_i^2}{N_c}\right)^2 Q_{xx} - 2 \cdot (\hat{x} - \bar{x}) \cdot \left(\hat{x}^2 - \frac{\sum x_i^2}{N_c}\right) \cdot Q_{x^3} \right\} \qquad (19)$$

Example:

1. *Calibration data*

i	x_i in mg/l	y_i in abs.
1	12	0.083
2	18	0.123
3	24	0.164
4	30	0.203
5	36	0.240
6	42	0.273
7	48	0.303
8	54	0.334
9	60	0.364
10	66	0.393

$\bar{x}$ = 39 mg/l
a = 0.00562 abs.
b = 0.00767 abs./(mg/l)
c = −0.000025 abs./(mg/l)^2
s_y = 0.00148 abs.
s_{xo} = 0.258617 mg/l
N_c = 10
Q_{xx} = 2970 (mg/l)^2
Q_{x^3} = 231660 (mg/l)^3
Q_{x^4} = 18753770 (mg/l)^4

2. Analytical result

The measured value of an unknown sample is $\bar{y} = 0.223$ abs. According to Eq. (19), the analytical result can be calculated as

$$\hat{x} = 33.46 \text{ mg/l}$$

Accordingly, the 95% prognosis interval with $t(f = 7,\ P = 95\%) = 2.36$ will be calculated as:

$$CI(\hat{x}) = \pm 0.643 \text{ mg/l}$$

1.2.4 Verification of the Fundamental Calibration

1.2.4.1 Verification of Linearity

If possible, one should attempt to work with a first-order calibration function. Second-order calibration functions should only be used in justified exceptions.

1.2.4.1.1 Visual Linearity Test

The type of calibration function is most simply determined through graphical representation of the calibration data, including a calibration line and a subjective evaluation. If the measured data were highly precise (see Figure 1-4) and the result is obviously nonlinear, one may do without a special statistical test of linearity. In doubtful cases, however, the linearity should be verified mathematically.

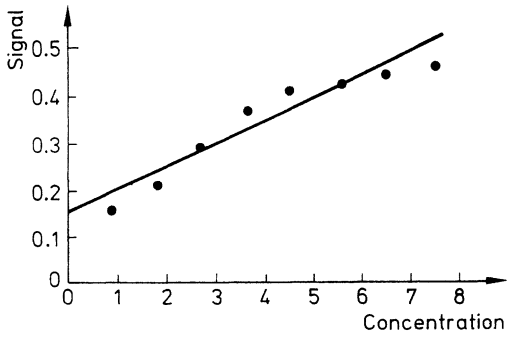


Fig. 1-4 Visual linearity test.

1.2.4.1.2 Mandel's Fitting Test

Mandel's fitting test is recommended for mathematical verification of linearity [47, 151, 190].

The test is based on the assumption that relatively large deviations of measured values from a straight line are caused by nonlinearity and may be reduced through the selection of a "better" regression model, in this case a second-order function.

For this, the first-order calibration function $y = a + bx$ and the second-order calibration function $y = a + bx + cx^2$, including their respective residual standard deviations s_y (see Sections 1.2.3.1 and 1.2.3.2), are used.

The difference of the variances DS^2 is calculated using the residual standard deviation s_{y_1} (from the first-order calibration function) and s_{y_2} (from the second-order calibration function):

$$DS^2 = (N-2)\, s_{y_1}^2 - (N-3)\, s_{y_2}^2, \text{ with the degree of freedom } f = 1 \tag{20}$$

The test value, *TV*, is calculated for the *F*-test:

$$TV = \frac{DS^2}{s_{y_2}^2} \tag{21}$$

and is compared with the value obtained from the table $F(f_1 = 1, f_2 = N-3, P = 99\%)$.

If $TV \leq F$, then the second-order calibration function will *not* provide a significantly better fit; the calibration function is linear.

If $TV > F$, then the individual steps of the analytical process should first be checked and improved upon if possible. If after that linearity is still not obtained (e. g., due to physicochemical laws), then a narrowing of the range should be attempted in order to maintain sufficient linearity. A last resort is the future evaluation of the measured data by way of a second-order calibration function (see Section 1.2.3.2). Residual analysis is an alternative to the Mandel fitting test for the verification of linearity.

1.2.4.1.3 **Residual Analysis**

Another possibility for testing whether the chosen functional approach of the calibration model adequately describes the measured results is residual analysis [113].

The residuals d_i are the vertical distances of the observations from the regression curve (see Figure 1-2)

$$d_i = \hat{y}_i - y_i \quad \text{for} \quad i = 1, \ldots, n$$

with y_i = observation and $\hat{y}_i$ = estimated value of y_i (from the regression function).

The residuals d_i are distributed normally [69, 113] if the chosen model approach is correct (see Figure 1-5 a). If the residuals show a trend (see Figure 1-5 b–d), then the underlying regression approach must be verified; for example, in case (d) of Figure 1-5, a second-order function must be calculated.

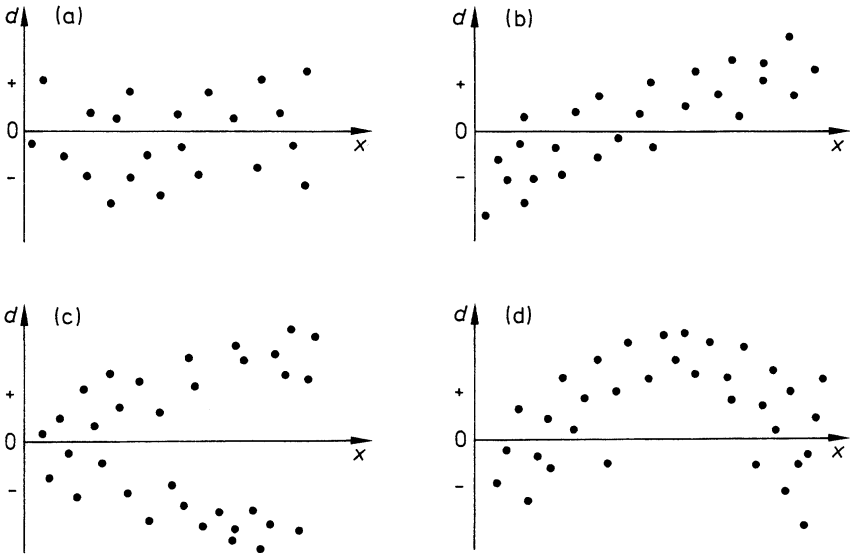


Fig. 1-5 Graphical representation of residuals dependent on the concentration *x*:
a) ideal course, i.e., model approach is correct;
b) linear trend, probably incorrect approach or calculation error;
c) increasing variances, i.e., inhomogeneity of variances;
d) nonlinear course, result of choosing an incorrect regression function.

1.2.4.2 **Verification of Precision**

1.2.4.2.1 **Homogeneity of Variances**

The described linear regression calculation assumes a constant (homogeneous) imprecision (variance of measured values) over the range.

Inhomogeneity leads not only to a higher imprecision (see Figure 1-6), but also to a higher inaccuracy through possible change in the linear slope.

In order to verify the homogeneity of variances, $n = 10$ standard samples of each of the lowest (x_1) and the highest (x_N) concentrations of the preliminary range are analyzed separately.

One obtains $2 \cdot n$ ($n = 10$) measurements (y_{ij}) from this series. The means, $\bar{y}_1$ and $\bar{y}_N$, and variances, s_1^2 and s_N^2, are calculated for both sets of data:

$$s_i^2 = \frac{\sum \left(y_{ij} - \bar{y}_i\right)^2}{n_i - 1} \quad \text{(for } i = 1 \text{ and } i = N, \text{ and with } j \text{ from 1 to 10 in each case)} \quad (22)$$

The variances of both series of measurements are checked for homogeneity using an *F*-test:

$$TV = \frac{s_N^2}{s_1^2} \quad (23)$$

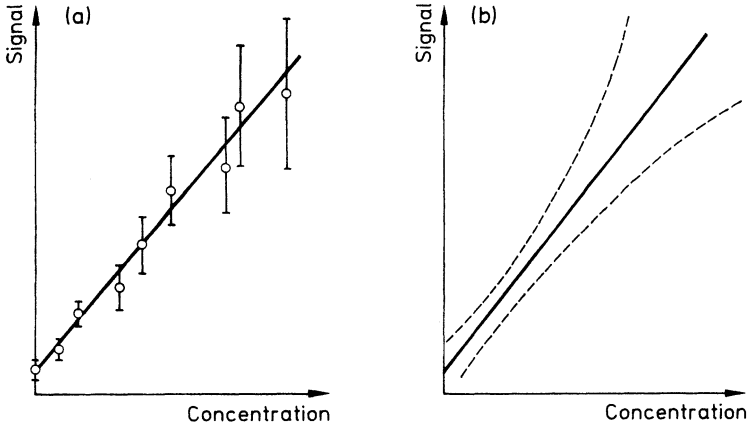


Fig. 1-6 Inhomogeneity of variances.
a) mean values and confidence intervals of the measured results (multiple analyses of the standards);
b) resulting prognosis bands.

