



Designation: D7366 – 08 (Reapproved 2019)

Standard Practice for Estimation of Measurement Uncertainty for Data from Regression-based Methods¹

This standard is issued under the fixed designation D7366; 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 the measurement uncertainty for applicable test methods in Committee D19 on Water. The practice does not provide a single-point estimate for the entire working range, but rather relates the uncertainty to concentration. The statistical technique of regression is employed during data analysis.

1.2 Applicable test methods are those whose results come from regression-based methods and whose data are intra-laboratory (not inter-laboratory data, such as result from round-robin studies). For each analysis conducted using such a method, it is assumed that a fixed, reproducible amount of sample is introduced.

1.3 Calculation of the measurement uncertainty involves the analysis of data collected to help characterize the analytical method over an appropriate concentration range. Example sources of data include: (1) calibration studies (which may or may not be conducted in pure solvent), (2) recovery studies (which typically are conducted in matrix and include all sample-preparation steps), and (3) collections of data obtained as part of the method's ongoing Quality Control program. Use of multiple instruments, multiple operators, or both, and field-sampling protocols may or may not be reflected in the data.

1.4 In any designed study whose data are to be used to calculate method uncertainty, the user should think carefully about what the study is trying to accomplish and much variation should be incorporated into the study. General guidance on designing studies (for example, calibration, recovery) is given in [Appendix X1](#). Detailed guidelines on sources of variation are outside the scope of this practice, but general points to consider are included in [Appendix X2](#), which is not intended to be exhaustive. With any study, the user must think carefully about the factors involved with conducting the

analysis, and must realize that the computed measurement uncertainty will reflect the quality of the input data.

1.5 Associated with the measurement uncertainty is a user-chosen level of statistical confidence.

1.6 At any concentration in the working range, the measurement uncertainty is plus-or-minus the half-width of the prediction interval associated with the regression line.

1.7 It is assumed that the user has access to a statistical software package for performing regression. A statistician should be consulted if assistance is needed in selecting such a program.

1.8 A statistician also should be consulted if data transformations are being considered.

1.9 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety, health, and environmental practices and determine the applicability of regulatory limitations prior to use.*

1.10 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

2. Referenced Documents

- 2.1 *ASTM Standards:*²
[D1129 Terminology Relating to Water](#)

3. Terminology

3.1 Definitions:

3.1.1 For definitions of terms used in this standard, refer to Terminology [D1129](#).

3.2 Definitions of Terms Specific to This Standard:

3.2.1 *confidence level, n*—the probability that the prediction interval from a regression estimate will encompass the true

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 Dec. 1, 2019. Published January 2020. Originally approved in 2008. Last previous approval in 2013 as D7366 – 08 (2013). DOI: 10.1520/D7366-08R19.

² For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

value of the amount or concentration of the analyte in a subsequent measurement. Typical choices for the confidence level are 99 % and 95 %.

3.2.2 *fitting technique, n*—a method for estimating the parameters of a mathematical model. For example, ordinary least squares is a fitting technique that may be used to estimate the parameters $a_0, a_1, a_2, \dots$ of the polynomial model $y = a_0 + a_1x + a_2x^2 + \dots$, based on observed $\{x, y\}$ pairs. Weighted least squares is also a fitting technique.

3.2.3 *lack-of-fit (LOF) test, n*—a statistical technique when replicate data are available; computes the significance of residual means to replicate y variability, to indicate whether deviations from model predictions are reasonably accounted for by random variability, thus indicating that the model is adequate; at each concentration, compares the amount of residual variation from model prediction with the amount of residual variation from the observed mean.

3.2.4 *least squares, n*—fitting technique that minimizes the sum of squared residuals between observed y values and those predicted by the model.

3.2.5 *model, n*—mathematical expression (for example, straight line, quadratic) relating y (directly measured value) to x (concentration or amount of analyte).

3.2.6 *ordinary least squares (OLS), n*—least squares, where all data points are given equal weight.

3.2.7 *prediction interval, n*—a pair of prediction limits (an “upper” and “lower”) used to bracket the “next” observation at a certain level of confidence.

3.2.8 *p-value, n*—the statistical significance of a test; the probability value associated with a statistical test, representing the likelihood that a test statistic would assume or exceed a certain value purely by chance, assuming the null hypothesis is true (a low p -value indicates statistical significance at a level of confidence equal to 1.0 minus the p -value).

3.2.9 *regression, n*—an analysis technique for fitting a model to data; often used as a synonym for OLS.

3.2.10 *residual, n*—error in the fit between observed and modeled concentration; response minus fit.

3.2.11 *root mean square error (RMSE), n*—an estimate of the measurement standard deviation (that is, inherent variation in the measurement system).

3.2.12 *significance level, n*—the likelihood that a measured or observed result came about due to simple random behavior.

3.2.13 *uncertainty (of a measurement), n*—the lack of exactness in measurement (for example, due to sampling error, measurement variation, and model inexactness); a statistical interval within which the measurement error is believed to occur, at some level of confidence.

3.2.14 *weight, n*—coefficient assigned to observations in order to manipulate their relative influence in subsequent calculations. For example, in weighted least squares, noisy observations are weighted downwards, while precise data are weighted upwards.

3.2.15 *weighted least squares (WLS), n*—least squares, where data points are weighted inversely proportional to their variance (“noisiness”).

4. Summary of Practice

4.1 Key points of the statistical protocol for measurement uncertainty are:

4.1.1 Within the working range of the method’s data set, the estimate of the method uncertainty at any given concentration is calculated to be plus-or-minus the half-width of the prediction interval.

4.1.2 The total number of data points in any designed study should be kept high. Blanks may or may not be included, depending on the data-quality objectives of the test method.

4.1.3 In applying regression to any applicable data set, the proper fitting technique (for example, ordinary least squares (OLS) or weighted least squares (WLS)) must be determined (for fitting the proposed model to the data).

4.1.4 The residual pattern and the lack-of-fit (LOF) test are used to evaluate the adequacy of the chosen model.

4.1.5 The magnitude of the half-width of the prediction interval must be evaluated, remembering that accepting or rejecting the amount of uncertainty is a judgment call, not a statistical decision.

5. Significance and Use

5.1 Appropriate application of this practice should result in an estimate of the test-method’s uncertainty (at any concentration within the working range), which can be compared with data-quality objectives to see if the uncertainty is acceptable.

5.2 With data sets that compare recovered concentration with true concentration, the resulting regression plot allows the correction of the recovery data to true values. Reporting of such corrections is at the discretion of the user.

5.3 This practice should be used to estimate the measurement uncertainty for any application of a test method where measurement uncertainty is important to data use.

6. Procedure

6.1 Introduction:

6.1.1 For purposes of this practice, only regression-based methods are applicable. An example of a module that is not regression-based is a balance. If an object is placed on a balance, the readout is in the desired units; that is, in units of mass. No user intervention is required to get to the needed result. However, for an instrument such as a chromatograph or a spectrometer, the raw data (for example, peak area or absorbance) must be transformed into meaningful units, typically concentration. Regression is at the core of this transformation process.

6.1.2 One additional distinction will be made regarding the applicability of this protocol. This practice will deal only with intralaboratory data. In other words, the variability introduced by collecting results from more than one lab is not being considered. The examples that are shown here are for one method with one operator. If the user wishes, additional operators may be included in the design, to capture multiple-operator variability.

6.1.3 A brief example will help illustrate the importance of estimating measurement uncertainty. A sample is to be analyzed to determine if it is under the upper specification limit of 5 (the actual units of concentration do not matter). The final test