# INTERNATIONAL STANDARD

# ISO 7066-1

First edition
1989-10-01

## Assessment of uncertainty in the calibration and use of flow measurement devices —

## Part 1 :
## Linear calibration relationships



*Évaluation de l'incertitude dans l'étalonnage et l'utilisation des appareils de mesure du débit —*

*Partie 1 : Relations d'étalonnage linéaires*



Reference number
ISO 7066-1 : 1989 (E)

## Contents







International Organization for Standardization
Case postale 56 • CH-1211 Genève 20 • Switzerland
Printed in Switzerland

## Foreword

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

Draft International Standards adopted by the technical committees are circulated to the member bodies for approval before their acceptance as International Standards by the ISO Council. They are approved in accordance with ISO procedures requiring at least 75 % approval by the member bodies voting.



International Standard ISO 7066-1 was prepared by Technical Committee ISO/TC 30, *Measurement of fluid flow in closed conduits.*



ISO 7066 consists of the following parts, under the general title *Assessment of uncertainty in the calibration and use of flow measurement devices*:

— *Part 1: Linear calibration relationships*

— *Part 2: Non-linear calibration relationships*

Annexes A, B and C form an integral part of this part of ISO 7066. Annexes D, E, F and G are for information only.

## Introduction

This International Standard has been drawn up according to the principles outlined in ISO 5168 and gives guidance on how the uncertainty in a calibration curve or in the mean of a number of measurements of the same flow-rate may be calculated. To achieve this it is assumed that the uncertainty in each individual measurement of flow-rate is calculated in accordance with ISO 5168.

This part of ISO 7066 deals only with calibration graphs which are linear or which can be linearized. ISO 7066-2 deals with non-linear calibration graphs.

# Assessment of uncertainty in the calibration and use of flow measurement devices —

## Part 1: Linear calibration relationships

## 1 Scope

This International Standard deals with methods of assessing the uncertainty in the calibration of any method of measuring flow-rate, either in closed conduits or in open channels. It also deals with the estimation of the uncertainty in one or more measurements which use the resulting calibration graph.

Only linear relations are considered in this part of ISO 7066; the uncertainty in non-linear relations is the subject of ISO 7066-2. Where a calibration curve is not linear, this part of ISO 7066 is therefore applicable only if

a) the variables may be transformed (for example by taking logarithms) to create a linear relationship between them;

b) the range over which the relationship is established may be subdivided in such a way that one variable varies linearly with the other within each subdivision; or

c) systematic deviations from linearity of the calibration graph are negligible in comparison with the uncertainty associated with the individual points forming the graph[1)].

Although it is assumed that the uncertainty in the independent and dependent variables for which the calibration graph is constructed is normally established prior to determining the calibration graph, consideration is given in 5.3 to how these uncertainties may sometimes be determined during the calibration procedure itself, when the uncertainty in an individual calibration point is not known.

For most of the calculations given in this part of ISO 7066, computer programs exist which are generally referred to in program libraries as "linear regression methods" or "linear curve fitting".

Examples are given in annexes A and B of how the principles in this part of ISO 7066 may be applied.

## 2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO 7066. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO 7066 are encouraged to investigate the possibility of applying the most recent editions of the standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards.



ISO 1100-2 : 1982, *Liquid flow measurement in open channels — Part 2 : Determination of the stage-discharge relation.*



ISO 5168 : 1978, *Measurement of fluid flow — Estimation of uncertainty of a flow-rate measurement.*

## 3 Symbols and definitions

The symbols and definitions used in this part of ISO 7066 have been taken from ISO 772 and ISO 4006.

The definitions given in 3.3 are only for terms used in some special sense or for terms the meaning of which it seems useful to emphasize.

### 3.1 Symbols

$a$ intercept of the calibration graph on the ordinate

$b$ gradient of the calibration graph

$C$ discharge coefficient

$d$ diameter of the orifice in an orifice plate flow-meter

$D$ diameter of the pipe

1) For example, a turbine meter calibration graph may have a minimum value after the "hump" before rising asymptotically to become horizontal, but the linear calibration range is often assumed to extend down to the flow-rate at which the extrapolation of the horizontal part of the graph intercepts the graph as it rises towards the maximum peak.