If the *F*-test shows a significant difference between the variances, i.e., $TV > F$ ($f_1 = 9, f_2 = 9$, $P = 99\%$), there are three possible ways to proceed:

1. Choose a narrower range and repeat the verification for the homogeneity of variances (*recommended method*). A typical cause of the inhomogeneity of variances lies, for example, in a change in the display or amplification range of the measuring instrument within the range being used. A display change alone can affect the variance by around a factor of ten.
2. Use a weighted regression [159].
3. Multiple curve fitting [160].

1.2.4.2.2 **Outlier Test**

As a matter of principle, calibration data must be free from outliers. Suspected outlier values can be tested by means of various outlier tests [149]. In each case, the suitable regression model must be determined beforehand (see Section 1.2.4.1), since the correctness of the chosen regression approach is a prerequisite for the application of outlier tests.

Residual analysis (see Section 1.2.4.1.3) can also be used to determine outliers during calibration [115]. To this end, the calibration curve with the residual standard deviation must be calculated using all data pairs. Potential outliers can be pre-selected mathematically through determination of the residuals ($y_i - \hat{y}_i$; see Figure A5; Appendix A1.2.4) and their graphical representation. Every data pair with a noticeably large residual is a potential outlier. After elimination of the suspect outlier pair (x_A, y_A) from the collected data, a new calibration line is calculated with the residual distribution $s_{y_{A2}}$. Either an *F*-test or a *t*-test may be used for verification. Both methods will give identical results.

F-Test: The residual distributions $s_{y_{A1}}$ and $s_{y_{A2}}$ of the two lines are checked for a significant difference.

The test value is calculated

$$TV = \frac{(N_{A1} - 2)\, s_{y_{A1}}^2 - (N_{A2} - 2)\, s_{y_{A2}}^2}{s_{y_{A2}}^2} \tag{24}$$

and compared with the value obtained from the table F ($f_1 = 1$, $f_1 = N_{A2} - 2$, $P = 99\,\%$). If $TV < F$, then with a 1% error probability no outlier exists and the eliminated values can be reincluded in the collected data.

t-Test: For the *t*-test, the prediction interval of the second regression line is calculated (after eliminating the outlier) for the concentration x_A and it is checked to see if the suspected outlier value lies within this prediction interval. If this is the case, then the eliminated value must be reincluded in the collected data.

Calculation of the prediction interval:

$$\begin{aligned} CI(\hat{y}_A) &= \hat{y}_A \pm t \cdot s_{y_{A2}} \cdot \sqrt{1 + \frac{1}{N_{A2}} + \frac{(x_A - \bar{x})^2}{\sum x_i^2 - \frac{1}{N_{A2}} \left(\sum x_i\right)^2}} \\ &= a_2 + b_2 \cdot x_A \pm t \cdot s_{y_{A2}} \cdot \sqrt{1 + \frac{1}{N_{A2}} + \frac{(x_A - \bar{x})^2}{\sum x_i^2 - \frac{1}{N_{A2}} \left(\sum x_i\right)^2}} \end{aligned} \tag{25}$$

t = value of *t*-distribution obtained from the table ($P = 95\,\%$, $f = N_{A2} - 2$, double-sided)
N_{A2} = $N - 1$ (N = original number of calibration data pairs)
x_A = standard concentration of the eliminated outlier observation
$\bar{x}$ = mean of all x_i (without x_A)

Note: If an outlier is statistically proven by means of an *F*- or *t*-test, then it is absolutely necessary to seek and eliminate the source of error and then repeat the entire calibration.

1.2.4.2.3 Securing the Lower Range Limit

A calibration function is only usable for quantitative analysis when all analytical results subsequently calculated with it are significantly different from zero. Therefore, the lower range limits are tested to determine if they are significantly different from zero. The calculation formulae for the test value x_a (see Figure 1-7) correspond to those used to determine the minimum detectable value (see Section 1.3.2.2).

$$x_a = 2\, CI_x\, (\gamma = \gamma_a) \quad \text{with } t\ (f = N - 2,\ P = 95\,\%, \text{ single-sided}) \tag{26}$$

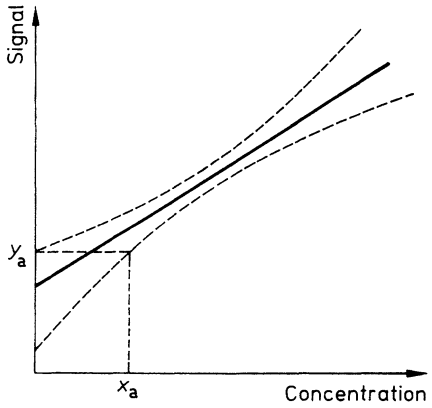


Fig. 1-7 Determination of x_a using the auxiliary value y_a.

$$x_a = 2 \cdot \frac{y_a - a}{b} \tag{27}$$

where

$$y_a = a + s_y \cdot t \cdot \sqrt{\frac{1}{N} + 1 + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}} \tag{28}$$

If $x_a < x_1$ (see Figure 1-8a), then the entire chosen range is statistically sound, i. e. the lower limit of the range x_1 is significantly different from the concentration zero (see Section 1.3.2, minimum detectable value).

If x_a lies above x_1 (see Figure 1-8b), then x_1 is not significantly different from the concentration zero and the range is therefore only sound for concentrations > x_a. In this case, quantitative analyses are only possible above this concentration value. A completely new calibration must be performed for this limited range. However, it makes more sense to examine and improve the analytical process, or the individual procedures therein.

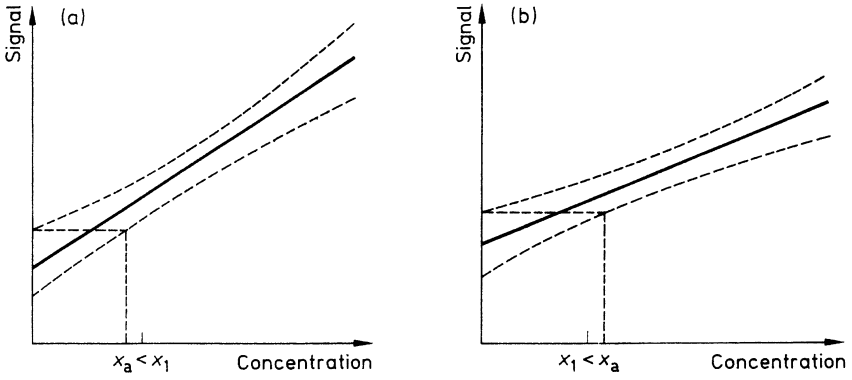


Fig. 1-8 Testing the lower range limit x_1. a) $x_a < x_1$; b) $x_a > x_1$.

1.2.4.2.4 Relative Analytical Imprecision

If a minimum precision $CI_{\text{rel,exp}}$ (relative, expected) is required for the analysis of the sought parameters, then the *relative analytical precision* at the lower end of the range (x_1) must be verified:

$$CI(x_1) = s_{xo} \cdot t \cdot \sqrt{\frac{1}{N} + 1 + \frac{(x_1 - \bar{x})^2}{\sum (x_i - \bar{x})^2}} \tag{29}$$

$$CI_{\text{rel}}(x_1) = \frac{CI(x_1)}{x_1} \cdot 100\,(\%) \tag{30}$$

If $CI_{\text{rel}}(x_1)$ is larger than $CI_{\text{rel,exp}}$, then either the analytical precision must be improved (optimization of individual procedure steps) or the lower range limit must be raised.

1.3 Analyses at Very Low Concentrations

In most analytical processes, the absolute precision of the measured data, and/or the analytical result, improves with decreasing concentration of the substance being analyzed. However, when working with extremely low concentrations (trace analysis) it is often a problem to significantly distinguish the signal of a sample from the signal of a blank or to obtain a quantitative result with sufficient precision.

If during an analysis an observation is obtained that lies negligibly above zero, then two contrasting interpretations are possible:

a) The sample does not contain the substance in question (analyte); the observation is only a result of the imprecision of the analytical procedure and is therefore included in the distribution range of the blank values.

b) The analyte is actually present in the sample; repeated analyses would in this case provide an analytical mean within the distribution range of which the first, and doubtful, observation would lie.

Both the probability of affiliation to the blank measurement values and also to an existing substance concentration can be statistically determined for each measured value at a known analytical precision (standard deviation). Decision-making criteria for or against the given probability of error are:

a) The probability that the sample really does not contain the analyte, even though a positive measurement has been obtained; this is known as α-error (false positive decision).

b) The probability that the tested sample really does contain the analyte, even though the analytical results are seen as zero (blank value); this is known as β-error (false negative decision), see Figure 1-9.

