Designation: D6091 – 07

AnAmerican National Standard

Standard Practice for 99 %/95 % Interlaboratory Detection Estimate (IDE) for Analytical Methods with Negligible Calibration Error[1]

This standard is issued under the fixed designation D6091; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This practice establishes a standard for computing a 99 %/95 % Interlaboratory Detection Estimate (IDE) and provides guidance concerning the appropriate use and application. The calculations involved in this practice can be performed with DQCALC, Microsoft Excel-based software available from ASTM.[2]

1.2 The IDE is computed to be the lowest concentration at which there is 90 % confidence that a single measurement from a laboratory selected from the population of qualified laboratories represented in an interlaboratory study will have a true detection probability of at least 95 % and a true nondetection probability of at least 99 % (when measuring a blank sample).

1.3 The fundamental assumption of the collaborative study is that the media tested, the concentrations tested, and the protocol followed in the study provide a representative and fair evaluation of the scope and applicability of the test method as written. When properly applied, the IDE procedure ensures that the 99 %/95 % IDE has the following properties:

1.3.1 *Routinely Achievable IDE Value* —Most laboratories are able to attain the IDE detection performance in routine analyses, using a standard measurement system, at reasonable cost. This property is needed for a detection limit to be practically feasible. Representative laboratories must be included in the data to calculate the IDE.

1.3.2 *Routine Sources of Error Accounted for*—The IDE should realistically include sources of bias and variation which are common to the measurement process. These sources include, but are not limited to: intrinsic instrument noise, some typical amount of carryover error, plus differences in laboratories, analysts, sample preparation, and instruments.

1.3.3 *Avoidable Sources of Error Excluded*—The IDE should realistically exclude avoidable sources of bias and variation, that is, those which can reasonably be avoided in routine field measurements. Avoidable sources would include, but are not limited to: modifications to the sample, measurement procedure, or measurement equipment of the validated method, and gross and easily discernible transcription errors (provided there was a way to detect and either correct or eliminate them).

1.3.4 *Low Probability of False Detection* —The IDE is a true concentration consistent with a measured concentration threshold (critical measured value) that will provide a high probability, 99 %, of true nondetection (a low probability of false detection, $\alpha = 1$ %). Thus, when measuring a blank sample, the probability of not detecting the analyte would be 99 %. To be useful, this must be demonstrated for the particular matrix being used, and not just for reagent water.

1.3.5 *Low Probability of False Nondetection*—The IDE should be a true concentration at which there is a high probability, at least 95 %, of true detection (a low probability of false nondetection, $\beta = 5$ %, at the IDE), with a simultaneous low probability of false detection (see 1.3.4). Thus, when measuring a sample at the IDE, the probability of detection would be at least 95 %. To be useful, this must be demonstrated for the particular matrix being used, and not just for reagent water.

NOTE 1—The referenced probabilities, α and β, are key parameters for risk-based assessment of a detection limit.

1.4 The IDE applies to measurement methods for which calibration error is minor relative to other sources, such as when the dominant source of variation is one of the following (with comment):

1.4.1 *Sample Preparation,* and calibration standards do not have to go through sample preparation.

1.4.2 *Differences in Analysts,* and analysts have little opportunity to affect calibration results (such as with automated calibration).

1.4.3 *Differences in Laboratories* , for whatever reasons, perhaps difficult to identify and eliminate.

1.4.4 *Differences in Instruments* (measurement equipment), which could take the form of differences in manufacturer, model, hardware, electronics, sampling rate, chemical processing rate, integration time, software algorithms, internal signal processing and thresholds, effective sample volume, and contamination level.

[1] This practice is under the jurisdiction of ASTM Committee D19 on Water and is the direct responsibility of Subcommittee D19.02 on Quality Systems, Specification, and Statistics.

Current edition approved March 1, 2007. Published April 2007. Originally approved in 1997. Last previous edition approved in 2003 as D6091 – 03. DOI: 10.1520/D6091-07.

[2] Available from ASTM International Headquarters. Order Adjunct No. ADJDQCALC. Original adjunct produced in 2007.

1.5 *Alternative Data Quality Objectives* —Other values forα, β, confidence, etc. may be chosen for calculating an IDE; however, this procedure addresses only the 99 %/95 % IDE.