$e(\ )$ uncertainty of variable contained in parentheses[1)]

$e_r(\ )$ random uncertainty of variable contained in parentheses[1)]

$e_s(\ )$ systematic uncertainty of variable contained in parentheses[1)]

$K$ calibration coefficient

$M$ number of repetitions of a measurement of flow-rate

$n$ number of measurement points used to establish calibration graph

$N_p$ number of pulses generated by a turbine meter per second

$Q$ flow-rate

$Re_d$ Reynolds number based on bore diameter

$s(\ )$ experimental standard deviation of variable contained in parentheses

$s_R$ standard deviation of points about the best straight line [see equation (17)]

$s(x, y)$ covariance of $x$ and $y$ [see equation (10)]

$t$ Student's $t$

$x$ independent variable

$y$ dependent variable

$\hat{y}$ the value of the dependent variable predicted by the calibration graph

$\nu$ number of degrees of freedom

## 3.2 Subscripts and superscripts

$i$ $i$th value of a variable

$k$ a specific value of a variable

— arithmetic mean value of a variable

^ the value of the variable predicted by an equation of a fitted curve

## 3.3 Definitions

**3.3.1 (absolute) error of measurement:** The result of a measurement minus the (conventional) true value of the measurand.

NOTES

1 The term relates equally to

— the indication,

— the uncorrected result,

— the corrected result.

2 The known parts of the error of measurement may be compensated by applying appropriate corrections. The error of the corrected result can only be characterized by an uncertainty.

3 The "absolute error", which has a sign, should not be confused with the absolute value of an error which is the modulus of an error.

**3.3.2 random error:** Component of the error of measurement which, in the course of a number of measurements of the same measurand, varies in an unpredictable way.

NOTE — It is not possible to correct for random error.

**3.3.3 systematic error:** Component of the error of measurement which, in the course of a number of measurements of the same measurand, remains constant or varies in a predictable way.

NOTE — Systematic errors and their causes may be known or unknown.

**3.3.4 spurious errors:** Errors which invalidate a measurement. They generally have a single cause such as the incorrect recording of one or more significant digits or malfunction of instruments.

**3.3.5 uncertainty:** An estimate characterizing the range of values within which the true value of a measurand lies.

**3.3.6 random uncertainty:** Component of uncertainty associated with a random error. Its effect on mean values can be reduced by taking many measurements.

**3.3.7 systematic uncertainty:** Component of uncertainty associated with a systematic error. Its effect cannot be reduced by taking many measurements.



**3.3.8 experimental standard deviation:** For a series of $n$ measurements of the same measurand, the parameter $s$ characterizing the dispersion of the results and given by the formula

$$s = \left[ \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1} \right]^{1/2}$$

where

$x_i$ is the result of the $i$th measurement;

$\bar{x}$ is the arithmetic mean of the $n$ results considered.

NOTES

1 The experimental standard deviation should not be confused with the population standard deviation $\sigma$ of a population of size $N$ and of mean $m$, given by the formula

$$\sigma = \left[ \frac{\sum_{i=1}^{N} (x_i - m)^2}{N} \right]^{1/2}$$

2 If the series of $n$ measurements is considered to be a sample of a population, $s$ is an estimate of the population standard deviation.

1) In some International Standards the symbols $U$ and $E$ have been used instead of $e$.

**3.3.9 variance**: The square of the standard deviation.

**3.3.10 confidence limits**: The lower and upper limits within which the true value is expected to lie, with a specified probability assuming negligible systematic error.

**3.3.11 calibration graph**: Locus of points obtained by plotting some index of the response of a flow-meter against some function of the flow-rate.

## 4 General

For a calibration to be meaningful, the systematic uncertainty in the calibrator shall be very much less than the systematic uncertainty in the device or system being calibrated. This is especially true when the procedures specified in 7.3 are used.

The calibration of a flow-metering device or system will result in a graph of the calibration coefficient which will subsequently be used to predict the flow-rate. As this subsequent flow-rate prediction has to have an uncertainty attached to it, then not only the functional relationship between calibration coefficient and flow-rate but also the uncertainty in the calibration coefficient shall be established during calibration.

There will exist a number of pairs of values $(x, y)$ where the uncertainties in $x$ and $y$ [$e(x)$ and $e(y)$ respectively[1)]] are known from one of the methods given in clause 5. The choice of the procedure by which the coefficients and the uncertainty of the calibration equation are calculated is determined by the relative magnitudes of the random components of the uncertainties $e_r(x)$ and $e_r(y)$, as described in clause 7.

When $e_r(x)$ can be ignored (as, for example, is normally the case in the calibration of an orifice plate), the calibration equation and the uncertainty in the calibration coefficient are computed by the methods specified in 7.2 and 9.3 respectively. When, however, the random uncertainties in $x$ and $y$ are of similar magnitude, the methods specified in 7.3 and 9.4 should be used. When the uncertainties in $x$ and $y$ are both significant but cannot be regarded as approximately equal, then the calculation of the uncertainty in the calibration graph is outside the scope of this part of ISO 7066.

A special case is that where $y$ is effectively independent of $x$; this is a common situation with flow-meters used in closed pipes, since there is an obvious advantage in having a calibration coefficient which is independent of flow-rate. In such cases, the method specified in 9.2 may be used for calculating the uncertainty.

