Table 12 — Error comparisons of examples one and two

	Example One — Facility	*Example Two — Facility*
1. Experimental standard deviation, kg_m/sec (s)	0.078 7	0.076 2
2. Degrees of freedom (v)	>30	>30
3. Systematic error, kg_m/sec (B)	0.245 7	0
4. Uncertainty, kg_m/sec	0.40	0.22

Annex B — Examples on estimating uncertainty in open channel flow measurement

B.1 General

Evaluation of the overall uncertainty of a flow in an open channel will be demonstrated by considering (1) the velocity-area method and (2) the weirs method.

The method of measuring the flow is such that it is impractical to eliminate interdependent variables from the equation before estimating flow uncertainty. Therefore, it involves evaluation of the interdependent uncertainties specified in 7.4. In addition, measurement conditions often make it impossible to obtain the replicate measurements needed for evaluation of experimental standard deviations. Thus, it is desirable to express the random errors as well as the systematic errors as error limits. Under these conditions, it also is appropriate to assume that all the random error limits are equivalent to two experimental standard deviations. Under this assumption, the random error limits can be propagated with each other by means of the same root-sum-square formulas as the systematic error limits (see equations 19-22).

B.2 Example one — velocity area method

B.2.1 The equation for discharge in an open channel — velocity area

The channel cross-section under consideration is divided into segments by m verticals. The breadth, depth and mean velocity associated with any vertical i are denoted by b_i, d_i and $\bar{v}_i$ respectively. (see figure 18) The product $Q_i = b_i d_i \bar{v}_i$ represents an approximation to the discharge (volumetric flow rate) in the i-th segment. The sum over all segments,

$$Q_{vo} = \sum_{i=1}^{m} Q_i = \sum_{i=1}^{m} b_i d_i \bar{v}_i \qquad (76)$$

represents an estimated or observed value of the total discharge.

If x and y are respectively horizontal and vertical coordinates of all the points in the cross-section, and A is its total area, then the precise mathematical expression for Q_v, the true volumetric flowrate (discharge) across the area, can be written as

$$Q_v = \iint_A v(x,y)\,dx\,dy \qquad (77)$$

The true discharge and the observed discharge are related by a proportionality factor representing the approximation of the integral equation (77) by the finite sum equation (76), thus:

$$Q_v = F_m Q_{vo} = F_m \sum_{i=1}^{m} b_i d_i \bar{v}_i \qquad (78)$$

where

$$F_m = \left[\iint_A v(x,y)\,dx\,dy\right] \Big/ \left[\sum_{i=1}^{m} b_i d_i \bar{v}_i\right]$$

In practice, F_m can be evaluated from analysis of measurements in which m is sufficiently large for the effects on Q_{vo} of omitting verticals, in stages, to be determined. F_m is subject to a random uncertainty.

It may be convenient in practice to take an F_m variation with m that is a mean value of values for sections of several different rivers, taken together. Then the actual variations of F_m from river to river, as compared with the meaned variation, will involve both systematic and random errors.

F_m is dependent on the number of verticals m, and tends to unity as m increases without limit. Thus, equation 78 can be written approximately as

$$Q_v = \sum_{i=1}^{m} (b_i d_i \bar{v}_i) \qquad (79)$$

with increasing accuracy as m increases.

This last form is the one that is given in ISO 748.

B.2.2 The overall uncertainty of the flow determination

It is plausible to assume that, at a given m, F and Q_{vo} can be treated as independent variables.

However, the Q_i in principle are not independent of one another, since the value corresponding to any one vertical will be related to the values of adjacent verticals. Furthermore, there is an interdependence between the d_i and $\bar{v}_i$ corresponding to any particular vertical. Thus, applying the principles for combining random errors (see clause 5) and denoting random error by S, the following expression for S_Q, the uncertainty of Q, can be derived from equation 78.

$$\left[\frac{S_Q}{Q_v} \right]^2 = \left[\frac{S_{F_m}}{F_m} \right]^2 + \sum_{i=1}^{m} \left(\frac{Q_i}{Q_{vo}} \right)^2 \left\{ \left[\frac{S_{b_i}}{b_i} \right]^2 + \left[\frac{S_{d_i}}{d_i} \right]^2 + \left[\frac{S_{\bar{v}_i}}{\bar{v}_i} \right]^2 \right\} + \frac{2}{Q_{vo}^2} \left\{ \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} S_{i,j} + \sum_{i=1}^{m} \left[\left(\frac{Q_i^2}{d_i \bar{v}_i} \right) S_{d,\bar{v}_i} \bar{v}_i \right] \right\} \quad (80)$$

where S_{ij} arise from the interdependence between Q_i and Q_j and $S_{d\bar{v}_i}$ from the interdependence between d_i and $\bar{v}_i$.

It is convenient to introduce the notation S′ for relative random error.

Thus S_{b_i}/b_i is written S'_{b_i}, S_{F_m}/F_m is written S'_{F_m}, and, neglecting S_{ij}, and $S_{d\bar{v}_i}$, equation (80) becomes

$$S'^2_Q = S'^2_{F_m} + \sum_{i=1}^{m} \frac{Q_i}{Q_{vo}} \left(S'^2_{b_i} + S'^2_{d_i} + S'^2_{\bar{v}_i} \right)$$

If the relative errors S'_b are all nearly enough equal, of value S'_{b_i}, and similarly for the S'_{d_i} and $S'_{d\bar{v}_i}$, then

$$S'^2_Q = S'^2_{F_m} + \left(S'^2_b + S'^2_d + S'^2_{\bar{v}} \right) \sum_{i=1}^{m} \left(Q_i/Q_{vo} \right)^2 \quad (81)$$

If the verticals are so located that $Q_i \approx Q_{vo}/m$, then

$$S'^2_Q \approx S'^2_{F_m} + \frac{1}{m} \left(S'^2_b + S'^2_d + S'^2_{\bar{v}} \right) \quad (82)$$

In multi-point velocity-area methods, velocity is measured at several points on a vertical, and the mean value is obtained by graphical integration or as a weighted average. The latter treatment can be expressed mathematically for a particular value as

$$\bar{v} = \sum_{p=1}^{k} w_p v_p$$

where the w_p are constant weighting factors. The suffix i that identifies the particular vertical is omitted to simplify the symbolism. The points usually are chosen so that $\Sigma w_p = 1$. This equation can also represent the single-point method, by taking k = 1.