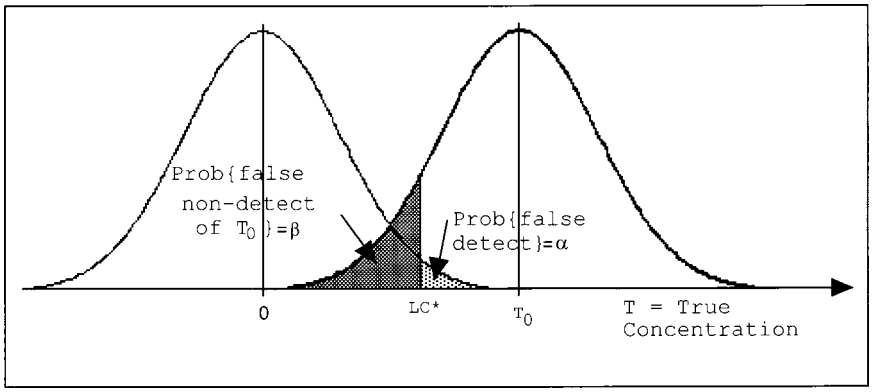

FIG. 1 Simplest Case of Reliable Detection

2. Referenced Documents

2.1 *ASTM Standards:*

D2777 Practice for Determination of Precision and Bias of Applicable Test Methods of Committee D19 on Water

2.2 *ASTM Adjuncts:*

DQCALC Microsoft Excel-based software for the Interlaboratory Quantitation Estimate (IQE)[2]

3. Terminology

3.1 *Definitions:*

3.1.1 *99 %/95 % Interlaboratory Detection Estimate (99 %/95 % IDE, also denoted LD for Limit of Detection in accordance with Currie* **(1)**[3]—The lowest concentration at which there is 90 % confidence that a single measurement from a laboratory selected from the population of qualified laboratories represented in an interlaboratory study will have a true detection probability of at least 95 % and a true nondetection probability of at least 99 %.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *Censored Measurement*—A measurement that is not reported numerically nor is reported missing but as a nondetect or a less-than, for example, "less than 0.1 ppb." The former means that an algorithm in the measurement system determined that the measurement should not be reported numerically for one of two reasons: either it was considered not sufficiently precise or accurate, or the identification of the analyte was suspect. A reported less-than may have the same meaning, but it also implies (perhaps erroneously) that any concentration greater than or equal to the accompanying value (for example, 0.1 ppb) can be measured and will be reported numerically.

3.2.2 *Detection Limit (DL) or Limit of Detection (LD)*—A numerical value, expressed in physical units or proportion, intended to represent the lowest level of reliable detection (a level which can be discriminated from zero with high probability while simultaneously allowing high probability of nondetection when blank samples are measured.

NOTE 2—In some cases, the discrimination may be from a value other than zero, such as a background level. Note also that a DL also depends on other characteristics of the measurement and detection process, such as described in 1.3.2. The IDE is an example of a DL.

3.2.3 *Probability of False Detection*—The false positive probability, denoted α, that a single measurement of a blank sample will result in a detection. (See Fig. 1.) This probability is often referred to as the Type 1 error probability and depends on the analyte, measurement system, analytical method, matrix, analyst, and measurement (recovery) threshold (measurement critical value) used to decide whether detection has occurred. This definition can be generalized to refer to unwanted detection from a single measurement of a sample at any nonzero concentration of the analyte rather than a blank sample, provided that the nonzero concentration is less than the detection limit or IDE.

3.2.4 *Probability of False Nondetection*—The false negative probability, denoted β or β (*T*), that a single measurement of a sample containing a nonzero concentration, *T*, of an analyte of interest will result in a nondetection. This is the complement of the probability of true detection. (See Fig. 1.) This probability function is often referred to as the Type 2 error probability function, and it depends explicitly on the concentration (*T*). It depends implicitly on the analyte, measurement system, analytical method, matrix, analyst, and critical value for detection.

3.2.5 *Probability of True Detection*—The probability, denoted 1-β or 1-β (*T*), that a single measurement of a sample containing a nonzero concentration, *T*, of an analyte of interest will result in a detection. (See Fig. 1.) This probability is often referred to as statistical power or the power of detection, and it depends explicitly on the concentration (*T*). It depends implicitly on the analyte, measurement system, analytical method, matrix, analyst, and critical value for detection.

