

CALIBRATION IN QUANTITATIVE ANALYSIS

Part 1. General Considerations

J. AGTERDENBOS

Analytisch Chemisch Laboratorium der Rijksuniversiteit, Croesestraat 77A, 3522 AD Utrecht (The Netherlands)

(Received 22nd December 1978)

SUMMARY

Mathematical procedures for calibration require assumptions to be made, e.g. the homogeneity of variances and the mathematical relationship between the analyte content x and the signal y . Little is known about the magnitude of errors arising from incorrect assumptions. The variation of the standard deviation of the analytical procedure with the content of the analyte, the selection of the type of mathematical relationship between x and y , and the types of errors made in testing hypotheses are discussed. In certain practical situations, the standard deviation (s.d.) is nearly independent of x if $x < 22p$ (p = detection limit) and the relative standard deviation (r.s.d.) is nearly independent of x if $x > 50p$. If the s.d. is constant, calibration relations of the type $y = a + bx$ are frequently to be preferred; with a constant r.s.d., relations of the type $\log y = a + b \log x$ have advantages.

In most types of quantitative instrumental analysis, calibration is vital. Generally a calibration curve is prepared or a corresponding mathematical relation between the content x (amount or concentration) of the analyte and the signal y is calculated and applied to the unknown sample. For the calculation of the relation between x and y , the mathematical procedures generally used are based on the "method of least squares". Only this type of procedure is discussed here. The calculation of the content x of the unknown sample frequently includes calculation of a confidence region for this content. The mathematical formulae applied in calibration have been given in many texts along with the mathematical criteria for their application [1–6]. The paper by Gottschalk [6] is part of a series on standardization in quantitative analytical procedures [7].

Analytical procedures do not generally obey the conditions imposed by mathematics, and some authors — recently Gottschalk [6] and Maessen et al. [5] — have therefore provided statistical methods to test the validity of the assumptions made. At least one point, however, seems to have escaped attention: a statistical test is not necessarily correct if it cannot be proved that it is incorrect. As has been shown by Agterdenbos [8], serious errors may result if this is not realized. Some other points are rarely discussed in analytical calibration. The mathematical relationship applied can be different from the

true relation between content x and signal y . Not much is known about the error caused by discrepancies, or about the way this error is influenced by factors such as the calibration range, or about the magnitude of the discrepancy. Calibrations made are justified only if certain mathematical conditions are fulfilled such as the homogeneity of variances (see below). Here, too, little is known about the error caused by unjustified assumptions and its dependence on the factors mentioned [8].

Another point rarely discussed is the importance of the location of the calibration points on the concentration axis if the number of these points is fixed. Shewell [9] calculated which location of five calibration points gave the best information (smallest confidence region) for the slope and intercept of a linear plot. However, he ignored the mutual dependence of these two parameters [10, 11]. In a recent paper, Franke and de Zeeuw [12] discussed a theory for achieving the optimum location of calibration points in the standard addition method. No general or more detailed discussion of the influence of the location of calibration points on accuracy seems to have been given.

A final point is "bracketing" the measurement of the concentration of an unknown sample by interpolation between two samples of known concentrations. The precision of such a procedure in dependence on some of the factors mentioned above was recently discussed by Schwartz [13].

GENERAL ASPECTS

Two steps may be distinguished in calibration: the preparation of the curve, i.e. calculation of the relation between x and y , and application of the relation to unknown samples. The preparation of the curve is sometimes preceded by or combined with an investigation of the type of mathematical relation between x and y . It always includes the calculation of the parameters a, b, \dots , and sometimes includes tests of the significance of the parameters [6] and of the homogeneity of variances [5, 6]. The application of the curve always includes the calculation of the analyte content of the unknown sample. Sometimes it also includes a comparison of the standard deviation found from the calibration measurements with the standard deviation found from the measurements of the unknown sample. Finally, a confidence region for the analyte content of the unknown sample is sometimes calculated [5, 6].

Types of mathematical relations in calibration

(The symbols and abbreviations used are listed in Table 1.)

Some mathematical relations that are frequently used are:

$$y = bx \quad (1)$$

$$y = a + bx \quad (2)$$

$$\log y = a + b \log x \quad (3)$$

(The Scheibe—Lomakin equation applied in emission spectrometry)

TABLE 1

Symbols and abbreviations

s.d.	= absolute standard deviation = σ or s in formulae ^{a,b}
r.s.d.	= relative standard deviation = σ_r or s_r in formulae (expressed as a fraction)
σ_t or s_t	= total s.d. in formulae
σ_{rt} or s_{rt}	= total r.s.d. in formulae
x	= content: y = signal
$\bar{a}, \bar{b}, \bar{c}$	are the best estimates and a, b, c are the population values of the calibration curve parameters
\bar{a}_j	= the same parameters, with special reference to equation j
Δ	= difference between the measured and calculated value of y
p	= limit of detection
k, q, z	= factors (see text)
α, β	= errors of the first and second kinds, respectively, in testing hypotheses

^a“Standard deviation” in the text means that both s.d. and r.s.d. are considered. The terms “population” or “sample” are added if their omission may lead to confusion.