In addition to determining the uncertainty in the coefficient or curve obtained during the calibration of a flow-meter or gauging station, it is necessary to determine the uncertainty in the particular value which is used as a coefficient or is read from the calibration curve when the flow-meter is used after having been calibrated. Where the value of the calibration coefficient to be used is determined completely independently of the measurement from which a flow-rate is to be obtained, then these two quantities are the same, provided that the conditions of use are identical with those of the calibration; if, however, some information from the test to measure the flow-rate is required before the calibration coefficient or curve can be used, then an additional uncertainty will be introduced. Annex C describes how this additional uncertainty is introduced.

The approaches to be used in these different circumstances are described in clause 9, and in clause 11 methods for assessing the uncertainty in the average of a number of measurements are described.

## 5 Uncertainties in individual calibration points

### 5.1 General

When a flow-meter is being calibrated, some function of its output may be plotted against either a reference measurement of the flow-rate or some function of this flow-rate, such as the Reynolds number.

In either case, it is necessary to establish the uncertainty in the coordinates of a single point in order to be able to compute the uncertainty in the calibration graph, and there are two ways in which this may be done:

— ISO 5168 may be used; or

— the information can sometimes be obtained from the calibration data directly.

It should be noted, however, that the uncertainty in a coordinate may vary with the value of the coordinate itself; thus, for example, where the reference flow-rate is measured by a diversion system involving the static weighing of a quantity of liquid collected over a measured period of time, the uncertainty due to the timing is usually less for long diversion periods (and consequently for low flow-rates if approximately the same weight of water is collected at each test point) than for short diversion periods.



1) The customary categories of independent and dependent variables, and of abscissae (horizontal) and ordinate (vertical) coordinates in a graph, are irrelevant for linear regressions in that the important distinction here is between variables that have significant uncertainties and variables that have negligible (or zero) uncertainties. When the uncertainty in one variable is significantly greater, in the manner described in clause 7, than the other, the former will be denoted by $y$ and the latter by $x$. Thus the regressions studied are all for $y$ on $x$ irrespective of whether a variable is considered to be independent, and irrespective of which variable is plotted "horizontally".

Justification shall always be provided where it is assumed that the uncertainty is constant throughout the range of a coordinate. Where the uncertainty cannot be regarded as constant, it shall be estimated for sufficient values of the coordinate to give a clear idea of how it varies with the value of the coordinate.

## 5.2 Use of ISO 5168

ISO 5168 describes in detail how an estimation of the uncertainty in a single measurement of flow-rate may be arrived at, and the procedures described in ISO 5168 may be used to calculate the uncertainties of both the independent and the dependent variables in the calibration graph.

Is it important, however, that the random and systematic contributions to the uncertainty are calculated separately and not combined; the various formulae for calculating the uncertainty in a calibration graph are first used to calculate only the random component of the uncertainty, and so only the random component of the uncertainty in the individual points is used at that stage. Subsequently, the systematic uncertainty in the individual points is added by the root-sum-square method to the value obtained to give the final combined value. The random component of the uncertainty is required separately in any case, since it is the relative magnitudes of random uncertainties in $x$ and $y$ which determine the calculation procedures to be used.

This method of calculating the uncertainty in the coordinates of a single experimental point may be used only when previous investigations have been carried out to establish the uncertainties in the various subsidiary measurements which have to be made, or when this information is available from some other source.

## 5.3 Use of calibration data

Where the measurement conditions can be kept constant, it is possible to determine the random components of the uncertainties of the coordinates of the calibration graph during the calibration by repeating measurements. Thus, for example in the calibration of an orifice plate, it might be possible to keep the Reynolds number constant and to take a series of readings of the data necessary to compute the calibration coefficient, and it might conversely be possible to keep the differential pressure across the orifice plate constant, while making a number of determinations of the reference flow-rate (and consequently the Reynolds number).

The uncertainty (due to random effects) in the calibration coefficient or Reynolds number may then be calculated from the standard deviations of the resulting measurements.

There is no alternative to using the methods of ISO 5168 in assessing systematic uncertainties if the assessment of the uncertainty is to be in accordance with this part of ISO 7066.

# 6 Linearity of the calibration graph

## 6.1 General

In considering the shape of any calibration graph, previous knowledge or supporting information appropriate to the flow-metering method being used (for example if there is a relevant published standard) should always be taken into account to justify any assumption made about the shape of the graph. Where the shape of the calibration graph is unusual, a sufficient number of calibration points shall be repeated to verify the shape for that particular flow-meter.

In the absence of such information, or where there is no reason to believe in advance that the form of the curve should be linear, simple techniques are available to establish whether or not the graph can be treated as linear, but these are applicable only when the experimental points are not grouped into sets. When the experimental points are grouped into sets, only visual observations may be used, since the extensive statistical tests which are otherwise required are outside the scope of this part of ISO 7066.