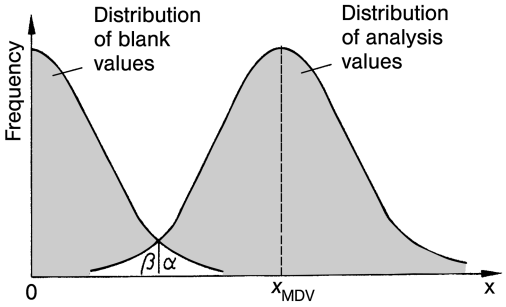


Fig. 1-9 Distribution of blank and analytical values with α- and β-errors.

In analytical practice, three characteristic values with different information content can be defined at a given precision:

- the decision limit x_{DL},
- the minimum detectable value x_{MDV},
- the limit of quantification x_{LQ}.

One must differentiate between two different courses of action for determining the decision limit and minimum detectable value depending on whether or not the analytical process is capable of being calibrated (see Figure 1-10).

In analytical processes that can be calibrated, the decision limit and the minimum detectable value will be derived from the calibration data. Therefore, a calibration can be carried out using a very low concentration that is close to the expected decision limit, minimum detectable value, and limit of quantification. This calibration range extends over at most one order of magnitude so that the variance homogeneity is given. It is usually situated well below the range for later routine analyses.

Certain analytical procedures, for example the determination of chemical oxygen demand (COD) in water analysis according to DIN 38409–41 [51] or the adsorption of organic halogens (AOX) according to DIN EN 1485 [59], cannot be calibrated because a suitable, defined (representative) standard substance does not exist.

In these cases, neither the decision limit nor the minimum detectable value can be calculated using calibration data. Instead, these are obtained by repeated analysis of blank samples.

As one can make predictions about the efficiency [88] of the future application of the analytical procedure using the characteristics "decision limit, minimum detectable value, and limit of quantification", the number of future multiple analyses per sample as N_a in the calculations must be taken into account: for single analyses $N_a = 1$, for triple analyses $N_a = 3$, and so on.

The determination and evaluation of blank values presumes:

- that the determination of blank values and measured values for the sample are independent of each other;
- that the blank value is obtained from the entire analytical process (i. e., including sample preparation and measurement);

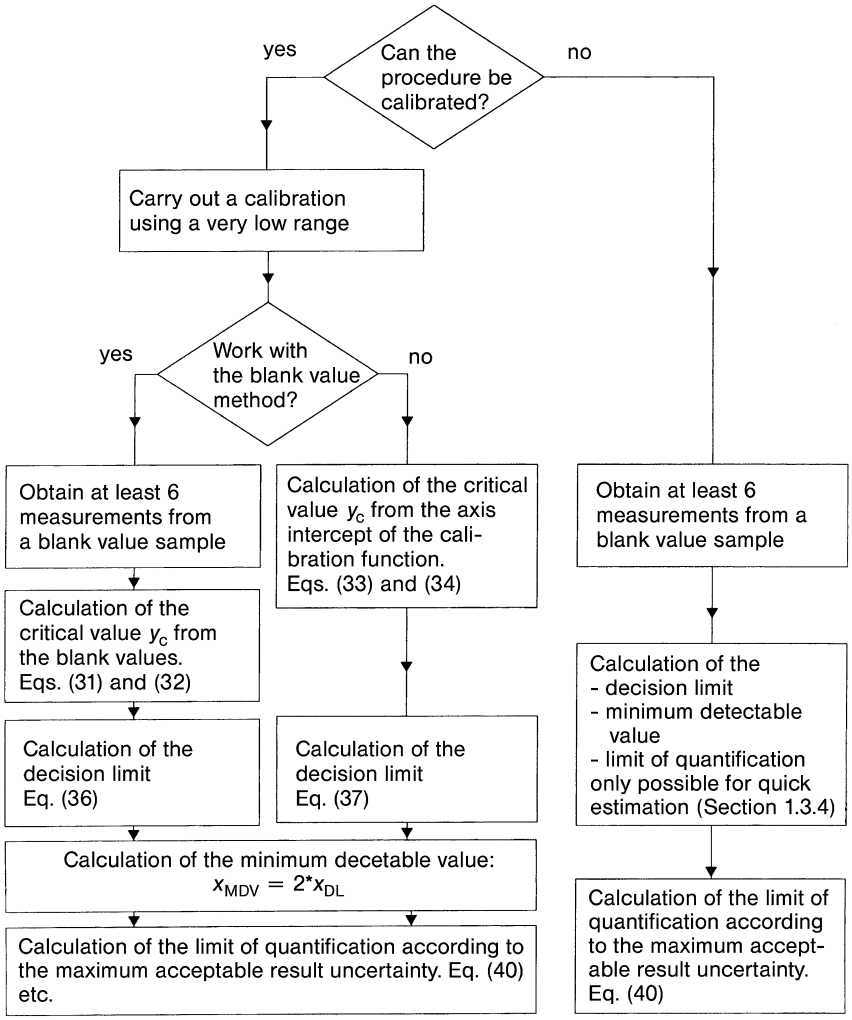


Fig. 1-10 Criteria for deciding between the blank value method and the calibration function method.

- that the measured values of the sample and the accompanying blank values are governed by a normal distribution [190];
- that the analytical result, calculated using the difference between sample values and blank values, also follows a normal distribution; and
- that the distribution of the blank values is not significantly different from the distribution of the analytical results at low concentrations.

1.3.1
Decision Limit [34, 120, 132]

The critical value y_c for measured values is defined as that observation (measured value) for which the β-error is exactly 50 % with an α-error of 5 % (see Figure 1-11).

Expressed in terms of error probability, this means that if the observation value y_c is used, the blank value has been exceeded ("substance proven"), and it is subject to an error probability of only α = 5 %. However, with repeated analysis of these samples approximately 50 % of all results lie below y_c (β = 50 %), in other words, the probability that the presence of substance *cannot* be proven is 50 %!

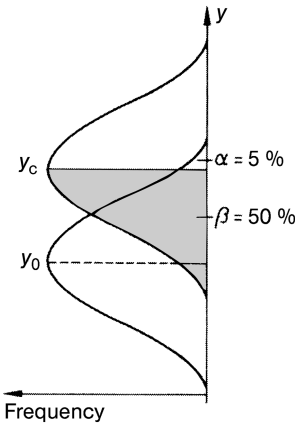


Fig. 1-11 Decision limit: α-error = 5 %, β-error = 50 %.

Practical Determination of the Decision Limit x_{DL}

First the critical value, y_c, is calculated. Inserting the critical value, y_c, into the calibration function and solving according to x yields the decision limit, x_{DL}.

$$x_{DL} = \frac{y_c - a}{b} \tag{31}$$

To derive y_c by blank values, the mean blank value, $\bar{y}_B$, and the prognosis interval for future blank values, $\Delta\bar{y}_B$, calculated by the standard deviation, s_B, are determined.

The mean and standard deviation should be based on at least N_B = 6 measured values obtained under repeat conditions.

$$\Delta\bar{y}_B = s_B \cdot t_{f,\alpha} \sqrt{\frac{1}{N_a} + \frac{1}{N_B}} \tag{32}$$

with $f = N_B - 1$ degrees of freedom gives

$$y_c = \bar{y}_B + \Delta y_B \tag{33}$$

When using the calibration line method, the critical value, y_c, is calculated as the sum of the axis intercept, a, and the predicted range of the axis intercept, Δa.

$$\Delta a = s_y \cdot t_{f,\alpha} \sqrt{\frac{1}{N_a} + \frac{1}{N_c} + \frac{\bar{x}^2}{Q_{xx}}} \tag{34}$$

with $f = N_c - 2$ degrees of freedom gives

$$y_c = a + \Delta a \tag{35}$$

After insertion of Eqs. (32) and (33) or (34) and (35) into Eq. (31), one obtains the decision limit x_{DL}:

a) Using the blank value method

$$x_{DL} = \frac{\bar{y}_B + s_B \cdot t_{f,\alpha} \sqrt{\frac{1}{N_a} + \frac{1}{N_B}} - \bar{y}_B}{b}$$

$$= \frac{s_B}{b} \cdot t_{f,\alpha} \sqrt{\frac{1}{N_a} + \frac{1}{N_B}} \tag{36}$$

with $f = N_B - 1$ degrees of freedom

b) Using the calibration line method

$$x_{DL} = \frac{a + s_y \cdot t_{f,\alpha} \sqrt{\frac{1}{N_a} + \frac{1}{N_c} + \frac{\bar{x}^2}{Q_{xx}}} - a}{b}$$

$$= \frac{s_y}{b} \cdot t_{f,\alpha} \sqrt{\frac{1}{N_a} + \frac{1}{N_c} + \frac{\bar{x}^2}{Q_{xx}}} \tag{37}$$

with $f = N_c$ - 2 degrees of freedom

Since only the upper limit of the distribution range is considered here, this is called a single-sided (or one-tailed) question, and this must be kept in mind when determining the t-value from the table.

In addition, the t-factor is dependent upon both the acceptable error probability, α, and the number of degrees of freedom, f. For reasons of conformity, especially in comparing boundaries, an α of 5% should be chosen. Both the error probability, α, and the number of degrees of freedom must be stated in every case.