^b σ is applied for population values, and s for sample values.

$$y = a + bx + cx^2 \quad (4)$$

(Found in atomic absorption spectrometry by Wagenaar et al. [14]; Agterdenbos and Vlogtman [15] found that it may be expected in absorption spectrometry if non-monochromatic light is the cause of non-linearity. The possible validity of eqn. (3) was not investigated by these authors.)

Other equations, rarely encountered, are:

$$\log y = a + \log x \quad (5)$$

$$\log y = a + b \log x + c (\log x)^2 \quad (6)$$

Calculations of the parameters in all the equations is based on the method of least squares, where for example, in eqn. (2) $\Sigma(\Delta_i)^2$ is a minimum for i measurements ($i > 2$) of y . The values calculated are the best estimates for the population parameters. Formulae for the calculation of the best estimates, their standard deviations and confidence intervals are given in standard texts. It should be remembered that these intervals are not independent of each other. For eqn. (2) this has been discussed in more detail [10, 11], and has been applied recently to practical analysis [5, 12]. More complicated relations between x and y may exist and the corresponding parameters can be calculated [13, 16].

With the procedure mentioned, good results are obtained only if some statistical conditions are fulfilled (see below) and if the true relation between x and y can be described adequately by a mathematical relationship.

Conditions to be fulfilled

Several conditions have to be fulfilled to justify the calculation procedure mentioned. One is that the precision of the measurement of the x values is

much better than the precision of the measurement of the y values. A second condition is that values found for y if parallel determinations are made at the same x value have a Gaussian distribution. In the present paper both conditions are assumed to be valid.

A third condition is the "homogeneity of variances" (variance = σ^2 or s^2), i.e. the s.d. of the measured y values must be the same for the whole range of x values covered by the calibration curve. The population s.d. should be considered, not the sample value or the r.s.d. Simple weighting procedures exist to take into account variations in the s.d. with x , if the values of \hat{a} and \hat{b} are calculated. However, if this is done, there is no good procedure for calculation of a confidence region for the analyte content in the unknown sample. Therefore such weighting procedures are not discussed here. The homogeneity of variances is discussed in more detail in the following section.

VARIATION OF THE S.D. OR R.S.D. IN y , WITH x

Some special cases

The condition that the s.d. of y should be independent of x was mentioned above. It will be shown below that by mathematical manipulation this condition may sometimes be changed to the condition that the r.s.d. of y should be independent of x . Therefore in the present section, the variation of the s.d. or r.s.d. of y within the range of x values applied for calibration is discussed in a more general way.

It seems difficult to give a single general discussion for all practical procedures, but some cases where the final measurement only is considered are well known. For counting errors, the s.d. increases with $N^{\frac{1}{2}}$, and the r.s.d. decreases with $N^{-\frac{1}{2}}$ (N = number of counts). In photometric analysis the Ringbom curve [17] is frequently cited, sometimes with some adaptations. In its original form, the Ringbom curve indicates that the r.s.d. in photometric measurements is reasonably constant if absorbance values between about 0.2 and 0.75 are considered. It seems unusual in the literature to consider the variation of the s.d. with absorbance; calculation shows that it is reasonably constant at lower absorbance values, but increases greatly at higher values (Table 2). (If the assumptions made are different from Ringbom's, other values are obtained, but the trend is similar.)

In many analytical procedures the sample preparation step may contribute more to the total error than the final measurement does. Here, too, a general theory for calculating the variation of the s.d. with the analyte content is not available. Agterdenbos [18] suggested that the total standard deviation may sometimes be found by (quadratic) addition of a s.d. and a r.s.d. component; quadratic addition is required because variances are additive. An obvious consequence of this consideration is that at smaller x values the s.d. is constant ("homogeneous") and at larger x values the r.s.d. is constant.