When the data can be grouped into sets, the following test may be used, but the sets shall be such that each one consists of a number of measurements made at one of a number of fixed values for one of the coordinates. Figure 1 illustrates a case where there are five sets, the number of measurements within the sets varying from four to six.

The test consists of comparing the variance of the means of the groups about the fitted straight line with the variance within the groups.



The variance within the groups, $s_g^2$, is given by

$$s_g^2 = \frac{\sum_{i=1}^{q} \sum_{j=1}^{n_i} (y_{i,j} - \bar{y}_i)^2}{n - q} \qquad \ldots (1)$$

where

$n$ is the total number of measurements;

$n_i$ is the number of measurements in the $i$th group;

$q$ is the number of groups;

$y_{i,j}$ is the $j$th measurement in the $i$th group;

$$\bar{y}_i = \frac{\sum_{j=1}^{n_i} y_{i,j}}{n_i}$$

The variance of the means of the group about the fitted straight line, $s_m^2$, is given by

$$s_m^2 = \frac{\sum_{i=1}^{q} n_i (\hat{y}_i - \bar{y}_i)^2}{q - 2} \qquad \ldots (2)$$

where

$n_i$ and $q$ have the same meaning as in equation (1);

$\hat{y}_i$ is the value of $y$ obtained from the fitted straight line for a given value of $x$.

The test quotient is $s_m^2/s_g^2$, and if this number equals or exceeds the value given in table 1 for $\nu_1 = q - 2$ and $\nu_2 = n - q$ degrees of freedom, then the best-fit equation through the data points cannot be assumed to be linear. If, however, the test quotient is less than the corresponding value given in table 1, the best-fit equation may be assumed to be linear at the 95 % confidence level.

If the relationship between the calibration coefficient and the independent variable is not linear, there are two possible ways in which the data may be made suitable for analysis in accordance with this part of ISO 7066. The first, to linearize the curve, may, however, only be done on the basis of some physical model.

## 6.2 Linearization of curve

The coordinates may be transformed to give two new variables which, when plotted one against the other, produce a straight line, or a function of one of the coordinates may be used instead of the coordinate itself so as to produce this result, although it is of course essential that the transformation be capable of being used easily when the resulting graph is subsequently applied to the use of the flow-meter or the flow measurement method.

**Table 1 — Values of the $F$ distribution for selected degrees of freedom — Probability level 0,05**

| $\nu_2$ \ $\nu_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 20 | 30 | 40 | 80 | 100 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 161,44 | 200 | 216 | 225 | 230 | 234 | 237 | 239 | 241 | 242 | 248 | 250 | 251 | 252 | 253 |
| 2 | 18,51 | 19 | 19,2 | 19,2 | 19,3 | 19,3 | 19,4 | 19,4 | 19,4 | 19,4 | 19,4 | 19,5 | 19,5 | 19,5 | 19,5 |
| 3 | 10,13 | 9,55 | 9,28 | 9,12 | 9,01 | 8,94 | 8,89 | 8,85 | 8,81 | 8,79 | 8,66 | 8,62 | 8,59 | 8,56 | 8,55 |
| 4 | 7,71 | 6,94 | 6,59 | 6,39 | 6,26 | 6,16 | 6,09 | 6,04 | 6 | 5,96 | 5,8 | 5,75 | 5,72 | 5,67 | 5,66 |
| 5 | 6,61 | 5,79 | 5,41 | 5,19 | 5,05 | 4,95 | 4,88 | 4,82 | 4,77 | 4,74 | 4,56 | 4,5 | 4,46 | 4,41 | 4,41 |
| 6 | 5,99 | 5,14 | 4,76 | 4,53 | 4,39 | 4,28 | 4,21 | 4,15 | 4,1 | 4,06 | 3,87 | 3,81 | 3,77 | 3,72 | 3,71 |
| 7 | 5,59 | 4,74 | 4,35 | 4,12 | 3,97 | 3,87 | 3,79 | 3,73 | 3,68 | 3,64 | 3,44 | 3,38 | 3,34 | 3,29 | 3,27 |
| 8 | 5,32 | 4,46 | 4,07 | 3,84 | 3,69 | 3,58 | 3,5 | 3,44 | 3,39 | 3,35 | 3,15 | 3,08 | 3,04 | 2,99 | 2,97 |
| 9 | 5,12 | 4,26 | 3,86 | 3,63 | 3,48 | 3,37 | 3,29 | 3,23 | 3,18 | 3,14 | 2,94 | 2,86 | 2,83 | 2,77 | 2,76 |
| 10 | 4,96 | 4,1 | 3,71 | 3,48 | 3,33 | 3,22 | 3,14 | 3,07 | 3,02 | 2,98 | 2,77 | 2,7 | 2,66 | 2,6 | 2,59 |
| 20 | 4,35 | 3,49 | 3,1 | 2,87 | 2,71 | 2,6 | 2,51 | 2,45 | 2,39 | 2,35 | 2,12 | 2,04 | 1,99 | 1,92 | 1,91 |
| 30 | 4,17 | 3,32 | 2,92 | 2,69 | 2,53 | 2,42 | 2,33 | 2,27 | 2,21 | 2,16 | 1,93 | 1,84 | 1,79 | 1,71 | 1,7 |
| 40 | 4,08 | 3,23 | 2,84 | 2,61 | 2,45 | 2,34 | 2,25 | 2,18 | 2,12 | 2,08 | 1,84 | 1,74 | 1,69 | 1,61 | 1,59 |
| 80 | 3,96 | 3,11 | 2,72 | 2,49 | 2,33 | 2,21 | 2,13 | 2,06 | 2 | 1,95 | 1,7 | 1,6 | 1,54 | 1,45 | 1,43 |
| 100 | 3,94 | 3,09 | 2,7 | 2,46 | 2,31 | 2,19 | 2,1 | 2,03 | 1,97 | 1,93 | 1,68 | 1,57 | 1,52 | 1,41 | 1,39 |