3.2.6 *Probability of True Nondetection*—The true negative probability, denoted 1-α, that a single measurement of a blank sample will result in a nondetection. This is the complement of

[3] The boldface numbers in parentheses refer to the list of references at the end of this standard.

the probability of false detection. (See Fig. 1.) This probability also depends on the analyte, measurement system, analytical method, matrix, analyst, and response threshold. The probability of true nondetection can be similarly generalized: it can apply to a single measurement of a sample at any nonzero concentration less than the detection limit or IDE.

3.2.7 *100(1-γ) %—Confidence Statistical Tolerance Limit for 100(1-δ) % of a Population (also known as a One-Sided Statistical Tolerance Interval)*—A statistically determined limit that will, with 100(1-γ) % confidence, exceed (or fall below) 100(1-δ) % of the population (the 100(1-δ) % quantile). See Hahn and Meeker **(2)** for further explanation and tables of values.

4. Summary of Practice

4.1 Every ASTM D-19 test method is evaluated to determine precision and bias by conducting a collaborative study in accordance with Practice D2777. That study, or a similar collaborative study, can also be used to evaluate the lowest concentration level of reliable detection for a test method, referred to herein as the Interlaboratory Detection Estimate. Such a study must include concentrations suitable for modeling the uncertainty of mean recovery of interlaboratory measurement (preferably without extrapolation). It must also be planned and conducted to allow the known, routine sources of measurement variability to be observed at typical levels of influence. After it is conducted, outlying laboratories and individual measurements should be eliminated using an accepted, scientifically based procedure for outlier removal, such as found in Practice D2777. The IDE computations must be based on retained data from at least six independent laboratories at each concentration level.

4.2 Retained data are analyzed to identify and fit one of three proposed interlaboratory standard deviation (ILSD) models which describe the relationship between the interlaboratory standard deviation of measurements and the true concentration. The identification process involves evaluating the models in order, from simplest to most complex: constant, straight-line, or exponential (all with respect to true concentration, *T*). Evaluation includes statistical significance and residual analysis.

4.3 The chosen model is used to predict interlaboratory measurement standard deviation at any true concentration within the study concentration range. If interlaboratory standard deviation is not constant, the predictions are used to generate weights for fitting the mean recovery relationship (the straight-line relationship between measured concentration and true concentration), using weighted least squares (otherwise, ordinary least squares is used). The mean recovery curve is evaluated for statistical significance and lack of fit and using residual analysis. An ILSD model prediction is also used to estimate the interlaboratory standard deviation of measurements of blanks. This estimate is used to compute *YC*, a measurement critical value for detection (see 6.4.1). The *YC* is the value that with approximately 90 % confidence will not be exceeded by 99 % of all measurements of blanks made by qualified laboratories as represented in the study. The *LC* computed from *YC* is the true concentration with expected measurement equal to *YC* (see 6.4.2). The model is also used to predict interlaboratory standard deviation at nonzero concentrations. The IDE is directly or iteratively computed to be the true concentration that with approximately 90 % confidence will produce measurements that will exceed *YC* at least 95 % of the time and simultaneously not exceed more than 1 % of the time when blank samples are measured.

5. Significance and Use

5.1 Appropriate application of this practice should result in an IDE achievable by most laboratories properly using the test method studied. This IDE provides the basis for any prospective use of the test method by qualified laboratories for reliable detection of low-level concentrations of the same analyte as the one studied in this practice and same media (matrix).

5.2 The IDE values may be used to compare the detection power of different methods for analysis of the same analyte in the same matrix.

5.3 The IDE provides high probability (approximately 95 %) that result values of the method studied which exceed the IDE represent presence of analyte in the sample and high probability (approximately 99 %) that blank samples will not result in a detection.

5.4 The IDE procedure should be used to establish the interlaboratory detection capability for any application of a method where interlaboratory detection is important to data use. The intent of IDE is not to set reporting limits.

6. Procedure

6.1 The procedure described as follows has stages described in the following sections: IDE Study Plan, Design and Protocol (6.2); Conduct the IDE Study, Screen the Data, and Choose a Model (6.3); and Compute the IDE (6.4). A flowchart of the procedure is shown in Fig. 2.