The situation considered here is similar to that discussed by Boumans [19, 20]. He suggests that inductively-coupled plasmas applied in emission

TABLE 2

Variation of s.d. and r.s.d. (both in arbitrary units) with absorbance A if the precision of the transmittance measurement is considered independent of transmittance

A	0.01	0.02	0.05	0.10	0.15	0.20	0.30
S.d.	0.444	0.455	0.487	0.547	0.613	0.688	0.867
R.s.d.	44.4	22.7	9.75	5.47	4.09	3.44	2.89
A	0.40	0.50	0.70	1.00	1.20	1.50	
S.d.	1.09	1.37	2.18	4.34	6.88	13.7	
R.s.d.	2.73	2.75	3.11	4.34	5.74	9.16	

spectroscopy behave as fluctuation-noise limited systems. Measurements with such a plasma, therefore, also give a total standard deviation composed of a s.d. and a r.s.d. component. For this consideration a relation between σ_{rt} , σ_r , p , k (see below) and x was derived [20].

Regions for constant s.d. and constant r.s.d.

The assumption that the total standard deviation is composed of a s.d. and a r.s.d. component seems useful for discussions of the homogeneity of variances in calibration. The following discussion shows which is the upper limit of the region where the s.d. is approximately constant and which is the lower limit of the region where the r.s.d. is approximately constant. By "approximately constant" is meant that σ_t or σ_{rt} does not vary by more than a factor $(z - 1)$ where z is either 1.1 or 1.2 over the range of x values considered. Also $\sigma < \sigma_t < z\sigma$ and $\sigma_r < \sigma_{rt} < z\sigma_r$, with $z = 1.1$ or 1.2.

In the calculation the detection limit p may be introduced: $p = k\sigma$, where k depends on the risk accepted for drawing false conclusions (often, $k = 2$ or 3). The total standard deviation is then

$$\sigma_t = \left\{ \left(\frac{1}{k} \cdot p \right)^2 + x^2 \sigma_r^2 \right\}^{\frac{1}{2}} \text{ or } \sigma_{rt} = \left\{ \left(\frac{1}{k} \cdot \frac{p}{x} \right)^2 + \sigma_r^2 \right\}^{\frac{1}{2}} \quad (7)$$

The equation for σ_{rt} is equivalent to that given by Boumans [20]. From eqns. (7) it follows that:

$$\sigma < \sigma_t < z\sigma \text{ if } 0 < x < p (z^2 - 1)^{\frac{1}{2}} / k\sigma_r \text{ and } \sigma_r < \sigma_{rt} < z\sigma_r$$

$$\text{if } \infty > x > p/k (z^2 - 1)^{\frac{1}{2}}.$$

From these relations, limits for x can be calculated. Some examples for $k = 3$ are: $\sigma < \sigma_t < 1.1\sigma$ if $x < 0.153 p/\sigma_r$; $\sigma < \sigma_t < 1.2\sigma$ if $x < 0.221 p/\sigma_r$; $\sigma_r < \sigma_{rt} < 1.1\sigma_r$ if $x > 0.73 p/\sigma_r$; and $\sigma_r < \sigma_{rt} < 1.2\sigma_r$ if $x > 0.50 p/\sigma_r$.

One conclusion is that with $\sigma_r = 1\%$ and $z = 1.2$ (i.e. a variation of 20% in σ_t or σ_{rt}) σ can be considered as a constant if $x < 22p$ and σ_r as a constant if $x > 50p$. The region of x values where neither σ nor σ_r can be considered constant is only a factor $q = 50:22 = 2.3$. The value of q is independent of σ_r . If $z = 1.1$ (i.e. a variation of 10% in σ_t or σ_{rt}), $q = 73:15.3 = 4.8$. Actually q is proportional to $(z^2 - 1)$.

SELECTION OF A CALIBRATION RELATION

In selecting a calibration relation, it is seldom realized that an important difference exists between the application of eqn. (1) or (2) on the one hand and eqn. (3) on the other hand, if the variation of the s.d. with x is considered. In Fig. 1 (a, b), linear relationships hold. In Fig. 1(a) the s.d. of y represented by vertical bars has the same value for all values of x ; the conditions of homogeneity of variances is fulfilled. In Fig. 1(b) however, the s.d. increases regularly with y ; the r.s.d. is constant, but the condition required is not fulfilled. Figure 1(c) is a logarithmic plot and the vertical bars show a constant value for the s.d. of $\log y$. Therefore when eqn. (3) is applied, the homogeneity of variances is fulfilled in this instance. However, the situation that the s.d. of $\log y$ is independent of x implies that the r.s.d. of y is independent of x . This is the same situation as in Fig. 1(b) but now conditions for application of the usual type of calculation are fulfilled.