In all cases, the estimates $\bar{v}$ so computed are subject to errors. These errors are due to improper placement of the meter at depth and to deviations of the actual velocity profile from the presumed profile. The effect of these errors can be expressed by means of a multiplicative coefficient P analogous to the coefficient F_m used for similar purposes in equation (78). The same analysis that led to equation (80) then yields the following expression for relative random error of the average velocity $\bar{v}$:

$$S'^2_{\bar{v}} = S'^2_p + S'^2_v \frac{\Sigma (w_p v_p)^2}{(\Sigma w_p v_p)^2}$$

in which S′ denotes relative random error in the subscript variable, v is measured point velocity, and the ratio of wv-sums expresses the variability of weighted velocity over the depth of the vertical. For a uniform k-point velocity profile, this ratio would equal 1/k. For an extremely non-uniform profile, in which a single term dominated all the others, the ratio would equal 1. The latter value is adopted, at least for small k values, for the sake of conservatism, with the result

$$S'^2_{\bar{v}} = S'^2_p + S'^2_v$$

This choice also helps to represent the effect of any unaccounted-for correlations among point-velocity errors in the same vertical.

In practice, the random error in the velocity measurement at a point is assumed to be due to a meter-calibration random relative error, S'_c, together with a stream pulsation random error S'_e. Then the random relative error for point velocities is

$$S'^2_v = S'^2_c + S'^2_e$$

The corresponding random relative error for average velocity in the vertical is

$$S'^2_{\bar{v}} = S'^2_p + S'^2_c + S'^2_e \qquad (83)$$

B.2.3 Calculation of uncertainty

It is required to calculate the uncertainty in a current-meter gauging from the following particulars:

Number of verticals used	20
Exposure time of current meter at each point in the vertical	3 min
Number of points taken in the vertical (single point, two points, etc.)	2
Type of current meter rating (individual or group)	individual
Average velocity in measuring section	above 0.3 m/s

Details of procedure are described in ISO 748.

The random and systematic errors are combined by the root-sum-square method as stated in 8.3, i.e., if S'_Q and B'_Q are the percentage overall random and systematic relative errors respectively, then U'_Q, the percentage uncertainty in the current meter gauging, is

$$U'_{Q_{95}} = \sqrt{(2S'_Q)^2 + B'^2_Q} \text{ and } U'_{Q_{99}} = B'_Q + 2S'_Q$$

B.2.3.1 The error equation used for evaluating the overall random error is (see equation (82).)

$$S'_Q = \sqrt{S'^2_m + \frac{1}{m}\left(S'^2_b + S'^2_d + S'^2_{\bar{v}}\right)}$$

where

S'_Q is the overall percentage random error

S'_m is the percentage random error due to the limited number of verticals used;

S'_b is the percentage random error in measuring width of segments;

S'_d is the percentage random error in measuring depth of segments;

$S'_{\bar{v}}$ is the percentage random error in estimating the average velocity in each vertical

$$S'_{\bar{v}} = \pm\sqrt{S'^2_p + S'^2_c + S'^2_e}$$

(see equation (85))

where

S'_p is the percentage error due to limited number of points taken in the vertical (in the present example the two-point method was used, i.e., at 0.2 and 0.8 from the surface respectively);

S'_c is the percentage error of the current meter rating (in the present example an individual rating was used at velocities of the order of 0.30 m/s);

S'_e is the percentage error due to pulsations (error due to the random fluctuation of velocity with time; the time of exposure in the present example was three one-minute readings of velocity.)

The percentage values of the above partial errors at the 95% confidence level are tabulated in B.2.3.2.

The equation for calculating the overall systematic error is

$$B'_Q = \sqrt{B'^2_b + B'^2_d + B'^2_c}$$

where

B'_Q is the overall percentage systematic uncertainty in discharge;

B'_b is the percentage systematic error in the instrument measuring width;

B'_c is the percentage systematic error in the instrument measuring depth; and

B'_d is the percentage systematic error in the current meter rating tank.

The systematic errors in the current meter gauging are confined to the instruments measuring width, depth and velocity and should be restricted to 1% as shown in B.2.3.2.

B.2.3.2 The values of the error elements affecting uncertainty in discharge are tabluated below as percentage errors at the 95% confidence level. The numerical values are taken from ISO 748. It is recommended, however, that each user determine independently the values of the errors for any particular measurement.