1.3.2
Determining the Minimum Detectable Value [34, 120]

A quantitative determination with a stated concentration is only possible when the analytical result is the same as or greater than the minimum detectable value (x_{MDV}), since only then is the required level of significance achieved (see Figure 1-12) [192].

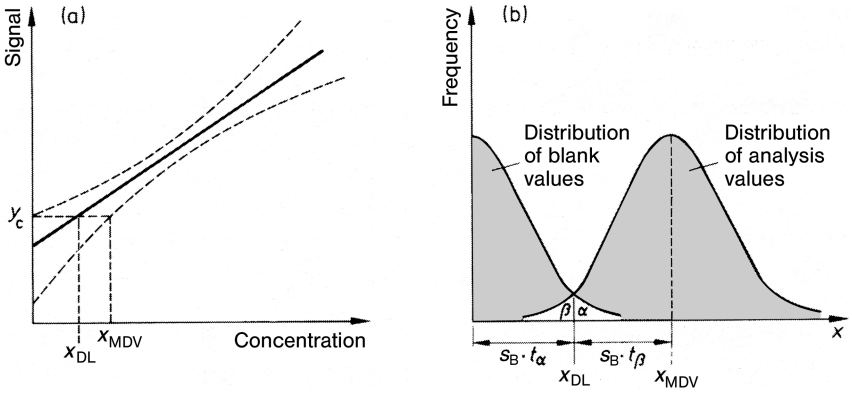


Fig. 1-12 Minimum detectable value.
a) obtained using the calibration function with its associated confidence interval;
b) obtained using the distribution of blank values; important assumption: *s* is identical.

This is also a single-sided (or one-tailed) question, which must again be kept in mind when determining the *t*-value from the table.

Like the decision limit, x_{DL}, the numerical value of the minimum detectable value, x_{MDV}, is dependent upon the chosen levels of significance and the number of degrees of freedom.

For equal significance levels α and β ($\alpha = \beta$), it follows that:

$$t_\alpha = t_\beta = t$$

and therefore

$$x_{MDV} = 2\, x_{DL}$$

If different significance levels are chosen, x_{MDV} must take into account the different values of *t*.

1.3.2.1 Minimum Detectable Value, Determined Using the Distribution of Blank Values

The general equation for the minimum detectable value, x_{MDV}, following the blank value method is:

$$x_{MDV} = x_{DL} + \frac{s_B}{b} \cdot t_{f,\beta} \sqrt{\frac{1}{N_a} + \frac{1}{N_B}} \tag{38}$$

($f = N_B - 1$ degrees of freedom)

1.3.2.2 Minimum Detectable Value, Obtained Using the Calibration Function

The general equation for the minimum detectable value following the calibration method is:

$$x_{MDV} = x_{DL} + s_{xo} \cdot t_{f,\beta} \sqrt{\frac{1}{N_a} + \frac{1}{N_c} + \frac{\bar{x}^2}{Q_{xx}}} \tag{39}$$

($f = N_c - 2$ degrees of freedom)

1.3.3 Limit of Quantification [34]

The limit of quantification is defined as the minimum concentration of a substance (x) that can be analyzed with a given maximum relative imprecision or result uncertainty, Δx_{rel}.

Example (approximately calculated):

- The determined result uncertainty is $\Delta x = 10$ mg/l.
- A relative result uncertainty, $\frac{\Delta x}{x}$, of at most 10 % is assumed.
- Valid analytical results x must therefore be larger or the same as
  $$\frac{\Delta x}{10\,\%} = \frac{10\,\text{mg/l}}{0.1} = 100\,\text{mg/l}$$
- The limit of quantification here amounts to $x_{LQ} = x = 100$ mg/l. The limit of quantification is $k = \frac{1}{10\%} = \frac{1}{0.1} =$ ten times larger than the result uncertainty.

Table 1-1 *k*-Factors for the determination of the limit of quantification.

Maximum acceptable relative result uncertainty (%)	*k*-Factor
5	20
10	10
15	6.7
20	5
25	4
33.3	3
50	2

Calculation of the Limit of Quantification

To calculate the limit of quantification with the aid of the calibration data, α and N_a are needed in addition to the k-factor (see Table 1-1). The latter is defined as:

$$\frac{\Delta x_{LQ}}{x_{LQ}} = \text{maximum acceptable relative result uncertainty} = \frac{1}{k} \tag{40}$$

The half-width of the double-sided prognosis interval of the limit of quantification for future analyses, Δx_{LQ}, is calculated as follows:

$$\Delta x_{LQ} = s_{xo} \cdot t_{f,\alpha} \sqrt{\frac{1}{N_a} + \frac{1}{N_c} + \frac{x_{LQ} - \bar{x}^2}{Q_{xx}}} \tag{41}$$

($f = N_c - 2$ degrees of freedom)

From the definition of the relative prognosis interval (see Eq. 40):

$$x_{LQ} = k \cdot \Delta x_{LQ} \tag{42}$$

which yields the quadratic equation

$$x_{LQ} = k \cdot s_{xo} \cdot t_{f=N_c-2,\alpha} \sqrt{\frac{1}{N_a} + \frac{1}{N_c} + \frac{(x_{LQ} - \bar{x})^2}{Q_{xx}}} \tag{43}$$

A fair approximation of the limit of quantification can be obtained by multiplying the previously calculated decision limit by the k-factor: $x_{LQ} = k \cdot x_{DL}$.

The limit of quantification must *always* be expressed together with the k-value.

1.3.4 Quick Estimation

Under similar conditions (N_c = 10, or N_B = 10 and N_a = 1), the decision limit, minimum detectable value, and the limit of quantification can be quickly approximated using the following formulae [34]:

	for α = 5%	for α = 1%
With the blank value method:	$x_{DL} = 1.9\ s_B/b$	$= 3\ s_B/b$
	$x_{MDV} = 3.8\ s_B/b$	$= 6\ s_B/b$
	$x_{LQ} = 1.9 \cdot k \cdot s_B/b$	$= 3 \cdot k \cdot s_B/b$
With the calibration method:	$x_{DL} = 2.3\ s_{xo}$	$= 3.6\ s_{xo}$
	$x_{MDV} = 4.6\ s_{xo}$	$= 7.2\ s_{xo}$
	$x_{LQ} = 2.3 \cdot k \cdot s_{xo}$	$= 3.6 \cdot k \cdot s_{xo}$

1.3.5
Estimation of the Decision Limit and Limit of Quantification Using the *S/N* Ratio

Because a quick estimation using the blank value method is unsuitable for chromatographic procedures, and using the calibration procedure is often too time-consuming, in the revised version of ISO/WD 13530 [136] a simple estimation protocol for these analytical procedures is described.

The decision limit x_{DL} is defined as that concentration for which the signal-to-noise ratio (S/N) equals 3. The limit of quantification results from

$$x_{LQ} = 3 \cdot x_{DL}$$

whereby the factor 3 corresponds to a 33.3% relative result uncertainty.

1.4
Validation of Individual Process Steps and Examination of Matrix Influences

A major quality criterion of an analytical process is its applicability to complex samples. It is then expected that the developer of a method examines the process for influences such as:

- required additional steps such as sample digestion or extraction,
- interferences or matrix effects.

Typical matrices (e.g., actual surface water of typical composition) should be chosen.

Process steps and matrix effects can manifest themselves as an increase in imprecision and/or as constant systematic or proportional systematic deviations of the analytical result from the "true" value.

Calculation of the recovery function allows one to detect systematic (constant systematic as well as proportional systematic) errors, to check individual process steps, and to ascertain the influence of the matrix [100, 209].

1.4.1
Systematic Errors

1.4.1.1 Constant Systematic Errors, Additive Deviations

For constant systematic errors (see Figure 1-13), the deviation is independent of the concentration of the analyzed components, which leads to a parallel displacement of the matrix calibration line 2 (with constant systematic deviation) in relation to the calibration line 1 (prepared with pure standard solutions). The cause of this additive deviation may be the co-detection of a matrix component; the analytical procedure is therefore insufficiently specific.

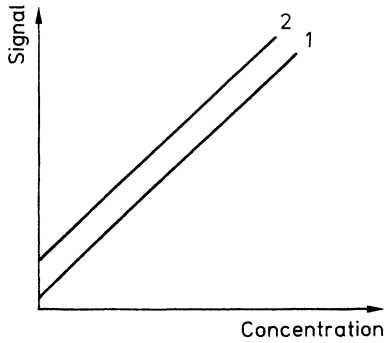


Fig. 1-13 Representation of a constant systematic deviation (no. 2).

1.4.1.2 **Proportional Systematic Errors, Multiplicative Deviations**

For proportional systematic deviations (see Figure 1-14), the extent of the deviation is dependent on the concentration of the analyzed components. This leads to a change in the slope of the matrix calibration line 2.

These multiplicative deviations from the true value can be the result of individual process steps (sample digestion, sample extraction, interference with frits or glassware, etc.) or matrix effects. Systematic errors can be detected by standard additions and/or the determination of the recovery function [100, 209].