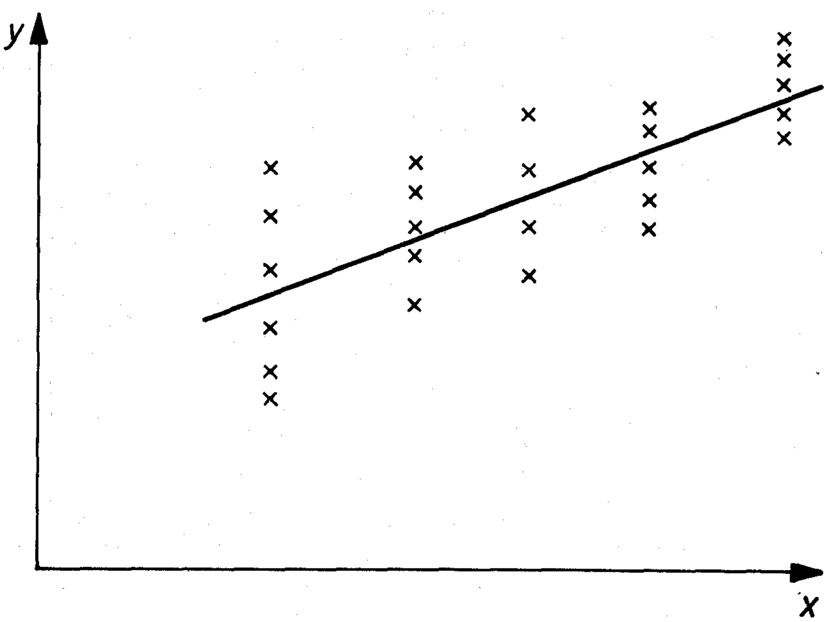


**Figure 1 — Example of grouping data to establish linearity of best-fit straight line**

An example of the first possibility would be to plot "log $y$" against "log $x$" where the basic relation is of the form "$y = ax^2$" ("$a$" being a constant).

The second method of linearizing a curve is commonly used when calibrating an orifice plate. In this case, it is known from experience that it is better to plot the discharge coefficient $C$ against the Reynolds number than to plot the differential pressure against the Reynolds number, since this gives a graph which is linear over a fairly wide range. The graph is, however, non-linear at low Reynolds numbers, and it has been found from experience that plotting the discharge coefficient against some function of the Reynolds number (for example $Re_d^{-0,5}$ or $Re_d^{-0,75}$) extends the linear range.

### 6.3 Subdivision of the curve

Although a calibration graph may not be linear over the full range of the calibration, it may well be that by subdividing the curve into a number of parts, each part can be regarded as linear; in this case, a separate calibration equation and confidence limits shall be calculated for each portion of the graph. Where possible, the number of points in each sub-section of the graph should be such that they give approximately the same uncertainties for the line through each sub-section. Since calibration graphs should never be extrapolated beyond the extreme data points unless there is extremely good reason for doing so, there shall be at least three points common to adjacent portions of the calibration graph.

## 7 Fitting the best straight line

### 7.1 General

Before a best straight line and its associated uncertainty are calculated, the data available shall first be examined, since they can fall into one of several classes, and different formulae apply for different types of data. Some basic principles are illustrated in figure 2.