Table 13 — Error elements affecting uncertainty in discharge

Error source	*Units*	*(2S') random error limit (2S:95%)*	*(B') percentage systematic error limit*
F_m, number of verticals	—	5.0	—
b, segment width	m	0.5	1.0
d, segment depth	m	0.5	1.0
v_p, number of profile points	m/s	7.0	—
v_c, meter calibration	m/s	2.0	1.0
v_e, meter exposure time	m/s	10.0	—

Then, the overall random error in discharge is given by

$$2S'_Q = 2\sqrt{S'^2_{F_m} + \frac{1}{m}(S'^2_b + S'^2_d + S'^2_p + S'^2_c + S'^2_e)}$$

$$= \sqrt{25 + \frac{1}{20}(0.25 + 0.25 + 49 + 4 + 100)}$$

$$= 5.7\%$$

The overall systematic error is

$$B'_Q = \sqrt{1^2 + 1^2 + 1^2}$$

$$= 1.7\%$$

The combination of both random and systematic errors then gives the overall percentage uncertainty in discharge, U'_Q.

$$U'_{Q_{95}} = \sqrt{(2S'_Q)^2 + B'^2_Q} \qquad U'_{Q_{99}} = B'_Q + 2S'_Q$$

$$= \sqrt{5.7^2 + 1.7^2} \qquad = 1.7 + 5.7$$

$$= 5.9\% \qquad = 7.4\%$$

B.2.3.3 The discharge measurement may be expressed in the following form:

Discharge	(Q) m^3/s
(Combined) uncertainty, $U'_{Q_{95}}$	5.9%
(Combined) uncertainty, $U'_{Q_{99}}$	7.4%
Random error ($2S'_Q$)	5.7%
Systematic error (B'_Q)	1.7%

Uncertainties calculated in accordance with ISO 5168.

B.3 Example two — weir measurement

B.3.1 Weir data

It is required to calculate the discharge and the uncertainty in discharge for a triangular profile weir given the following details: (see figure 19)

Gauged head, h	0.67m
Breadth of weir, b	10m
Crest height, P	1m
Coefficient of discharge, Cd	1.163
Coefficient of velocity, Cv	1.054

The discharge equation is

$$Q = (2/3)^{3/2} C_d C_v \sqrt{g}\, b\, h^{3/2} \tag{84}$$

Details of the procedure are described in ISO 4360.

B.3.2 Uncertainty equations

Taylor series analysis of the discharge equation yields the following uncertainty equations, which can be used for both random and systematic errors:

$$S'_Q = \sqrt{S'^2_{c_dc_v} + S'^2_b + (3/2)^2 S'^2_h}$$

and

$$B'_Q = \sqrt{B'^2_{c_dc_v} + B'^2_b + (3/2)^2 B'^2_h} \qquad (85)$$

in which S′ and B′ denote percentage errors of the subscript variables.

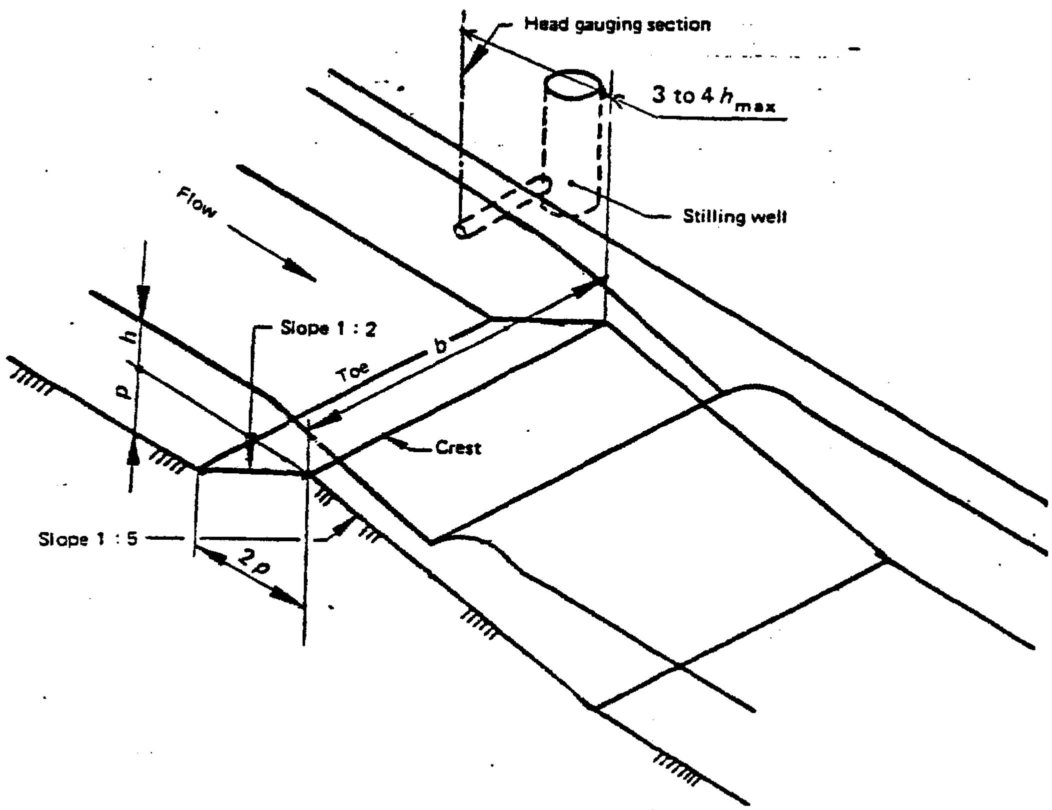


Figure 19 — Triangular profile weir

B.3.3 Evaluation of discharge and uncertainties

The values of the error elements affecting this problem are tabulated below as error limits at the 95% confidence level. The numerical values are based on information given in ISO 4360. It is recommended, however, that each user determine independently the values of the errors for any particular measurement. (See table 14)

Table 14 — Error element values

Variable	*Units*	*Nominal value*	*(2S′) random error limit (2S:95%)*	*(B′) systematic error limit*
h	m	0.67	0.003 0.45%	0.003 0.45%
b	m	10.00	0.	0.01 0.1%
C_dC_v	—	1.226	0.5%	1.5%
g	m/s^2	9.81	0.	0.