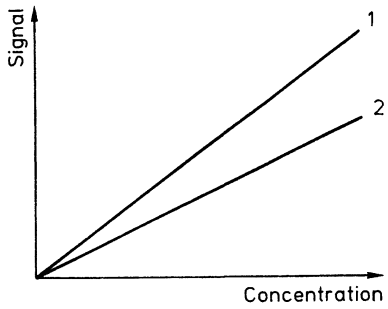


Fig. 1-14 Proportional systematic deviation (no. 2).

1.4.2
Establishment and Assessment of the Recovery Function

The objective of a recovery experiment and the establishment of a recovery function is the determination of the influence of a procedure or sample modification (also the influence of the matrix) on the analytical process. The experiments are carried out over the entire working range. Initially, the calibration function of the fundamental analytical procedure is determined:

$$y = a_c + b_c \cdot x_c \tag{44}$$

Each individual calibration sample is then subjected to the modified analytical procedure.

The analytical results x_f are then calculated using the found signal values y_f and the analysis function (the calibration function solved for x):

$$x_f = \frac{y_f - a_c}{b_c} \tag{45}$$

If the "found concentrations" (x_f) are plotted on the ordinate versus the original calibration concentrations (x_c) on the abscissa, then the *recovery curve* is obtained, which can be described mathematically by the recovery function (linear regression line):

$$x_f = a_f + b_f \cdot x_c \tag{46}$$

In the ideal case, the recovery function results in a line with intercept $a_f = 0$ and slope $b_f = 1$, as well as a residual standard deviation, s_{y_f}, that corresponds to the standard process deviation of the fundamental analytical procedure, s_{xo_c}.

1.4.2.1 Prerequisites for the Interpretation of the Recovery Function

An important prerequisite for the significance of the recovery function is the equivalence of the process standard deviation, s_{xo_c}, of the calibration function of the fundamental analytical procedure, and s_{y_f} of the calibration function of the spiked matrix or the calibration function of individual preliminary sample treatments (e. g., digestion or extraction). A matrix or a preliminary sample treatment could lead to markedly higher imprecision of the calibration (see Figure 1-15), which could possibly mask any systematic errors present.

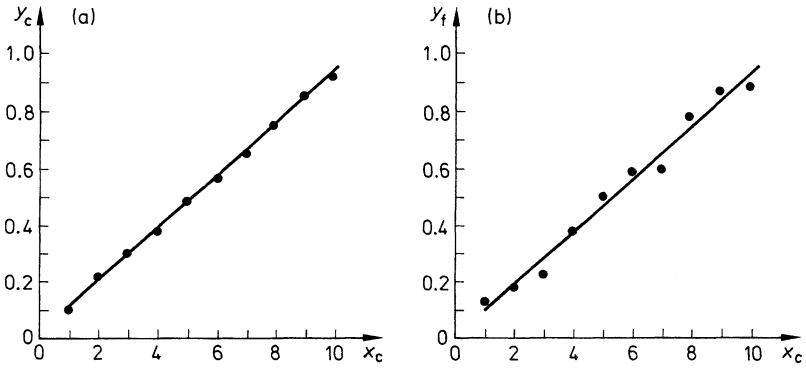


Fig. 1-15 Calibration functions
a) for aqueous standard solutions, b) for a matrix.

Checking Analytical Precision

The process standard deviation of the calibration function of the fundamental procedure, s_{xo_c}, and the residual standard deviation of the recovery function, s_{y_f}, are tested for significant difference:

$$TV = \left(\frac{s_{y_f}}{s_{xo_c}}\right)^2 \tag{47}$$

If $TV > F$ ($f_1 = f_2 = N_c - 2$, $P = 99\,\%$), then there is a significant difference between the two standard deviations.

In this case, no final decision with respect to the presence or absence of systematic deviations can be made. Instead, the cause of the high imprecision must be found and a new recovery function determined.

Note: The comparison of s_{y_f} and s_{xo_c} by means of an *F*-test is allowable here, since both standard deviations are calculated in concentration units.

1.4.2.2 Testing for Systematic Errors

Since measurements always contain random errors, i.e. they fall within a range, the ideal values of $a_f = 0$ and $b_f = 1$ for intercept and slope, respectively, are never obtained. In order to make a statement about the presence of systematic deviations, the confidence intervals for a_f and b_f must be determined [14].

$$CI\,(a_f) = a_f \pm t_{f,P} \cdot s_{a_f} \tag{48}$$

$$CI\,(a_f) = a_f \pm t_{f,P} \cdot s_{y_f} \cdot \sqrt{\frac{1}{N_f} + \frac{\bar{x}_c^2}{Q_{xx}}} \tag{49}$$

with

$$s_{y_f} = \sqrt{\frac{\sum [x_{if} - (a_f + b_f \cdot x_{ic})]^2}{N_f - 2}} \tag{50}$$

and N_f = the number of concentration levels.

The confidence interval for b_f can be represented by:

$$CI\,(b_f) = b_f \pm t_{f,P} \cdot s_{b_f} \tag{51}$$

$$CI\,(b_f) = b_f \pm \frac{t_{f,P} \cdot s_{y_f}}{\sqrt{Q_{xx}}} \tag{52}$$

(t = Student's t-factor: $f = N_f - 2$, $P = 95\,\%$)

The presence of systematic errors can be tested for by means of the calculated confidence interval.

If the confidence interval $CI(a_f)$ does *not* include the value $a_f = 0$, a constant systematic error is present with 95 % statistical certainty.

By the same reasoning, if the confidence interval $CI(b_f)$ does *not* include the value $b_f = 1$, then a proportional systematic deviation is present with 95 % certainty.

1.4.3 Application of the Recovery Function

1.4.3.1 Checking Individual Process Steps

If constant or proportional systematic errors are found during the checking of individual process steps (e. g., extraction), then the cause(s) of these errors should be sought if possible. The analytical process should then be optimized and the measurements should be repeated in order to determine the recovery function.

If systematic errors cannot be eliminated, then the process documentation must *clearly* indicate this and in practice either the calibration must be performed over the entire process (including sample preparation) or the method of standard addition must be applied.

For *exclusively* proportional systematic errors, the recovery rate (RR) can be given. This is determined from the slope b_f:

$$RR = b_f \cdot 100\,(\%) \tag{53}$$

1.4.3.1.1 Meaning of the Recovery Rate

The recovery rate is an assessment criterion for a given analytical process or individual process step [154]. If a recovery rate of 100 % is obtained when verifying individual process steps, and the process in question is free from constant systematic errors, the determination of analytical results need not be performed using the method of standard addition (see Section 3.4.1.1).

However, if a proportional systematic error is found, future analytical results must be obtained using the method of standard addition; if a constant systematic error is discovered, a corresponding warning must be entered in the analysis protocol.

The recovery rate may also be used as a control quantity for a quality control chart (see Section 2.6.7.1.5.3). The central line can be chosen as either $RR = 100\,\%$ or as equal to the mean recovery rate. This is dependent on the results of the preliminary period.

1.4.3.1.2 Impact of a Constant Systematic Error on the Recovery Rate

Attempting to describe the trueness of an analytical process by using the recovery rate (RR) derived by replicate analyses of one sample alone leads to incorrect results if an additional constant systematic error is present (see Figure 1-16).

For every recovery function having either

- a positive intercept ($a_f > 0$) and a slope $b_f < 1$ (1)
- or a negative intercept ($a_f < 0$) and slope $b_f > 1$ (2)

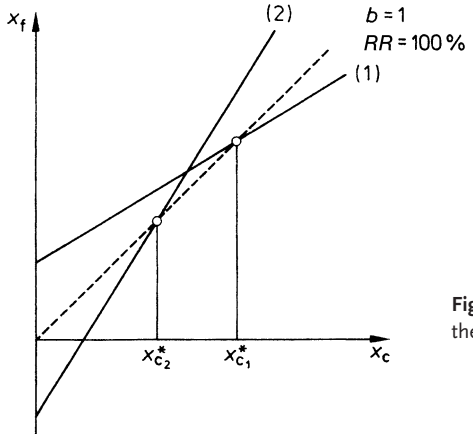


Fig. 1-16 Dependence of the recovery rate on the slope.

an RR of 100 % can be found for a single concentration $x_c^* > 0$ (intersection point of the recovery curve with the angle bisect). For all concentrations $x_c \neq x_c^*$ however, the $RR \neq 100\,\%$.