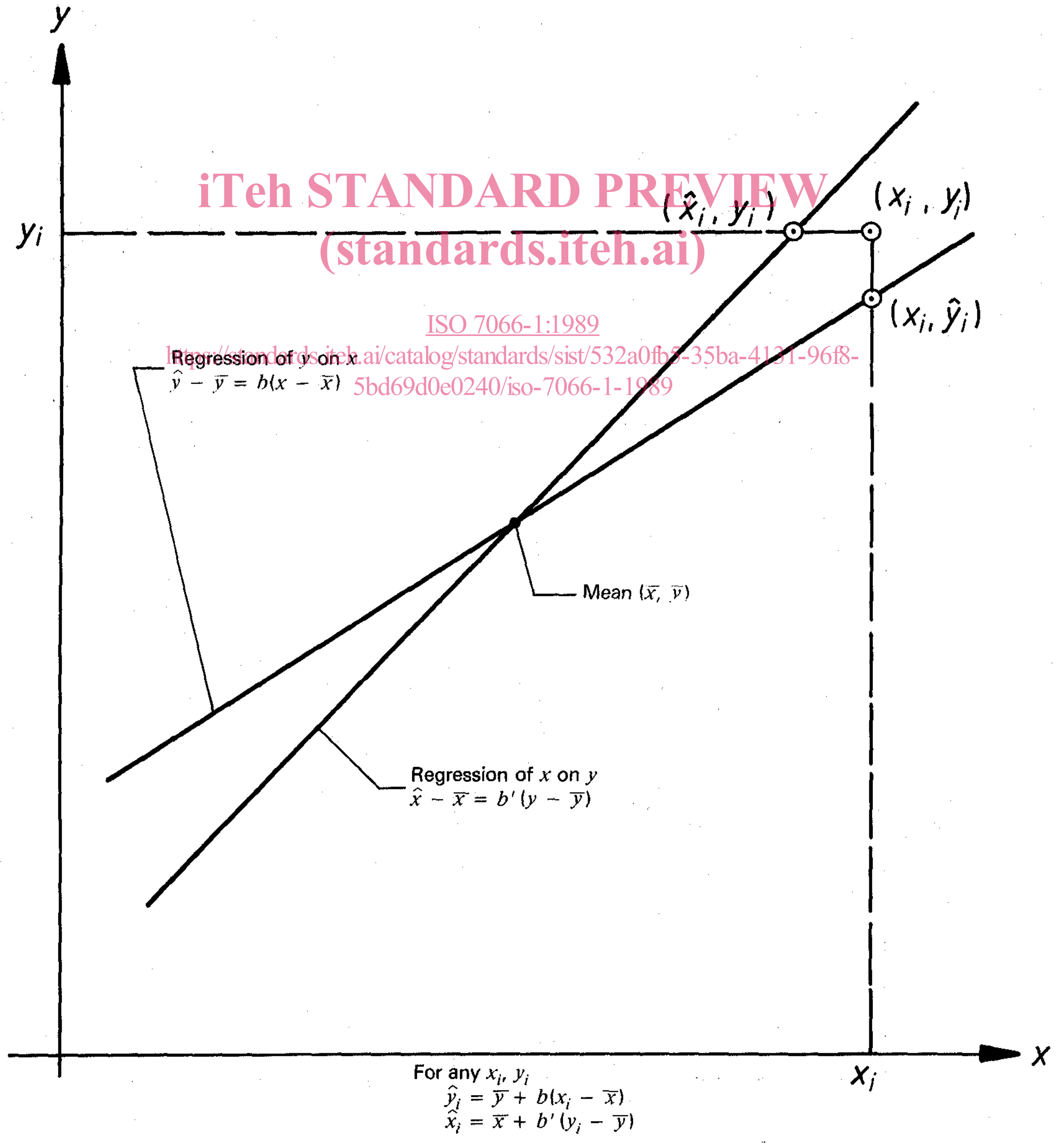




**Figure 2 — Basic principles for best-fit straight lines**

The best straight line fit to a sample of $n$ points, $(x_i, y_i)$, when $1 \leqslant i \leqslant n$, is given by the regression of $y$ on $x$:

$$\hat{y}_i - \bar{y} = b(x_i - \bar{x}) \qquad \ldots (3)$$

where

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \ldots (4)$$

$\hat{y}_i$ is the value of $y$ on the line for a measured value $x_i$;

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} \qquad \ldots (5)$$

Equation (3) may also be written as

$$\hat{y}_i = a + bx_i \qquad \ldots (6)$$

where

$$a = \bar{y} - b\bar{x} \qquad \ldots (7)$$

The method to be used for computing the values of the coefficients $a$ and $b$ depends on the magnitude of the random uncertainties in $x$ and $y$.

The most common case occurs when the random uncertainties in $x$ and $y$ are both significantly different from zero. Fortunately, if both variables have random uncertainties significantly different from zero, it is normally possible to assume that

a) one random uncertainty is negligible in comparison with the other, or

b) these random uncertainties are of approximately equal magnitude.

The former assumption leads to the same formulae as for the cases where only one variable has a random uncertainty significantly different from zero; this is the situation which normally applies.