0260r

Substitution of the nominal values into the discharge equation yields

$$Q = (3/2)^{3/2} \times (1.226) \times \sqrt{9.81} \times 10 \times (0.67)^{3/2}$$

$$= 11.46\ m^3/s$$

Evaluation of the random errors yields

$$2S'_Q = \sqrt{(0.5)^2 + (3/2)^2\,(0.45)^2}$$

$$= 0.84\%$$

Evaluation of the systematic errors yields

$$B'_Q = \sqrt{(1.5)^2 + (0.1)^2 + (3/2)^2\,(0.45)^2}$$

$$= 1.65\%$$

Combining the random and systematic errors by the root-sum-square (RSS) method yields

$$U'_{Q_{95}} = \sqrt{(2S'^2_Q) + B'^2_Q} \qquad U'_{Q_{99}} = B'_Q + 2S'_Q$$

$$= \sqrt{(0.84)^2 + (1.65)^2} \qquad = 1.65 + 0.84$$

$$= 1.85\% \qquad = 2.49\%$$

B.3.4 Presentation of results

The discharge Q may be reported as follows:

Discharge	$6 m^3/s$
(Combined) uncertainty, $U'_{Q_{95}}$	%
(Combined) uncertainty, $U'_{Q_{99}}$	2.5%
Random error $(2S'_Q)$	0.8%
Systematic error (B'_Q)	1.6%

Uncertainties calculated in accordance with ISO 5168.

Annex C — Small sample methods

C.1 Student's t.

When the experimental standard deviation is based on small samples ($N \leq 30$), uncertainty is defined as:

$$U_{ADD} = B + t_{95}S \tag{86}$$

$$U_{RSS} = \sqrt{B^2 + (t_{95}S)^2} \tag{87}$$

For these small samples, the interval $[\overline{X} - t_{95}S/\sqrt{N}, \overline{X} + t_{95}S/\sqrt{N}]$ will contain the true unknown average, μ, 95% of the time. If the systematic error is negligible, this statistical confidence interval is the uncertainty interval. t_{95} is the 95th percentile point for the two-tailed Students t-distribution. For small samples, t will be large, and for larger samples t will be smaller, approaching 1.96 as a lower limit. The t-value is a function of the number of degrees of freedom (ν) used in calculating S. Since 30 degrees of freedom (ν) yield a t of 2.05 and infinite degrees of freedom yield a t of 1.96, an arbitrary selection of $t = 2$ is used for simplicity for values of ν from 30 to infinity. See table 15.

C.2 Degrees of freedom for small samples

In a sample, the number of degrees of freedom (ν) is the sample size, N. When a statistic is calculated from the sample, the degrees of freedom associated with the statistic is reduced by one for every estimated parameter used in calculating the statistic. For example, from a sample of size N, $\overline{X}$ is calculated and has N degrees of freedom, and the experimental standard deviation, S, is calculated using equation (1), and has N-1 degrees of freedom because $\overline{X}$ is used to calculate S. In calculating other statistics, more than one degree of freedom may be lost. For example, in calculating the standard error of a curve fit, the number of degrees of freedom which are lost is equal to the number of estimated coefficients for the curve, $N - 2$.

When all random error sources have large sample sizes (i.e., $\nu_{ij} > 30$) the calculation of is unnecessary and 2 is substituted for t_{95}. However, for small samples, when combining experimental standard deviations by the root-sum-square method (see equation (20) for example), the degrees of freedom (ν) associated with the combined experimental standard deviations is calculated using the Welch-Satterthwaite formula (88).

For example: the degrees of freedom for the calibration experimental standard deviation (S_1) given by equation (20), is:

$$v_1 = \frac{\left(\sum_{i=1}^{4} S_{i1}^2\right)^2}{\sum_{i=1}^{4} \frac{S_{i1}^4}{v_{i1}}} = \frac{(S_{11}^2 + S_{21}^2 + S_{31}^2 + S_{41}^2)^2}{\frac{S_{11}^4}{v_{11}} + \frac{S_{21}^4}{v_{21}} + \frac{S_{31}^4}{v_{31}} + \frac{S_{41}^4}{v_{41}}} \tag{88}$$

where v_{i1} is the degrees of freedom of each elemental experimental standard deviation in the calibration process.

The degrees of freedom for the measurement experimental standard deviation (S), as given by equation (21) is:

$$v = \frac{\left(\sum_{j=1}^{3} \sum_{i=1}^{K} S_{ij}^2\right)^2}{\sum_{j=1}^{3} \sum_{i=1}^{K} \frac{S_{ij}^4}{v_{ij}}} \tag{89}$$

If the test result is an average, $\bar{X}$, based on a sample of size N,

$$S_{\bar{x}} = \frac{S_X}{\sqrt{N}} \tag{90}$$

As $\sqrt{N}$ is a known constant, the degrees of freedom of $S_{\bar{x}}$ is the same as S, i.e.

$$v_{S_{\bar{x}}} = v \tag{91}$$

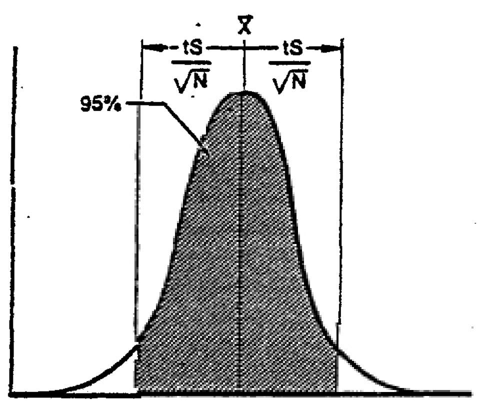
Table 15 — Two-tailed student's "t" table

SMALL SAMPLE METHODS

Degrees of freedom < 30

Degrees of freedom	t_{95}	Degrees of freedom	t_{95}
1	12.706	17	2.110
2	4.303	18	2.101
3	3.182	19	2.093
4	2.776	20	2.086
5	2.571	21	2.080
6	2.447	22	2.074
7	2.365	23	2.069
8	2.306	24	2.064
9	2.262	25	2.060
10	2.228	26	2.056
11	2.201	27	2.052
12	2.179	28	2.048
13	2.160	29	2.045
14	2.145	∞	1.96
15	2.131		
16	2.120		