6.2 *IDE Study Plan, Design, and Protocol:*

6.2.1 *Choose Analyte, Matrix, and Method*—At least one analyte of interest is selected, typically one for which there is interest in trace levels of concentration, such as toxic materials that are controlled and regulated. For each analyte, an approximate maximum true concentration is selected based on the following considerations:

6.2.1.1 The anticipated IDE should be exceeded by a factor of 2 or more,

6.2.1.2 A single model (ideally a straight-line model in true concentration, *T*) should describe mean recovery from zero to that maximum concentration,

6.2.1.3 A single model in true concentration should describe interlaboratory measurement standard deviation from zero to that maximum concentration, and

6.2.1.4 The range must be sufficient to enable statistically significant coefficients to be estimated for the ILSD model and mean recovery model. One or more matrices of interest are also selected, and an accepted standard analytical method for those analytes is selected for study. If there is no possibility of matrix interference, then it may only be necessary to determine a list of acceptable matrices which can be used instead of selecting a specific matrix. For example, for a particular analyte,

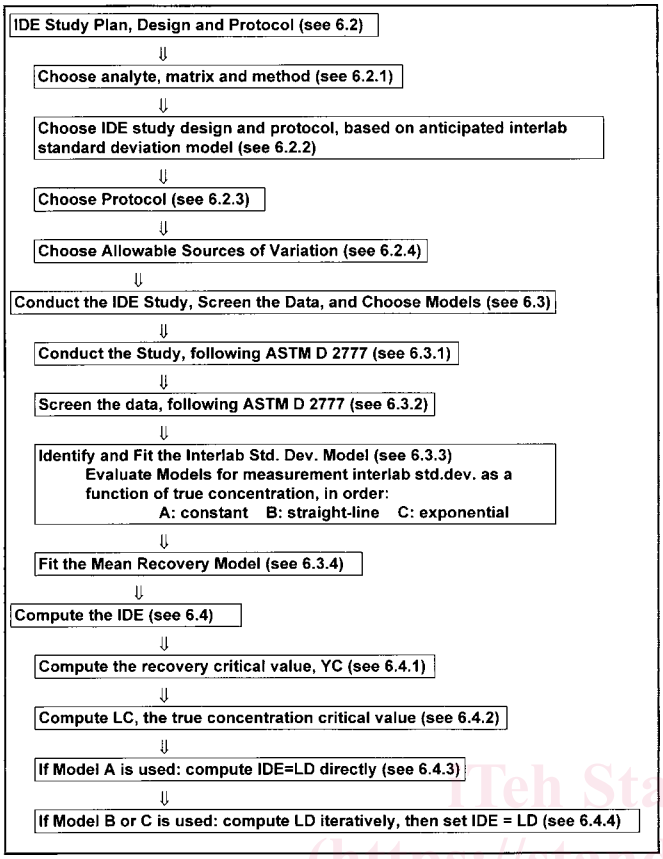

FIG. 2 Flowchart of IDE Procedure

concentration range, and method it may be supposed that reagent waters from different laboratories are indistinguishable, but for another analyte or another concentration range that assumption may not hold.

6.2.2 *Choose IDE Study Design and Protocol,* based (if possible) on anticipated interlaboratory standard deviation (ILSD) model. Section 7 of Practice D2777 can be followed for the study design and protocol. The anticipated form of the ILSD model (the relationship between interlaboratory measurement standard deviation and true concentration) can help in choosing an IDE study design. Three models are proposed herein for the interlaboratory measurement standard deviation with respect to true concentration: constant, straight-line (increasing), and exponential (increasing). Chemistry, physics, empirical evidence, or informed judgment may make one model more likely than others. However, it may not be possible to anticipate the relationship between standard deviation and true concentration.

6.2.2.1 Select an IDE study design that has enough distinct concentrations to assess statistical lack of fit of the models (see Draper and Smith (**3**)). Recommended designs are: (*a*) The semi-geometric design at five or more true concentrations, {T_1, T_2, and so forth}, such as: {0, $IDE_0/4$, $IDE_0/2$, IDE_0, $2 \times IDE_0$, $4 \times IDE_0$}, where IDE_0 is an initial estimate of the IDE (such as $10 \times s`$, where $s`$ is the interlaboratory measurement standard deviation at a trace-level, nonzero concentration), (*b*) equi-spaced design: {0, $IDE_0/2$, IDE_0, $(3/2) \times IDE_0$, $2 \times IDE_0$, $(5/2) \times IDE_0$}, and (*c*) any other design with at least five concentrations, provided that the design includes blanks, at least one concentration approximately equal to $2 \times IDE_0$, and at least one nonzero concentration below IDE_0.