In general, and with respect to the intercept, a_f, it follows that

$$RR = \frac{x_f}{x_c} \cdot 100\% = \left(\frac{a_f}{x_c} + b_f\right) \cdot 100\% \tag{54}$$

The RR is independent of x_c and is determined solely by b_f only if $a_f \approx 0$.

a) Proportional errors only:

Example:

$x_f = 0.001 + 0.78 \cdot x_c$

$RR = \left(\frac{0.001}{x_c} + 0.78\right) \cdot 100\%$

$\boldsymbol{RR_1}$ $(x_c = \mathbf{10}) = (0.0001 + 0.78) \cdot 100\,\% = \mathbf{78.010\,\%}$

$\boldsymbol{RR_2}$ $(x_c = \mathbf{50}) = (0.00002 + 0.78) \cdot 100\,\% = \mathbf{78.002\,\%}$

$\boldsymbol{RR_3}$ $(x_c = \mathbf{100}) = (0.00001 + 0.78) \cdot 100\,\% = \mathbf{78.001\,\%}$

b) With additional significant constant error:

Example:

$x_f = 1.0 + 0.78 \cdot x_c$

$RR = \left(\frac{1.0}{x_c} + 0.78\right) \cdot 100\%$

$\boldsymbol{RR}$ $(x_c = \mathbf{10}) = (0.1 + 0.78) \cdot 100\,\% = \mathbf{88\,\%}$

$$\begin{aligned} \boldsymbol{RR}\ (x_c = \mathbf{50}) &= (0.02 + 0.78) \cdot 100\,\% \\ &= \mathbf{80}\,\% \\ \boldsymbol{RR}\ (x_c = \mathbf{100}) &= (0.01 + 0.78) \cdot 100\,\% \\ &= \mathbf{79}\,\% \end{aligned}$$

Consequence: the constant systematic error a_f adds to the RR as $\frac{a_f}{x_c}$.

This portion of the error has its greatest value at the lower end of the range, $x_c = x_1$, and decreases hyperbolically with increasing x_c.

If, for example, the RR is allowed to be subject to a maximum inaccuracy due to a_f of 1%,

$$\frac{a_f}{x_c} \leq 0.01$$

then the constant systematic error must be $a_f \leq 0.01 \cdot x_1$.

Example of Checking a Process Step: Optimization of an Extraction Step [140]

A solid-phase extraction (using EXTRELUT® columns) was performed before the quantitative HPTLC determination of phenobarbital in urine. In order to check solely the extraction step, six standard solutions of phenobarbital were submitted to the entire extraction procedure: 5 ml standard solutions were diluted with 15 ml of sodium phosphate buffer (pH 6.0), applied to the EXTRELUT® columns, and eluted after 10 minutes with 40 ml of a mixture of a dichloromethane/2-propanol (93 : 3, *v*/*v*). The extracts were then concentrated to dryness in a nitrogen stream, and the residues were taken up in 500 µl of methanol and applied to HPTLC silica gel plates.

Non-extracted standards of the same concentration were also concentrated under nitrogen, and the residues were redissolved in methanol and analyzed in exactly the same way as the extracted standard solutions.

A calibration function was determined using the measured values of the non-extracted standards and their respective concentrations. The recovery function

$$x_f = -100.2\ \text{ng/spot} + 1.098 \cdot x_c$$

was obtained by linear regression of the found concentrations x_f versus the calibration concentrations x_c (see Figure 1-17).

The confidence intervals of a_f and b_f were found to be:

$$CI(a_f) = (-100.2 \pm 13.698)\ \text{ng/spot}$$

$$CI(b_f) = 1.098 \pm 0.0969$$

that is, both a constant systematic and a small proportional systematic error were present.

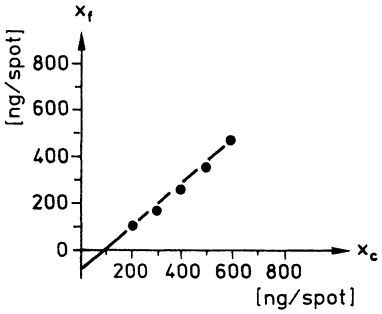


Fig. 1-17 Recovery function 1 for the phenobarbital extraction.

The cause of these systematic deviations was found to be insufficient loading of the EXTRELUT® column. Approximately 1 ml of the EXTRELUT® material at the lower end of the column remained unwetted when the column was charged with a sample volume of only 20 ml. This unwetted part of the absorbing material reabsorbed a portion of the eluted phenobarbital during elution of the column.

For this reason, an additional 1.5 ml of phosphate buffer was added to the original sample volume (20 ml) and the recovery experiment was repeated. The recovery function (see Figure 1-18)

$$x_f = 2.37 \text{ ng/spot} + 0.98 \cdot x_c$$

with the confidence intervals

$$CI(a_f) = (2.37 \pm 13.975) \text{ ng/spot}$$

$$CI(b_f) = 0.98 \pm 0.082$$

was determined. Constant or proportional systematic errors were no longer present.

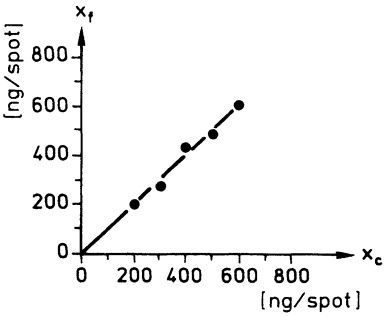


Fig. 1-18 Recovery function 2 for the phenobarbital extraction.

1.4.3.2 Determination of the Recovery Function to Prove the Influence of a Matrix

In order to prove the possible influence of a matrix, various typical sample matrices should be prepared that do not contain the analyte (possibly a synthetic sample matrix) and these should be divided into ten equal aliquots. A *concentrated* standard solution is then added to each of these ("spiking"), so that each of the matrix partial samples contains the same analyte concentration as in the aqueous solutions used in the calibration.

The "samples" thus produced are then analyzed using the appropriate analytical process. The measured values, y_f, are then converted into concentrations, x_f, using the analysis function (Eq. 45).

The recovery function as well as the residual standard deviation and the confidence interval of the intercept a_f and slope b_f are calculated and evaluated as described in Section 1.4.2.2 (see Eqs. 48 to 51).

Note: If an actual matrix that contains the analyte is used for these experiments, then it is not possible to detect the presence of a constant systematic error.

If this check for matrix influence confirms the presence of a proportional or constant systematic error, then the analysis may *not* be interpreted later using a calibration function determined with aqueous standard solutions, but the method of addition of a standard [138] must be applied in each case.

Example: HPTLC Determination of Selenium in Human Serum [100]

To determine the selenium content in human serum, a wet-chemical sample preparation is necessary. In order to determine the influence of this biological matrix on the recovery rate, the recovery function must first be determined. A serum sample is divided into six parts, five of which are spiked with various seleno-cysteine concentrations. The five spiked samples and one unspiked sample are first concentrated to dryness at 100 °C and then treated with nitric acid and hydrogen peroxide.

The individual samples are then extracted by means of an EXTRELUT® column and the extracts are applied to an HPTLC silica gel 60 plate, separated, and scanned. The following recovery function was obtained (see Figure 1-19):

$$x_f = 10.609 \text{ pg/spot} + 1.002 \cdot x_c$$

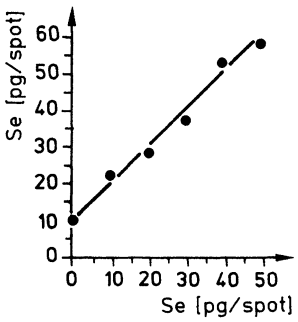


Fig. 1-19 Recovery function for selenium in human serum.

The confidence interval of the slope b_f ($0.87 < b_f < 1.14$) shows that no proportional systematic error is present. In other words, the determination of selenium in human serum does not have to be performed using the method of standard addition. A statement about the presence of a constant systematic error is not possible, since the matrix used for this recovery experiment was not demonstrably free of selenium.

1.5 Additional Statistical Methods

In cases of proven inhomogeneity of variances (see Section 1.2.4.2.1) as well as for disruptions in quantification, e.g. sub-optimal separation of signals (interferences, peak overlap), other more demanding and more complex statistical methods may be used [146, 153]. These include:

- weighted regression [159],
- multiple curve fitting [160],
- multiple regression [146],
- multivariate standard addition [146],
- process functions instead of process data [146].

However, these methods are beyond the scope of this book. Besides the increased calculational demands, the analytical requirements associated with these procedures are also very high.

1.6 Use of Internal Standards [50]

Multiple component procedures are analysis procedures whereby a single sample is analyzed for more than one component at a time. Gas chromatography and ICP-AES are examples. In principle, strategies involving calibration of the fundamental analytical procedure and verification by means of the recovery function, as described in the previous sections, are applicable to these procedures.

For these analyses, additional mutual influences must also be taken into account. Quality assurance measures here also include the control of equipment parameters. This is achieved, for example, through the use of internal standards.

1.6.1 Definition, Purpose

An internal standard is a substance which is known not to be present in the analysis sample and which is added in defined form and quantity to each calibration and analysis sample to be quantitatively co-analyzed. It is assumed that both the standard and the analyte are subject to the same physicochemical influences,

such as the parameter to be examined. Analytical results obtained with internal standards are used:

- For checking purposes: Assuming that the internal standard is added *after all preparation steps*, immediately before the measurement, then the results obtained with an internal standard can be related to the control of the sample application (e.g., injection) or sample manipulation within the equipment, or to the detection step. In contrast, the addition of standards *before sample preparation* makes it possible to check specific preparation steps.
- Under certain preconditions, for mathematical correction of results when systematic equipment errors occur.