Two-tailed student's "t" table



30 or more use 2.0

C.3 Propagating the degrees of freedom

The Student's t value of table 16 to be used in calculating the uncertainty of the test result (equations (86) or (87)) is based on v_r, the degrees of freedom of S_r. If the degrees of freedom of any measurement standard deviation is less than 30, the degrees of freedom of the result also may be less than 30. In such cases, the following small sample method may be used to determine v_r. This is defined for the absolute experimental standard deviation according to the Welch-Satterthwaite formula by:

$$v_r = \frac{S_r^4}{\sum_{i=1}^{J} \frac{(\theta_i S_{\bar{P}_i})^4}{v_{\bar{P}_i}}} \tag{92}$$

0260r

and for the relative experimental standard deviation by:

$$v_r = \frac{(S_r/r)^4}{\sum_{i=1}^{J} \frac{(\theta_i S_{\bar{P_i}} / \bar{P}_i)^4}{v_{\bar{P_i}}}} \tag{93}$$

where

$$S_r = \sqrt{\Sigma (\theta_i S_i)^2}$$

and the degrees of freedom of the experimental standard deviation (S_{P_i}) of the independent measurements is usually given by:

$$v_{\bar{P_i}} = (N_i - 1) \tag{94}$$

NOTE: The degrees of freedom for the relative and absolute experimental standard deviations are identical.

Welch-Satterthwaite degrees of freedom may contain fractional, decimal parts. The fractions should be dropped or truncated as rounding down is conservative with Student's t, i.e. $v = 13.6$ should be treated as $v = 13.0$.

Annex D — Outlier treatment

D.1 General

All measurement systems may produce spurious data points. These points may be caused by temporary or intermittent malfunctions of the measurement system or they may represent actual variations in the measurement. Errors of this type should not be included as part of the uncertainty of the measurement. Such points are meaningless as test data. They should be discarded. Figure 20 shows a spurious data point called an outlier.

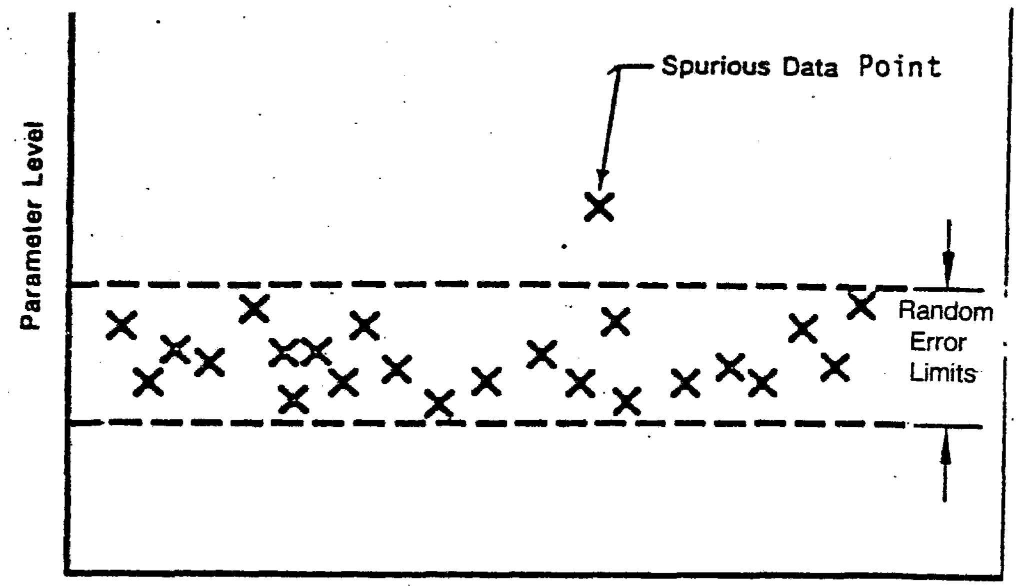


Figure 20 — Outlier outside the range of acceptable data

All data should be inspected for spurious data points as a continuing check on the measurement process. Points should be rejected based on engineering analysis of instrumentation, thermodynamics, flow profiles and past history with similar data. To ease the burden of scanning large masses of data, computerized routines are available to scan steady-state data and flag suspected outliers. The flagged points should then be subjected to an engineering analysis.

The effect of these outliers is to increase the random error of the system. A test is needed to determine if a particular point from a sample is an outlier. The test should consider two types of errors in detecting outliers:

(1) Rejecting a good data point
(2) Not rejecting a bad data point

The probability for rejecting a good point is usually set at 5%. This means that the odds of rejecting a good point are 20 to 1 (or less). The odds will be increased by setting the probability of (1) lower. However, this practice decreases the probability of rejecting bad data points. The probability of rejecting a good point will require that the rejected points be further from the calculated mean and fewer bad data points will be identified. For large sample sizes, several hundred measurements, almost all bad data points can be identified. For small samples (five or ten), bad data points are hard to identify.

One test in common usage for determining whether spurious data are outliers is Grubbs' Method.

D.2 Grubbs' method

Consider a sample (X_i) of N measurements. The mean ($\overline{X}$) and an experimental standard deviation (S) are calculated by equation (1). Suppose that (X_j), the j-th observation, is the suspected outlier; then, the absolute statistic calculated is:

$$T_n = \left[\frac{X_j - \overline{X}}{S} \right]$$

Using table 16, a value of T_n is obtained for the sample size (N) and the 5% significance level (P). This limits the probability of rejecting a good point to 5%. (The probability of not rejecting a bad data point is not fixed. It will vary as a function of sample size.

The test for the outlier is to compare the calculated T_n with the table T_n.

If T_n calculated is larger than or equal to T_n table, we call X_j an outlier.