It is possible, therefore, to apply the usual method of calculation even if the r.s.d. is constant, but in this case the logarithmic eqn. (3) has to be used. A limitation of the latter equation, however, is that it cannot be used when $y \neq 0$ at $x = 0$. In such cases all y values measured can be diminished by the y value found for $x = 0$ ("blank"). This subtraction can be applied in many problems but it may suggest too high a precision in the results if the blank s.d. is too large. A difference between eqns. (2) and (3) is that eqn. (3) represents a non-linear relation between x and y in the general case that $b_3 \neq 1$. In this respect it seems somewhat equivalent to eqn. (4), and so it could be concluded that for non-linear relations between x and y , it is preferable to apply eqn. (3) if the r.s.d. of y is independent of x , and to apply eqn. (4) if the s.d. of y is independent of x .

Another important consideration in selecting eqn. (3) or (4) as a calibration equation is that the equation applied must represent correctly the true relationship between x and y . Some preliminary calculations [8] showed that if the curvature is only slight, the validity of one of these equations

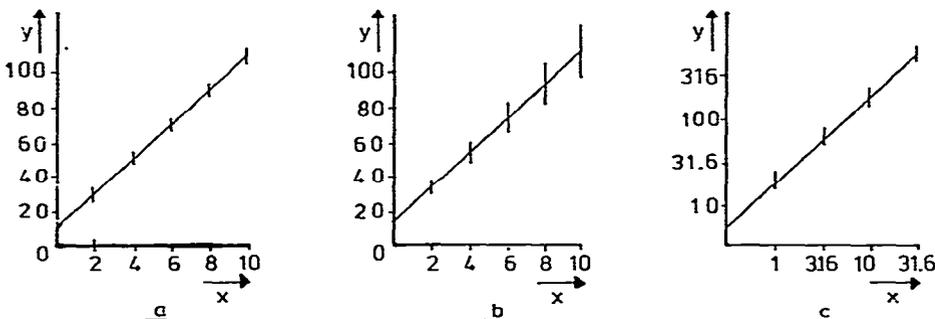


Fig. 1. Variation of the standard deviation in the signal y with the content x : (a) s.d. of y independent of x ; (b) s.d. of y proportional to y , r.s.d. of y independent of x ; (c) logarithmic plot with s.d. of $\log y$ independent of x , and r.s.d. of y independent of x .

implies that the other is also valid to a good approximation (of course after subtraction of any blank present). With large curvature, however, only one of the equations can be valid.

A relation that may be useful in the case of a constant r.s.d. of y if eqn. (3) does not apply, is eqn. (6), which was found satisfactory in some atomic absorption measurements [8]. For a linear relation between x and y and a constant r.s.d., eqn. (5) can be useful.

The conclusion that logarithmic calibration curves may be applied advantageously in situations where the r.s.d. is approximately independent of x , suggests that analytical instruments designed to perform the types of analyses where this situation can be expected, should be designed with appropriate calculation units.

Finally, it should be considered that for a straight-line plot (Fig. 1, a, b, c) a constant value of the s.d. of y or $\log y$ corresponds to a constant value of the s.d. of x or $\log x$. For a non-linear plot (eqn. 4 or 6), however, a constant value of s.d. or r.s.d. in y results in a gradually varying value of the s.d. or r.s.d. of x .

Errors in testing hypotheses

In calibration, statistical tests are sometimes applied to obtain information on the validity of assumptions or approximations. Such tests have been applied [5, 6] in testing the homogeneity of variances, in testing if a relation between x and y is curved, and in testing if the s.d. of the calibration procedure differs from the s.d. of the analysis of the real sample.

If hypotheses are tested, two types of errors can be made. First, the hypothesis is correct but it is assumed that it is false. This is an error of the first kind, and its probability is represented by α . Secondly, the hypothesis is incorrect but it is not rejected and is assumed to be correct (an error of the second kind with probability β). The maximum acceptable values for α and β should depend on the consequence of making a false decision. For α , a value of 0.10, 0.05 or 0.01 is often considered acceptable. Acceptable values for β are seldom given, however. This may be due partly to the fact that calculation of β is possible only for rather simple cases. An important reason why calculation of β is difficult is that it depends on the degree of incorrectness of the hypothesis, and this degree is difficult to describe in most cases.

An unwanted consequence of the difficulty of calculating β is that in many cases the existence of β is completely ignored. In such cases the hypothesis is considered correct if it cannot be proved that it is incorrect, even if β is large. This way of drawing statistical conclusions is comparable in some respects to drawing the analytical conclusion that a component is not present in a mixture because it cannot be detected, if one is not aware of the detection limit of the analytical method applied. Application of a different analytical method may give a positive result. Analogously, if it is considered necessary to test the hypothesis that homogeneity of variance exists, the result may depend on the type of statistical test applied [8]. It

is clear that other factors such as the number of observations also influence the probability that inhomogeneity will be detected. In the literature on analytical calibration no attention seems to have been given to this point.