6.2.2.2 The study concentration levels must either be: known (true concentration levels), or knowable, after the fact. A concentration is considered known if reference standards can be purchased or constructed and knowable if an accurate determination can be made (for example, the median value from many laboratories, or results from a recognized laboratory, such as NIST, using a high-accuracy method).

6.2.3 *Choose Protocol*— The protocol should follow Section 7 of Practice D2777. It should include design run order and details on when the system is to be purged, have extra blanks run, and so forth. It should take into consideration possible problems with carryover, study cost (in time and money), and time constants of measurement system drift or sample degradation.

6.2.3.1 For purposes of the collaborative study, the study supervisor should provide instructions to participating laboratories to disable (if possible) any internal measurement system thresholds (such as an instrument detection limit or peak-area threshold) that are used to determine whether a numerical measurement is to be reported as a nondetect or less-than, or as a number (censoring). If censoring is unavoidable, the laboratory censoring threshold must be reported with its study data. However, qualitative criteria used by the method to identify and discriminate analytes are separate criteria and must be satisfied according to the method.

6.2.4 *Choose Allowable Sources of Variation*—It is assumed that collectively the many sources of variation will contribute to cause interlaboratory measurements at any true concentration to be normally distributed. Representative between-laboratory variation can only be seen if the number of laboratories providing usable data is maximized. Ordinary within-laboratory variation must be allowed to affect the measurement process as happens in routine measurement. Ideally, there would be many laboratories, and each measurement at each laboratory would be an unsuspecting blind measurement made by a different analyst using a different (qualified) measurement system on a different day, in random order.

6.2.4.1 As emphasized in Practice D2777, maximizing the number of participating laboratories is often the most important thing that can be done to guarantee a successful study, and there are several reasons why the number of participating laboratories will somewhat exceed the number of laboratories providing a full set of usable data. A minimum of ten participating laboratories is recommended.

6.2.4.2 If possible, the study should be conducted completely blind, particularly if the method is labor-intensive, as opposed to a highly automated method. That is, not only should the analysts not be aware of the true concentrations of the samples they are measuring, but they should not even be aware of the fact that they are measuring special, study samples. This is to minimize the extra care distortion of data so common in analytical studies.

6.2.4.3 For each laboratory, the maximum number of qualified analysts possible should be involved in the study since

there are variations which may be allowed by the method, may be practiced by different analysts, and will be seen in routine analyses.

6.2.4.4 For each laboratory, the maximum number of qualified measurement systems should be used since there are model-to-model and instrument-to-instrument differences in equipment and maintenance, as will be seen in routine analyses.

6.2.4.5 For each laboratory, the IDE study should be scheduled to span the maximum possible number of days consistent with holding time constraints since day-to-day changes in analytical laboratory environmental conditions, contamination, solvent purity, and other factors can affect measurements, and will be seen in routine analyses.

6.3 *Conduct the IDE Study, Screen the Data, and Choose a Model:*

6.3.1 The IDE study should be conducted in accordance with Section 9 of Practice D2777. Blank correction should not be performed by the laboratories, unless the method requires this subtraction in order to perform the test. Each laboratory should supply method blank data along with the uncorrected measurement values, and the study supervisor can determine whether the reported measurements should be corrected.

6.3.2 The IDE study data should be screened in accordance with the initial subsections relating to removing data, Section 10 of Practice D2777. Skip to 6.5 if, for any concentration, more than 10 % of the retained measurements are nondetects or less-thans.

6.3.3 *Identify and Fit the ILSD Model* —The ILSD model should be identified, and its coefficients should be estimated by using the following procedure. See Caulcutt and Boddy **((4))** for more discussion of standard deviation modeling and weighted least squares (WLS) in analytical chemistry. This model is an attempt to characterize the unknown (or partly known) function between interlaboratory measurement standard deviation and true concentration, $\sigma = G(T)$. It is used for two purposes: to provide weights for the WLS regression to fit the mean recovery model and to provide the interlaboratory standard deviation estimates crucial to determining critical values and the IDE.

6.3.3.1 Three ILSD models are proposed. The identification process considers (fits and evaluates) each model in turn, from simplest to most complex, until a suitable model is found. Prior knowledge can be combined with empirical results to influence the selection of a model if a suitable refereed publication can be cited. See Carroll and Ruppert **((5))** for further discussion of standard deviation modeling. The model order is as follows:

Model A (Constant ILSD Model):

$$s = g + \text{error} \quad (1)$$

where: g is a fitted constant. Standard deviation does not change with concentration, resulting in a relative standard deviation that declines with increasing T.