If T_n calculated is smaller than T_n table, we say X_j is not an outlier.

Table 16 — Rejection values for Grubbs' method

Sample size N	*5% (1-sided)*	*Sample size*	*5% (1-sided)*
3	1.150	20	2.56
4	1.46	21	2.58
5	1.67	22	2.60
6	1.82	23	2.62
7	1.94	24	2.64
8	2.03	25	2.66
9	2.11	30	2.75
10	2.18	35	2.82
11	2.23	40	2.87
12	2.29	45	2.92
13	2.33	50	2.96
14	2.37	60	3.03
15	2.41	70	3.09
16	2.44	80	3.14
17	2.47	90	3.18
18	2.50	100	3.21
19	2.53		

D.3 Example

In the following sample of 40 values,

26	79	58	24	1	−103	−121	−220
−11	−137	120	124	129	−38	25	−60
−148	−52	−216	−12	−56	89	8	−29
−107	20	9	−40	40	2	10	166
126	−72	179	41	127	−35	<u>334</u>	<u>−555</u>

suspected outliers are 334 and -555 (underlined).

To illustrate the calculations for determining whether -555 is an outlier from figure 21.

$$\begin{aligned} \text{Mean } (\overline{X}) &= 1.125 \\ \text{Exp. Std. Dev.} &= 140.813\ 6 \\ \text{Sample Size} &= 40 \end{aligned}$$

$$T_{n_{calc}} = \frac{-555 - 1.125}{140.813\ 6} = 3.95$$

from table 16 using Grubbs' Method for N = 40 @ 5% level of significance (one-sided),

$$T = 2.87$$

Therefore, since 3.95 > 2.87

$$(T_{n_{calc}}) > (T_{n_{table}})$$

−555 is an outlier according to Grubbs' test.

Suspected outlier	Calculated Tn	Table Tn P=5	Sample size (N)	Experimental standard deviation(s)	Mean X
−555	3.95	2.87	40	140.8	1.125
334	2.91 (stop)	2.86	39	109.6	15.385
−220	2.33	2.85	38	97.5	7.000

Figure 21 is a normal probability plot of this data with the suspected outliers indicated. In this case, the engineering analysis indicated that the −555 and 334 readings were outliers, agreeing with the Grubbs' test results.

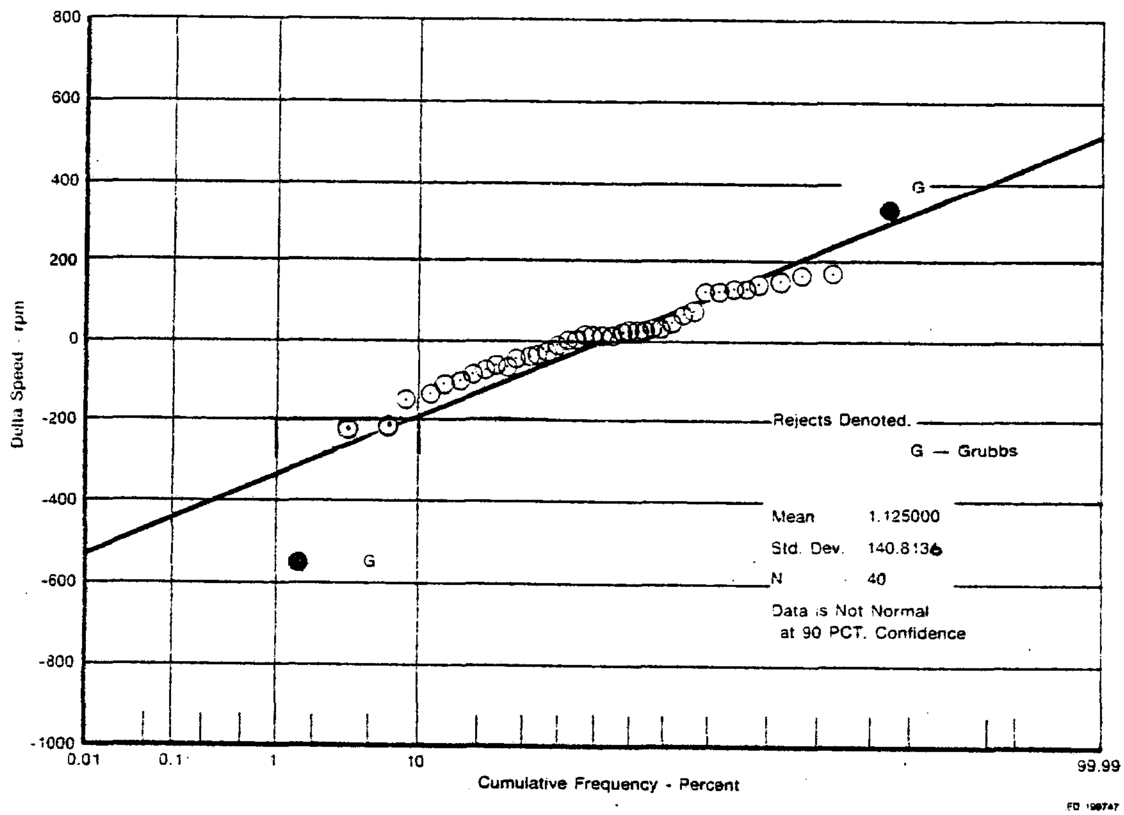


Figure 21 — Results of outlier tests

Annex E — Statistical uncertainty intervals

It is usually impossible to determine the statistical distribution of the systematic errors (β) because they are usually subjective judgments, i.e. not based on data. However, if there is information to justify a distribution assumption, it is possible to use rigorous statistical methods to calculate the uncertainty interval. The validity of this assumption must be left to the judgement of the reader. The purpose of this annex is to describe the methods, given the assumption.