1.6.2 Conditions and Limitations of the Use of Internal Standards

A substance which is to be used as an internal standard must fulfill certain conditions:

- It must not, with a high degree of certainty, be present in the natural sample.
- It may not itself cause matrix effects.
- Its physicochemical properties must be as similar as possible to those of the analyte. For example, this includes a similar boiling point (in chromatography), a similar but not identical retention time, similar detectability, etc. If sample preparation steps are to be checked using internal standards, then similar chemical properties are necessary as well.
- It must be possible to add the internal standard in highly concentrated form in order to avoid volumetric errors. Otherwise, the additional volume in the sample must be considered when calculating the results of the analysis.
- The concentration of the internal standard in the sample must be suited to the measurement problem.
- A calibration and determination of the process data is obviously also necessary for an internal standard.

The use of an internal standard does not dispense with the need to perform a complete calibration on each parameter to be analyzed as well as perform experiments to determine process and matrix influences. Internal standards are also not a substitute for calibration functions.

1.6.3 Procedure

In order to monitor or correct an analytical process with the aid of one or more internal standards, it is necessary to submit the entire analytical process to a complete calibration and checking of process data. The influence of the individual process steps and matrices on the analytical process requires special scrutiny. If despite all efforts instabilities or matrix effects cannot be brought under control,

then one may consider the use of one or more internal standards. Two or more internal standards may be used simultaneously when different influence quantities are to be monitored selectively. For example, internal standards of varying retention times can be used to monitor the entire separation and the stability of the detector for time-intensive chromatographic separations.

After the selection of suitable standard substances, a basic calibration is performed for these and tests are carried out to determine the influences of individual process steps and especially matrices. After the process data have been obtained for the substances to be used as internal standards, a calibration is performed on the standard mixtures. Repeated analyses of known samples complete the preliminary experiments.

If the internal standards serve only for checking purposes, then the results of these analyses form the basis of appropriate quality control systems/charts in routine analysis (see Section 2.6).

If internal standards are used for the systematic correction of results, then their recovery rates are included in each individual result [50]. The recovery rate of an internal standard is regarded as representative of the recovery rate of the analyzed substances. To obtain the corrected analytical results, the uncorrected results should be divided by the recovery rate of the internal standard.

Example:

- The uncorrected result is 50 µg/l.
- A similar substance used as the internal standard is present at a concentration of 78 µg/l. The analytical result for this internal standard is obtained as 62 µg/l, indicating a recovery rate, RR, of $\frac{62}{78} = 0.795$.
- Using this RR to correct the result of 50 µg/l, the adjusted concentration is $\frac{50\ \mu g/l}{RR} = \frac{50\ \mu g/l}{0.795} = 62.9\ \mu g/l$.

Provided that the concentration of the internal standard is known and is constant throughout, the quotient

$$\frac{\text{measured value of parameter in question}}{\text{measured value of internal standard}}$$

can be used for the calibration and evaluation instead of the original measured values of the parameter to be examined.

1.7
Preparing for Routine Analysis

1.7.1
Examination of the Time Dependency of Measured Values

A quality objective for routine analysis is to maintain constant precision and accuracy of the results over a longer time period, or in other words to achieve a reliable analytical process.

The calibration is, as a rule, performed under so-called repetitive conditions, but nevertheless within a small time frame. The observation of analysis quality over a longer time period makes apparent a higher imprecision in the analytical result, due perhaps solely to such factors as longer pauses between individual series of analyses, environmental influences, and the varying daily "state of mind" of personnel. It may be possible that this additional imprecision prevents any kind of comparison of the analytical results. Additionally, temporal instability of analysis parameters is noticeable for some analytical procedures. The effects of aging, contamination, etc., can lead to systematically lower values within a certain operational time, even within a working day.

The objective in preparing for a routine analysis is the recognition of such effects and influences, their quantification, and, where possible, their elimination. So-called control samples (see Section 2.5) are necessary for their detection; these are samples which remain stable over longer periods of time and are analyzed as controls in each analysis series.

1.7.1.1 Comparison of the "Within Batch" Standard Deviation (s_w) with the "Between Batches" Standard Deviation (s_b) [215]

For precision analyses, the standard deviation within batch, s_w, and the standard deviation between batches, s_b, are determined and tested for significant difference using an F-test.

The determination of s_b and s_w is performed on $N = 20$ consecutive days (analysis series of a day = batch; at least 6 to 10 different analysis days or analysis series) and multiple, or at least duplicate, determinations are made for each control sample. The individual standard deviation (s_i) within batch i, the within batch standard deviation (s_w) of all batches, the between batch standard deviation (s_b), and the total standard deviation (s_t) are all calculated from the results.

The standard deviation of N_i determinations of a control sample within batch i:

$$s_i = \sqrt{\frac{1}{f_i} \cdot \sum \left(x_{ij} - \bar{x}_i\right)^2} \qquad \text{with } f_i = N_i - 1 \tag{55}$$

for duplicate determinations: $f_i = 2 - 1 = 1$

Within batch standard deviation:

$$s_w = \sqrt{\frac{1}{f_w} \cdot \sum \left(f_i \cdot s_i^2\right)} \tag{56}$$

where $f_w = \sum f_i$; for duplicate determinations $f_w = N$ (57)

Between batch standard deviation (from batch to batch or day to day):

$$s_b = \sqrt{\frac{1}{f_b} \sum \left(\bar{x}_i - \bar{\bar{x}}\right)^2} \tag{58}$$

where $f_b = N - 1$ and $\bar{\bar{x}} =$ grand average $= \frac{1}{N} \cdot \sum \bar{x}_i$ (59)

The total standard deviation is thus given by:

$$s_t = \sqrt{\frac{1}{f_t}\left(s_b^2 \cdot f_b + s_w^2 \cdot f_w\right)} \tag{60}$$

where $f_t = f_b + f_w$; for duplicate determinations: $f_t = 2 \cdot N - 1$ (61)

Evaluation:

1. The total standard deviation must fulfill the quality requirements (maximum allowable imprecision).
2. The between batch standard deviation should be no more than twice the within batch standard deviation; if necessary, the variance *F*-test may be applied to s_b and s_w in order to assess the precision:

   Calculation of the test value: $TV = \frac{s_b^2}{s_w^2}$

Assessment of the test value: if $TV \leq F\,(f_b, f_w,\ P = 95\,\%)$, then s_b is only coincidentally larger than s_w.

If $TV > F\,(f_b, f_w,\ P = 95\,\%)$, then the imprecision between the series significantly influences the total imprecision. In the event of the total standard deviation being unacceptably high, significant quality improvements can be achieved by tracking and correcting analysis influences, and/or by readjusting or even recalibrating each series.

A Practical Example from Water Analysis

To determine dissolved organic carbon (DOC), a control sample with a content of 5 mg/l was chosen and a duplicate determination was performed on each of ten consecutive days. The results were:

$s_w = 0.099$ mg/l

$s_b = 0.224$ mg/l

$s_t = 0.158$ mg/l

$$TV = \frac{s_b^2}{s_w^2} = 5.119 \quad \text{(table value: } F(9, 10, 99\%) = 4.94\text{)}$$

Since $TV > F(9, 10, 99\%)$, s_b and s_w are significantly different.

The cause of the increased imprecision is in this case an unstable calibration function. Daily recalibration could substantially improve the analytical results.

1.7.1.2 Determining the Need for Daily Adjustment of Analytical Equipment

If a result is determined to be time dependent, then a daily adjustment before each series of analyses using one or two standard solutions is recommended in order to determine the *current slope* of the calibration function. A prerequisite for daily adjustment is, however, that the process standard deviations, s_{xo}, for various slopes are not significantly different (see Figure 1-20b). This can be tested by means of an *F*-test.

For a constant residual standard deviation s_y (case (a) in Figure 1-20), before applying a daily adjustment one must determine by how much the *current slope* of the calibration function (determined using the daily adjustment) may be decreased while still allowing evaluation over the entire chosen analytical range. Verification by means of the maximum acceptable imprecision is recommended (see Section 1.3.3) [22].

If the slope of the calibration function decreases with constant residual standard deviation s_y (Figure 1-20a), then it may be that the required precision $CI_{rel,req}$ for x_1 can no longer be maintained.