Model B (Straight-line ILSD Model):

$$s = g + h \times T + \text{error} \quad (2)$$

where: g and h are fitted constants. Standard deviation increases linearly with concentration, resulting in an asymptotically constant relative standard deviation as T increases.

Model C (Exponential ILSD Model):

$$s = g \times \exp\{h \times T\} + \text{error} \quad \text{or} \quad (3)$$

$$s = g \times \exp\{h \times T\} \times \text{error} \quad (4)$$

where: g and h are fitted constants. Interlaboratory standard deviation increases exponentially with concentration, resulting in a relative standard deviation that may initially decline as T increases but eventually increases as T increases. Error can be additive or multiplicative.

(a) In all cases, it is assumed that $g > 0$. A value of $g < 0$ has no practical interpretation and may indicate that a different ILSD model should be used. Furthermore, it is assumed that g is not underestimated due to censored data among measurements of blanks or other low-concentration samples. (Censoring is addressed in 6.2.3.1, 6.3.2, and 6.5.)

(b) If $h < 0$, it must not be statistically significant, and Model A should be evaluated.

6.3.3.2 *ILSD Model Identification and Fitting Procedure:*

(*1*) Merge all retained IDE study data (after possible elimination of some data in accordance with 6.3.2).

(*2*) For each true concentration, T_k, compute the adjusted interlaboratory sample standard deviation, s_k, an estimate of the true underlying interlaboratory measurement standard deviation, σ_k. The adjusted interlaboratory sample standard deviation is the sample standard deviation s_k, multiplied by the bias-correction factor, a'_n found in Table 1. In this Practice, all references to computed and fitted values of the interlaborator sample standard deviation refer to adjusted values. Note that a simplifying approximation can be used if the number of retained replicates is the same for each spike level; unadjusted sample standard deviations can be sued, and the final IDE can be multiplied by the adjustment factor (see the example). The larger the number of replicates, the better the approximation.

(*3*) Plot s_k versus T_k.

(*4*) Using ordinary least squares (OLS) (see Caulcutt and Boddy **(4)**), regress s_k on T_k, temporarily assuming that a straight-line model is valid. This provides coefficients, g and h, in the relationship:

TABLE 1 Bias-Correction Adjustment Factors for Sample Standard Deviations Based on n Measurements (at at particular concentration)[A]

n	2	3	4	5	6	7	8	9	10
a'_n	1.253	1.128	1.085	1.064	1.051	1.042	1.036	1.031	1.028

[A] For each true concentration T_k, the adjusted value $s_k = a'_n s'_k$ should be modeled in place of sample standard deviation s'_k. For $n > 10$, use the formula $a'_n = 1 + [4(n - 1)]^{-1}$. See Johnson and Kotz **(7)**.

$$s_k = g + h \times T_k + \text{error} \quad (5)$$

(*e*) Evaluate the reasonableness of Model A (the constant ILSD model) by doing two things. Note the *p*-value associated with slope estimate *h*, from the OLS regression. If it is less than 5 %, there is statistically significant slope, and Model A should be rejected; proceed to the next step. Secondly, examine the plot produced in step (*c*), or a plot of the residuals from the OLS fit. If obvious systematic curvature is present (for example, quadratic or exponential-like behavior), Model A should be rejected; proceed to step (*h*). If Model A is not rejected, skip to 6.3.4.

(*f*) Model A is rejected, due to statistically significant slope. Compute residuals:

$$r_k = s_k - (g + h \times T_k) \quad (6)$$

Plot r_k versus T_k.

(*g*) Evaluate the reasonableness of Model B (the straight-line ILSD model). Examine the plot produced in step (*f*). If obvious systematic curvature is present (for example, quadratic or exponential-like behavior), with a minimum that appears to be within the concentration range, Model B should be rejected; proceed to step (*h*). If Model B is not rejected, skip to 6.3.4.