E.1 Assumed systematic error distribution

If it is assumed that the systematic errors (B) are actually the maximum possible upper and lower limit of the true, unknown systematic error (β), and that β is equally probable anywhere within the limits, then the standard deviation of the systematic error may be determined by

$$\sigma_B = \frac{B}{\sqrt{3}} \tag{95}$$

As depicted in figure 22.

The validity of this assumption cannot be proved or disproved. It is a matter of judgement.

E.2 U_{RSS}

The systematic error limit of the measurement result may be calculated as before

$$B = \sqrt{\sum_{i}^{N} (\theta_i B_i)^2} \tag{96}$$

The experimental standard deviation of the systematic error is estimated as:

$$S_B = \frac{B}{\sqrt{3}} \tag{97}$$

The uncertainty is

$$U_{RSS} = \sqrt{(1.645\ S_B)^2 + (2S)^2} \tag{98}$$

for large samples, where S is the experimental standard deviation of the random error.

Assuming there are many sources of systematic and random errors, say ten or more, the Central Limit Theorem states that sums of samples taken from any distribution(s) will tend toward normality. Therefore, the true error (δ) should be distributed as a normal distribution with standard deviation equal to the root-sum-square of the systematic and random error experimental standard deviations. This will be illustrated in E.4. If small samples are used to estimate the random error experimental standard deviations, Student's t and the Welch-Satterthwaite approximation will be needed as described in annex C.

E.3 U_{ADD}

With the additive model of uncertainty, the assumed distribution does not affect the answer. The systematic error, B, is still determined as equation (96) and there is no advantage to calculating a standard deviation of systematic error.

$$U_{ADD} = B + t_{95}S \tag{99}$$

E.4 Monte Carlo example

To illustrate the Central Limit Theorem, the sum of a random sample from each of the ten rectangular distributions with means zero was repeated 1000 times. In sets of three, the distributions had $\sigma = 0.5$, 1.0, 2.0 respectively, and the tenth, $\sigma = 4.0$. If the tendency toward normality and the Monte Carlo simualtion were both perfect

$$\sigma = \sqrt{3(0.5^2 + 1.0^2 + 2.0^2) + 4^2}$$

$$= 5.585.$$

The average S for 1000 trials was $S = 5.671$. The results are shown in figure 23. The bell shape of the normal distribution is apparent. A goodness-of-fit test could not reject normality at the 90% level of confidence.

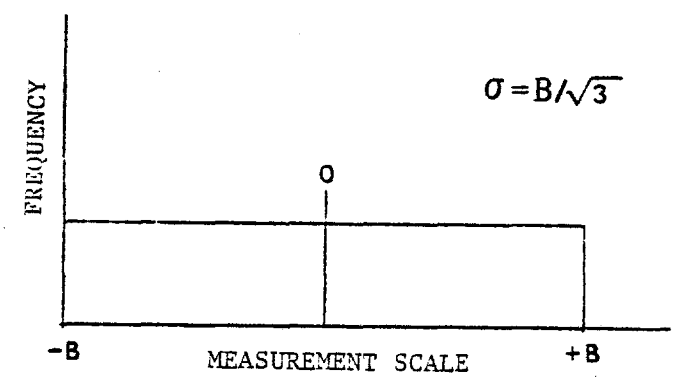


Figure 22 — The assumed frequency rectangular distribution of the systematic error (β) as a function of the limit B.