To assess the relative magnitude of the random uncertainties in $x$ and $y$, first calculate $e_r(x)$ and $e_r(y)$ according to the principles of ISO 5168.

Obtain an approximate value for the gradient $b$ from a graph or from equation (8) or from equation (11). If the absolute value of $be_r(x)$ is less than approximately one-fifth of $e_r(y)$, the formulae given in 7.2 shall be used to fit a straight line; if not, the formulae given in 7.3 shall be used.

## 7.2 Random uncertainty negligible or small in one variable in comparison with the other

The gradient of the line is, in this case, given by

$$b = \frac{s(x, y)}{s^2(x)} \qquad \ldots (8)$$

where

$$s^2(x) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \qquad \ldots (9)$$

$$s(x, y) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \qquad \ldots (10)$$

The intercept, $a$, is given by equation (7) :

$$a = \bar{y} - b\bar{x}$$



## 7.3 Random uncertainty in both variables of similar magnitude

The gradient is calculated from

$$b = \pm \left[ \frac{s^2(y)}{s^2(x)} \right]^{1/2} \qquad \ldots (11)$$

where

$$s^2(y) = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2 \qquad \ldots (12)$$

$s^2(x)$ is given by equation (9);

the intercept, $a$, is again given by equation (7).

The sign of $b$ is the same as that of $s(x, y)$.

NOTE — The following alternative formulae for $s^2(x)$, $s^2(y)$ and $s(x, y)$ are easier to use, if the computation is carried out manually, but great care should be taken since they are prone to rounding errors:

$$s^2(x) = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right]$$

$$s^2(y) = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right]$$

$$s(x, y) = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)\right]$$

## 8 Detection of outliers

Spurious errors (see ISO 5168) are errors, such as human errors or instrument malfunction, which invalidate a measurement; they may be due to, for example, the transposing of numbers in recording data or the presence of pockets of air in leads from a water line to a manometer. Such errors cannot be incorporated into any statistical analysis and the measurement shall be discarded. Where the error is not large enough to make the result obviously invalid, some rejection criterion should be applied to decide whether the data point should be rejected or retained.

Whenever it is suspected that one or more results have been affected by errors of this nature, a statistical "outlier" test should be applied. For the purposes of this part of ISO 7066, either the Dixon test or the Grubbs extreme deviation outlier test may be used. The Dixon test is easy to use by hand, but when a set of values are being processed by computer, the Grubbs test is more suitable, since it is more reliable, easier to program and takes up less storage.

Details of these tests are given in annex E.

## 9 Uncertainty of calibration

### 9.1 General

In general, the uncertainty in the best straight line arises from uncertainties in the values for the intercept, $a$, and for the gradient, $b$.

From equation (3)

$$\hat{y}_i = \bar{y} + b(x_i - \bar{x}) \qquad \dots \quad (13)$$

Thus, combining the variances of $\bar{y}$, $\bar{x}$ and $b$ by the root-sum-square method (see ISO 5168),

$$s^2(\hat{y}) = s^2(\bar{y}) + b^2 s^2(\bar{x}) + (x_k - \bar{x})^2 s^2(b) \qquad \dots \quad (14)$$

where $x_k$ is the value of $x$ at which the uncertainty in $\hat{y}$ is required.

Thus $s^2(\bar{y})$ is the variance of $\bar{y}$ which would be obtained from the scatter of several different determinations of $\bar{y}$ obtained using the same range and values of $x$, and with the same number of measurements of $y$, and is given by

$$s^2(\bar{y}) = \frac{\sum_{i=1}^{n} (\bar{y}_i - \bar{y})^2}{n-1}$$

$s^2(\bar{x})$ is defined similarly.

Note that $s^2(\bar{x})$ and $s^2(\bar{y})$ are quite different from $s^2(x)$ and $s^2(y)$ defined in equations (9) and (12). $s^2(y)$, for example, is the variance of all the values of $y$ over the range of the calibration graph relative to their mean value, and $s^2(x)$ has a similar meaning, whereas $s^2(\bar{y})$ is associated only with the scatter of different determinations of $\bar{y}$ and is essentially an indication of the random uncertainty in $y$.

From equation (7), it can be seen that the variance of $a$ is given by

$$s^2(\bar{y}) + b^2 s^2(\bar{x})$$

and the contribution of the variance in $b$ to the variance in $\hat{y}$ is

$$(x_k - \bar{x})^2 s^2(b)$$

Thus the random component of the uncertainty in $\hat{y}$, $e_r(\hat{y})$, is given at the 95 % confidence level by

$$e_r(\hat{y}) = \pm t\,[s^2(\bar{y}) + b^2 s^2(\bar{x}) + (x_k - \bar{x})^2 s^2(b)]^{1/2} \qquad \dots \quad (15)$$

where $t$ is obtained from table 2 for $n - 2$ degrees of freedom.