(*h*) To evaluate the reasonableness of Model C (the exponential ILSD model), the model must first be fit. There are two approaches. The simplest approach is to do OLS regression on the log of the interlaboratory sample standard deviations:

$$\ln s_k = \ln g + h \times T_k + \text{error} \quad (7)$$

This corresponds to the multiplicative error assumption, which is generally a good assumption. The fit will provide *h* directly and $g` = \ln g$ which is converted, $g = \exp\{ g`\}$. Alternatively, the fit can be done using nonlinear least squares (NLLS), by Newton-Raphson iteration or another method. This approach corresponds to the less-plausible additive error assumption. In either case, the fit should satisfy two types of evaluation. First, the *p*-value for *h* should be less than 5 %. Secondly, a plot of the residuals, in log form, should be constructed. Plot r_k versus T_k, where:

$$r_k = \ln s_k - (\ln g + h \times T_k) \quad (8)$$

The plot should show no systematic behavior (for example, curvature). If the fit satisfies both types of evaluation, proceed to 6.3.4. Otherwise, a different and possibly more complex model will have to be used. One possibility is the Rocke and Lorenzato **(6)** model, which has:

$$s \approx (g + h \times T^2)^{1/2} \quad (9)$$

This model has nearly constant (slightly increasing) ILSD for low true concentrations, changing to standard deviation nearly proportional to concentration for higher concentration levels. It can be fit and evaluated using NLLS or maximum likelihood. If there are enough true concentrations, a model with more coefficients could be considered, such as quadratic (strictly increasing with increasing concentration), or even cubic.

6.3.4 *Fit the Mean Recovery Model* —The mean recovery model is a simple straight line:

$$\text{Model } R: Y = a + b \times T + \text{error} \quad (10)$$

The fitting procedure depends on the model selection from 6.3.3. If Model A was selected for ILSD, then OLS can be used to fit Model R for mean recovery (see Caulcutt and Boddy **(4)**). If a nonconstant ILSD model was selected, such as Model B or C, then WLS should be used to fit mean recovery. This approximately provides the minimum variance unbiased linear estimate of the coefficients *a* and *b*. The WLS procedure appears in 6.3.4.1.

6.3.4.1 *Weighted Least Squares Procedure, Using the Interlaboratory Standard Deviation Model:*

(a) Using the ILSD model and coefficient estimates from 6.3.3, compute predicted interlaboratory standard deviation, $\hat{s}_k$ for each true concentration, T_k:

$$\text{Model } B: \hat{s}_k = g + h \times T_k \quad (11)$$

$$\text{Model } C: \hat{s}_k = g \times \exp\{h \times T_k\} \quad (12)$$

(b) Compute weights for WLS:

$$w_k = (\hat{s}_k)^{-2} \quad (13)$$

(c) Note that if this is done using computer software, the default setting for weights may be different. For example, instead of supplying the values, $(\hat{s}_k)^{-2}$ as weights, the software may require the user to supply values $(\hat{s}_k)$ or $(\hat{s}_k)^2$ as weights that are internally transformed by the software.

(d) Carry out WLS computations analogous to OLS computations. See Table 2 or Caulcutt and Boddy **(4)**. The result will be coefficient estimates, *a* and *b*, for the mean recovery model, Model R.

(e) There are three approximate approaches to WLS commonly practiced but that are not acceptable for this application. One approach uses the reciprocal squared sample standard deviations as weights. In this context, since a standard deviation model is explicitly evaluated and selected, the predicted value for s_k is probably more precise than a sample value. The

TABLE 2 Computations to Estimate Straight-Line Model Coefficients By Means of Least Squares—Ordinary and Weighted

Ordinary Least Squares, OLS	Weighted Least Squares, WLS
$\bar{T} = \frac{1}{n}\sum_{i=1}^{n} T_i,$	$\bar{T}_w = \sum_{i=1}^{n} w_i T_i / \sum_{i=1}^{n} w_i$
$\bar{y} = \bar{T} = \frac{1}{n}\sum_{i=1}^{n} T_i$	$\bar{y}_w = \sum_{i=1}^{n} w_i y_i / \sum_{i=1}^{n} w_i$
$S_{TT} = \sum_{i=1}^{n} (T_i - \bar{T})^2$	$S_{wTT} = \sum_{i=1}^{n} w_i (T_i - \bar{T})^2$
$S_{TY} = \sum_{i=1}^{n} (T_i - \bar{T})(y_i - \bar{y})$	$S_{wTY} = \sum_{i=1}^{n} w_i (T_i - \bar{T})(y_i - \bar{y})$
Slope = $b = S_{TY}/S_{TT}$ Intercept = $a = \bar{y} - b\bar{T}$	Slope = $b = S_{wTY}/S_{wTT}$ Intercept = $a = \bar{y}_w - b\bar{T}_w$