A last point to be mentioned is that Gottschalk [6] applied a "heuristic" procedure. This can be demonstrated with an example: if eqn. (4) is applied and parameter c is found to be significantly different from zero, the conclusion may be that the application of the simpler eqn. (2) is not allowed. However, if c is so small that it may be neglected without significantly influencing the conclusions drawn about the analyte content of the sample, it may be preferable that eqn. (2) is applied. It seems that this is a good philosophy but good criteria for neglecting c must be given.

CONCLUSIONS

Calibration curves can be prepared in several ways. The simplest seems to be the graphical presentation, but very often mathematical procedures are applied, which can have different degrees of complexity. A discussion and comparison of graphical presentation and several mathematical procedures has been given by Kelly [21].

One limitation of the present paper is that only calibrations based on the method of least squares have been discussed. A second limitation is that the mathematical relations between x and y are rather simple. More complex relationships may be required in some practical situations. Some of these have been discussed recently by Schwartz [13]. However, the limitation of the present paper seem justified for at least two reasons: (a) the method of least squares is applied frequently in practice; (b) too frequently, analysts are not sufficiently aware of the assumptions and approximations made in their calculations and of the errors made in the final conclusions on the content of the unknown sample, if deviations arising from the assumptions made are too large.

The problem has been discussed only qualitatively here, and a more quantitative discussion is required. Some of the points to be included in such a more quantitative discussion, which will be discussed in further papers are the following: methods of detecting non-linearity of calibration curves and if possible a comparison of the "quality" of such methods; errors caused by the application of an improper calibration curve; dependence of the error on the degree of curvature and of the range of the calibration curve; methods of detecting inhomogeneity of variances and errors caused by unjustified assumptions of homogeneity; optimum location of calibration points on the x -axis if the number of observations is fixed.

REFERENCES

- 1 W. J. Dixon and F. J. Massey, Jr., *Introduction to Statistical Analysis*, 2nd edn., McGraw-Hill, New York, 1957.
- 2 W. J. Youden, *Statistical Methods for Chemists*, J. Wiley, New York, 1951.

- 3 K. Doerffel, *Beurteilung von Analyseverfahren und Ergebnissen*, Springer, Berlin, 1962.
- 4 V. V. Nalimov, *The Application of Mathematical Statistics to Chemical Analysis*, Pergamon, Oxford, 1963.
- 5 F. J. M. J. Maessen, J. W. Elgersma and P. W. J. M. Boumans, *Spectrochim. Acta*, Part B, 31 (1976) 179.
- 6 G. Gottschalk, *Fresenius Z. Anal. Chem.*, 276 (1975) 81.
- 7 G. Gottschalk, *Fresenius Z. Anal. Chem.*, 275 (1975) 1: 276 (1975) 257; 278 (1976) 1: 280 (1976) 205; 282 (1976) 1.
- 8 J. Agterdenbos, Unpublished work presented at XXth C.S.I., 7 I.C.A.S., Strbske Pleso, 1977 and Euroanalysis III, Dublin 1978.
- 9 C. T. Shewell, *Anal. Chem.*, 32 (1960) 1535.
- 10 J. Mandel and F. J. Linnig, *Anal. Chem.*, 29 (1957) 743.
- 11 J. Mandel, *The Statistical Analysis of Experimental Data*, Interscience, New York, 1964.
- 12 J. P. Franke and R. A. de Zeeuw, *Anal. Chem.*, 50 (1978) 1374.
- 13 L. M. Schwartz, *Anal. Chem.*, 49 (1977) 2062.
- 14 H. C. Wagenaar, I. Novotny and L. de Galan, *Spectrochim. Acta*, Part B, 29 (1974) 301.
- 15 J. Agterdenbos and J. Vlogtman, *Talanta*, 21 (1974) 231.
- 16 L. M. Schwartz, *Anal. Chem.*, 48 (1976) 2287.
- 17 A. Ringbom, *Fresenius Z. Anal. Chem.*, 115 (1938) 332.
- 18 J. Agterdenbos, *Fresenius Z. Anal. Chem.*, 154 (1957) 401; 157 (1957) 161: 159 (1958) 202.
- 19 P. W. J. M. Boumans, *Spectrochim. Acta*, Part B, 31 (1976) 147.
- 20 P. W. J. M. Boumans, *Fresenius Z. Anal. Chem.*, in press
- 21 P. C. Kelly, *Anal. Chem.*, 44 (1972) 28A.