**Table 2 — Value of Student's $t$ at 95 % confidence level**

| Number of degrees of freedom, $\nu$ | $t$ |
|---|---|
| 1 | 12,7 |
| 2 | 4,3 |
| 3 | 3,2 |
| 4 | 2,8 |
| 5 | 2,6 |
| 6 | 2,4 |
| 7 | 2,4 |
| 10 | 2,2 |
| 15 | 2,1 |
| 20 | 2,1 |
| 30 | 2 |
| 60 | 2 |
| ∞ | 1,96 |

Equation (15) is the basic equation for uncertainty which is used either directly or in a modified form in the various paragraphs below.

As noted in clause 4, the method to be used for calculating the uncertainty in a calibration depends on the magnitude of the random uncertainty in the variable $x$, and on whether or not the value of $y$ is independent of the value of $x$, that is on whether or not $b$ is zero. The first step is therefore to establish whether or not the gradient, $b$, of the line is significantly different from zero, and this is carried out as follows.

The standard deviation of $b$, $s(b)$, shall be calculated from

$$s^2(b) = \frac{s_R^2}{(n-1)s^2(x)} \qquad \ldots \quad (16)$$

where $s_R$ is the standard deviation of the points about the best straight line, i.e.

$$s_R = \left[ \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-2} \right]^{1/2} \qquad \ldots \quad (17)$$

An alternative formula for $s(b)$ is

$$s(b) = \left[ \frac{s^2(y)s^2(x) - s^2(x,y)}{(n-2)s^4(x)} \right]^{1/2} \qquad \ldots \quad (18)$$

This is more convenient, but the calculation of the numerator is prone to rounding errors, so it is important to ensure that enough significant figures are used in the computation.

The 95 % confidence limits for $b$ are then given by

$$\left.\begin{array}{l} b - ts(b) \\ b + ts(b) \end{array}\right\} \qquad \ldots \quad (19)$$

where $t$ is the value obtained from table 2 for $n - 2$ degrees of freedom.

If these limits include zero and there is independent evidence that the particular flow-meter or type of flow-meter is expected to have a constant coefficient (for example from the information in a relevant standard or a previous calibration), it can be assumed that the calibration coefficient has a constant value. If there is no independent evidence that a constant calibration coefficient is expected, then, whether or not the limits include zero, the formulae given in 9.3 (if the uncertainty in $x$ can be ignored) or in 9.4 shall be used.

In order to calculate the uncertainty in the calibration coefficient the random and systematic components shall first be evaluated separately. Random components are calculated from the formulae given in 9.2 to 9.4 and shall include the contributions from the random uncertainties in the instruments used during the calibration, so that these do not have to be allowed for separately. In particular, hysteresis effects in instruments will contribute to the random uncertainty component obtained in this way. Systematic components are calculated in accordance with ISO 5168.

NOTE — Equation (16) is strictly valid only when the random uncertainty in $x$ contributes significantly less to the calibration graph uncertainty than the random uncertainty in $y$ does, since it is a simplification of a more general formula. The expansion of equation (17), used to develop equation (18), also makes this assumption. The use of equations (18) and (19) will, however, give a smaller uncertainty for $b$ than the full formulae would, and so if they include zero, the more general formula would lead to the same conclusion. If they do not include zero, it is possible that the gradient will be treated as having a non-zero value, when it might have been acceptable to take its value as zero, but this would be a rare situation, and the only penalty incurred would be that the more complicated formulae given in 9.4 would be used instead of the simpler method given in 9.2. The results obtained for the uncertainty of the best-fit line would in such a case be virtually identical no matter whether the formulae given in 9.2 or 9.4 were used.

## 9.2 Calibration curve with zero gradient

In this case, the calibration coefficient has a single value, and all of the estimates of this value can be analysed together, irrespective of the value of the independent variable to which they correspond. An example of where this often occurs is in the calibration of a turbine meter for use in water.

The best estimate of the value of the calibration coefficient is given by equation (5):



$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$



In this case, $b = s(b) = 0$, and so, from equation (15), the random component of the uncertainty in $\hat{y}$ is given at the 95 % confidence level by

$$e_r(\hat{y}) = \pm\, ts(\bar{y}) = \pm\, t \frac{s(y)}{n^{1/2}} \qquad \ldots \quad (20)$$

where

$$s(y) = \left[ \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1} \right]^{1/2} \qquad \ldots \quad (21)$$

and $t$ is obtained from table 2 for $n - 1$ degrees of freedom.

The systematic component of the uncertainty in $\hat{y}$, $e_s(\hat{y})$, is calculated as described in clause 5 and the uncertainty in the calibration coefficient, $e(\hat{y}_c)$, is then given by

$$e(\hat{y}_c) = [e_r^2(\hat{y}) + e_s^2(\hat{y})]^{1/2} \qquad \ldots \quad (22)$$