0260r

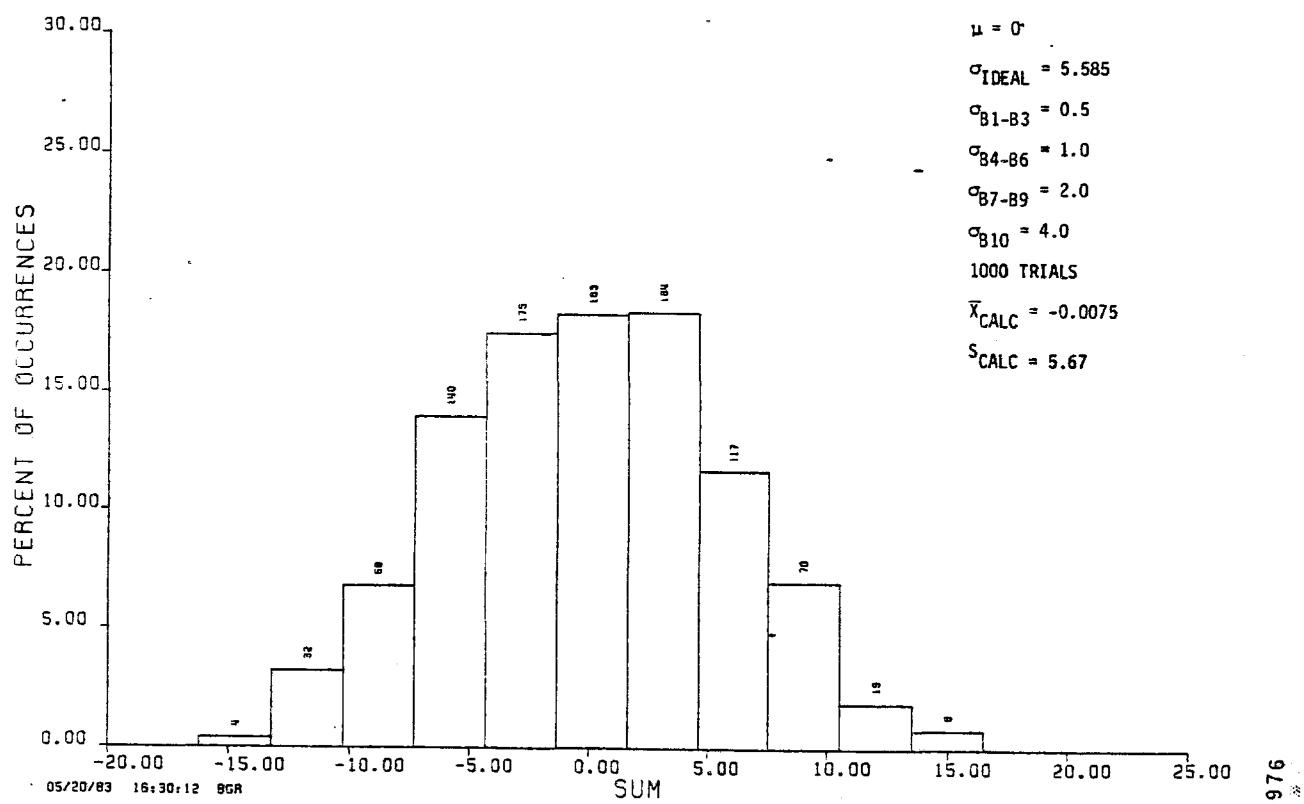


Figure 23 — Distribution of sum of 10 rectangular systematic errors

Annex F: Uncertainty interval coverage

Introduction

A rigorous calculation of confidence level or the coverage of the true value by the interval is not possible because the distributions of systematic error limits, based on judgement, cannot be rigorously defined. Monte Carlo simulation of the intervals can provide approximate coverage* based on assuming various systematic error limits.

F.1 Simulation results

As the actual systematic error and systematic error limit distributions will probably never be known, the simulation studies were based on a range of assumptions. The result of these studies comparing the two intervals are:

* Coverage as used herein is the proportion of Monte Carlo trials where the measurement uncertainty interval contains the true value.

a) U_{99} averages approximately 99.1% coverage while U_{95} provides 95.0% based on systematic error limits assumed to be 95%.

For 99.7% systematic error limits, U_{99} averages 99.7% coverage and U_{95}, 97.5%.

b) The ratio of the average U_{99} interval size to U_{95} interval size is 1.35:1.

c) If the systematic error is negligible, both intervals provide a 95% statistical confidence (coverage).

d) If the random error is negligible, both intervals provide 95% or 99.7% depending on the assumed systematic error limit size.

0260r

Assumptions and Simulation Cases Considered

(1) From 3 to 10 error sources, both systematic and random

(2) Systematic errors distributed both normally and rectangularly

(3) Random error distributed normally

(4) Systematic error limits at both 95% and 99.7% for both the normal and the rectangular distributions

(5) Sample standard deviations based on sample sizes from 3 to 30

(6) Ratio of random to systematic errors at 1/2, 1.0 and 2.0.

F.2 Non-symmetrical interval

If there is a non-symmetrical systematic error limit, the uncertainty (U) is no longer symmetrical about the measurement. The interval is defined by the upper limit of the systematic error interval $(B)^+$. The lower limit is defined by the lower limit of the systematic error interval $(B)^-$. (see clause 7.3)

Figure 24 shows the uncertainty (U) for non-symmetrical systematic error limits. (See table 17.)

$$U^+ = B^+ + t_{95}S \tag{100}$$

$$U^- = B^- - t_{95}S \tag{101}$$

Table 17 — Uncertainty intervals defined by non-symmetrical systematic error limits

B^-	B^+	$t_{95}S_x$	U_{99}^- (Lower limit for U)	U_{99}^+ (Upper limit for U)
0 deg K	+10 deg K	2 deg K	−2 deg K	+12 deg K
−3 Kg	+13 Kg	2 lb	−7 Kg	+17 Kg
0 P_a	+7 P_a	2 P_a	−2 P_a	+9 P_a
−8 deg K	0 deg K	2 deg K	−10 deg K	+2 deg K

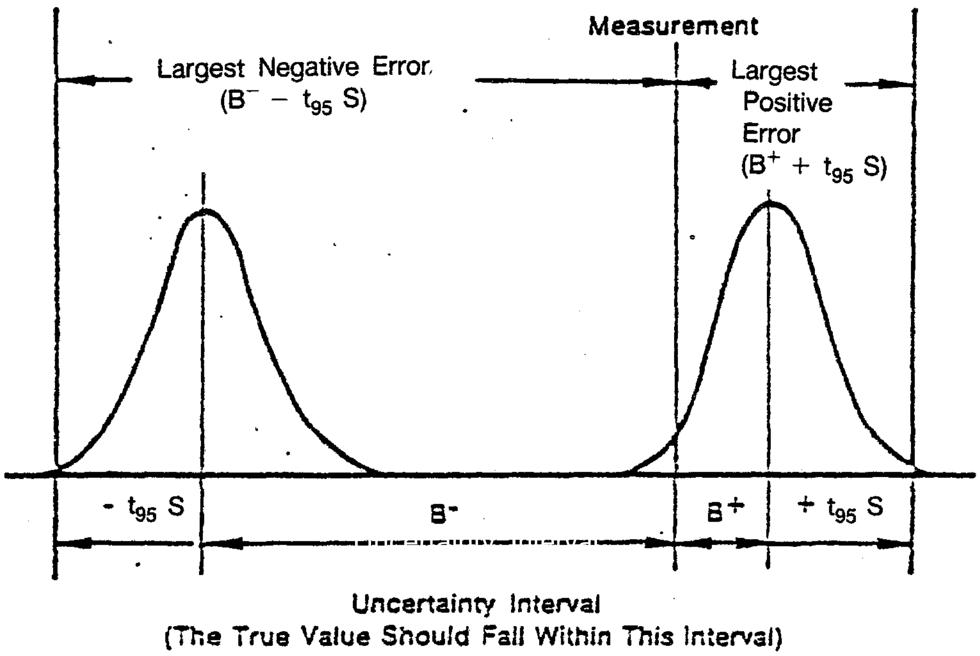


Figure 24 — Measurement uncertainty; non-symmetrical systematic error