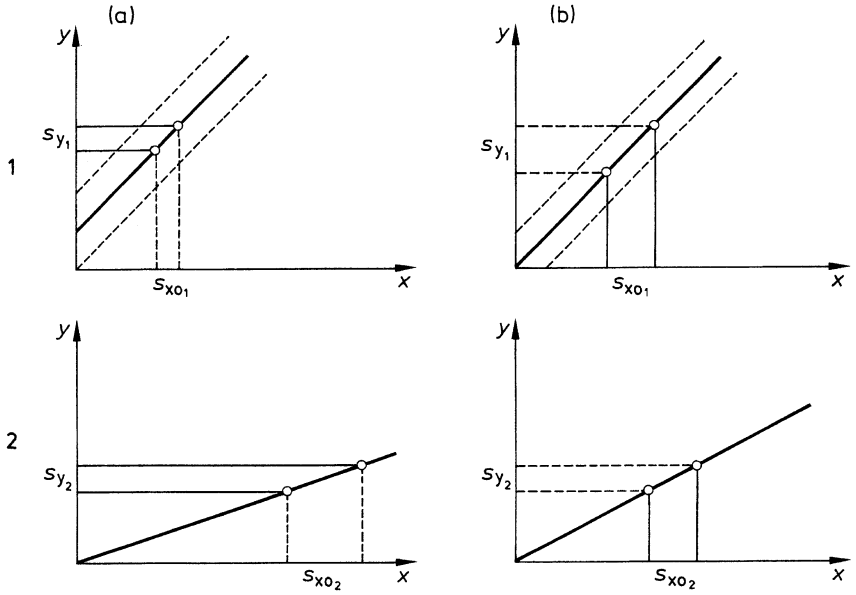
If no imprecision limit is available from $CI_{rel,req}$, then the lower range limit must be tested using the test value x_a (see Section 1.2.4.2.3 and Fig. 1-21). Below the concentration x_a, no definitive analytical statement can be made. The practitioner may, however, perform quantitative analysis above this point provided that the values stay above the limit of quantification of the analytical process.

1.7.1.3 The Trend Test

The procedure described in Section 1.7.1.1 is not suitable to distinguish between a time-dependent *systematic trend* and a simple imprecision. Therefore, a simple trend test (according to Neumann [190]) should be applied.

A prerequisite for the execution of the trend test is the availability of a series of temporally successive analytical results of a control sample $x_1, x_2, \ldots, x_n$, which originated from a normally distributed base entity. At least $n = 20$ analysis results should be included, if possible.

In addition to the standard deviation, *s*, of the *n* values, the mean quadratic of the $n - 1$ differences of consecutive values (single or mean values of consecutive series; successive difference dispersion) Δ^2 is calculated:



(a)
$s_{y_1} = s_{y_2}$; $b_1 > b_2$
$\frac{s_{y_1}}{b_1} = s_{xo_1} < s_{xo_2} = \frac{s_{y_2}}{b_2}$

(b)
$s_{y_1} > s_{y_2}$; $b_1 > b_2$
$\frac{s_{y_1}}{b_1} = s_{xo_1} = s_{xo_2} = \frac{s_{y_2}}{b_2}$

Fig. 1-20 Precision and sensitivity are mutually dependent.

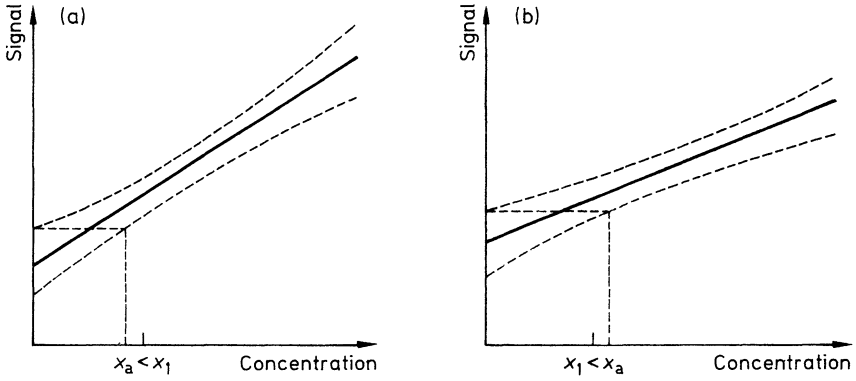


Fig. 1-21 Statistical check of a range by means of x_a for daily adjustment.
a) x_1 is significantly different from the concentration zero,
b) x_1 is *not* significantly different from the concentration zero.

$$\Delta^2 = \frac{\sum (x_i - x_{i+1})^2}{(n-1)} \tag{62}$$

If the consecutive values are independent, then $\Delta^2 \approx 2s^2$.
The hypothesis that consecutive values are independent must be rejected in favor of the alternative hypothesis that a trend exists, if the quotient

$$\frac{\Delta^2}{s^2} = \frac{\sum (x_i - x_{i+1})^2}{\sum (x_i - \bar{x})^2} \tag{63}$$

is less than or equal to the critical limit (n, $\alpha = 1\%$) (see Table 1-2).

Table 1-2 Critical limits for the quotients from the mean quadratic successive difference dispersion and the variance [190].

n	0.1%	1%	5%	*n*	0.1%	1%	5%
4	0.5898	0.6256	0.7805	33	1.0055	1.2283	1.4434
5	0.4161	0.5379	0.8204	34	1.0180	1.2386	1.4511
6	0.3634	0.5615	0.8902	35	1.0300	1.2485	1.4585
7	0.3695	0.6140	0.9359	36	1.0416	1.2581	1.4656
8	0.4036	0.6628	0.9825	37	1.0529	1.2673	1.4726
9	0.4420	0.7088	1.0244	38	1.0639	1.2763	1.4793
10	0.4816	0.7518	1.0623	39	1.0746	1.2850	1.4858
11	0.5197	0.7915	1.0965	40	1.0850	1.2934	1.4921
12	0.5557	0.8280	1.1276	41	1.0950	1.3017	1.4982
13	0.5898	0.8618	1.1558	42	1.1048	1.3096	1.5041
14	0.6223	0.8931	1.1816	43	1.1142	1.3172	1.5098
15	0.6532	0.9221	1.2053	44	1.1233	1.3246	1.5154
16	0.6826	0.9491	1.2272	45	1.1320	1.3317	1.5206
17	0.7104	0.9743	1.2473	46	1.1404	1.3387	1.5257
18	0.7368	0.9979	1.2660	47	1.1484	1.3453	1.5305
19	0.7617	1.0199	1.2834	48	1.1561	1.3515	1.5351
20	0.7852	1.0406	1.2996	49	1.1635	1.3573	1.5395
21	0.8073	1.0601	1.3148	50	1.1705	1.3629	1.5437
22	0.8283	1.0785	1.3290	51	1.1774	1.3683	1.5477
23	0.8481	1.0958	1.3425	52	1.1843	1.3738	1.5518
24	0.8668	1.1122	1.3552	53	1.1910	1.3792	1.5557
25	0.8846	1.1278	1.3671	54	1.1976	1.3846	1.5596
26	0.9017	1.1426	1.3785	55	1.2041	1.3899	1.5634
27	0.9182	1.1567	1.3892	56	1.2104	1.3949	1.5670
28	0.9341	1.1702	1.3994	57	1.2166	1.3999	1.5707
29	0.9496	1.1830	1.4091	58	1.2227	1.4048	1.5743
30	0.9645	1.1951	1.4183	59	1.2288	1.4096	1.5779
31	0.9789	1.2067	1.4270	60	1.2349	1.4144	1.5814
32	0.9925	1.2177	1.4354	∞	2.0000	2.0000	2.0000

Meaning of the Trend Test

The trend test may be used to check if, compared to the within batch standard deviation (s_w), a larger standard deviation between batches (s_b, see Section 1.7.1.1) is a result of a temporal systematic deterioration of the detection system. As an example, the aging of a detector in gas chromatography can be determined in this way.

Consequence

If a temporal trend can be proven, then the quality of the analytical results can be significantly improved by readjustment or even recalibration in each series.

1.8 Summary of the Results of Phase I (Process Development): Documentation

The final documentation of a fundamental analytical process includes, in addition to the description of the analysis procedure, all relevant information about the analysis quality.

Since the quality of analytical results is closely related to the quality of the instructions, the latter should be clear, understandable, and practically oriented [17, 123] so that an unacquainted analyst can follow them to obtain satisfactory results.

The instructions are to be completed with the following information, for the most part obtained in Phase I:

- the substance to be analyzed (analyte),
- area of application, type of possible bulk (or gross) samples,
- tolerances (e.g., "add 6.0 ml ± 0.1 ml"), in order to obtain the necessary precision at each analytical step,
- measuring instructions, including physical quantity and unit of the measured value,
- the procedure to determine the calibration function and the blank value,
- the working range,
- calibration instructions: number, type, preparation of standard solutions,
- coefficients of the calibration function with information about the linearity, sensitivity, and residual standard deviation,
- the process standard deviation,
- the decision limit, with information about how it was obtained including how many measurements were used,
- the relative imprecision at the lower limit of the range ($CI_{rel}(x_1)$) for later comparison with the required limit of quantification,
- instructions for the evaluation and description of the results,
- frequently seen process errors and accompanying countermeasures,
- quantitative information about interfering substances,

- qualitative and quantitative information about constant and proportional systematic deviations as well as their causes and elimination,
- temporal stability (perishability) and storage information for samples, reagents, standard solutions, and reference materials,
- special notes:
  - probable limits on the applicability of the process,
  - special procedures such as readjustment, recalibration, use of internal standards (including evaluation),
  - determination and use of a blank,
- notes about disposal of reagents and occupational safety,
- literature references with further information (e. g., examples of applications).

The documentation is to be supplemented at a later time by means of interlaboratory test data, the routine analyses to follow, and their respective quality assurance processes (see Phase IV).