Designation: D3670 – 91 (Reapproved 2022)

Standard Guide for Determination of Precision and Bias of Methods of Committee D22[1]

This standard is issued under the fixed designation D3670; 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 standard provides guidance to task groups of Committee D22 on Sampling and Analysis of Atmospheres in planning and conducting collaborative testing of candidate methods.

1.2 It is intended for use with other ASTM practices for the determination of precision and bias.

1.3 It is applicable to most manual and automated methods and to most components of monitoring systems. It is recognized that the evaluation of monitoring systems may provide special problems. Practice D3249 should be considered for general guidance in this respect.

1.4 It is directly applicable to chemical methods and in principle to most physical methods, sampling methods, and calibration procedures.

1.5 The processes described are for the general validation of methods of test. A user has the obligation and responsibility to validate any method it uses for a specific application and to demonstrate its own competence in the use of validated methods.

1.6 *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]

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

D3249 Practice for General Ambient Air Analyzer Procedures

E177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods

E180 Practice for Determining the Precision of ASTM Methods for Analysis and Testing of Industrial and Specialty Chemicals (Withdrawn 2009)[3]

E691 Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method

E1169 Practice for Conducting Ruggedness Tests

3. Terminology

3.1 The terms used in this practice are consistent with those defined in Practices D2777, E177, E180, and E691.

3.2 *Definitions:*

3.2.1 *accuracy, n*—the degree of conformity of a value generated by a specific procedure to the assumed or accepted true value. It includes both precision and bias.

3.2.2 *bias, n*—a systematic (nonrandom) deviation of the method average value or the measured value from an accepted reference value.

3.2.3 *candidate method, n*—an analytical method or measurement process being considered for standardization. A method is a "candidate" until completion of all phases of the consensus process specified by ASTM regulations for a proposal, an emergency standard, or a standard.

3.2.4 *collaborative test, n*—an interlaboratory study of a test method wherein the participants analyze or make measurements on sub-samples of the same test material. If the test method includes the sampling of atmospheres, the participants should sample the same test atmosphere, as possible.

3.2.5 *laboratory bias, n*—systematic differences between the true value and a value reported by a laboratory due to errors of application such as losses, contamination, miscalibration, and faulty manipulations, for example.

3.2.6 *method bias, n*—systematic departures of the limiting mean from the true value of the parameter measured, caused by physical or chemical phenomena inherent in the methodology.

[1] This guide is under the jurisdiction of ASTM Committee D22 on Air Qualityand is the direct responsibility of Subcommittee D22.01 on Quality Control.

Current edition approved Sept. 1, 2022. Published September 2022. Originally approved in 1978. Last previous edition approved in 2014 as D3670 – 91 (2014). DOI: 10.1520/D3670-91R22.

[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.

[3] The last approved version of this historical standard is referenced on www.astm.org.

3.2.7 *over-all precision, n*—a value including components of within-laboratory and between-user variability.

3.2.8 *precision, n*—the degree of mutual agreement between individual measurements using an analytical method or measurement process. In practice, the standard deviation of an entire array of reviewed and acceptable data is calculated to provide the value to be stated as the precision of the method.

3.2.9 *ruggedness test, n*—a factorial test designed to explore the sensitivity of the method to variations in the procedure (see Youden and Steiner, 1987).[4]

3.2.10 *single-operator precision, n*—a measure of the replication of repeated measurements obtained by a single operator on a given sample.

3.2.10.1 *Discussion*—Other classifications of precision which are useful in evaluating a method, a measurement, or performance within a single laboratory are: multioperator precision, single or multi-apparatus precision, and single or multi-day precision.

3.2.10.2 *Discussion*—The terms "repeatability" and "reproducibility" are not standardized, but have generally come to mean "single-laboratory-operator-material precision" and "multi-laboratory-multi-operator-single material precision," respectively. Such usage is maintained in the text of this practice.

3.2.10.3 *Discussion*—Further classifications of bias which are useful in evaluating performance are: operator bias, apparatus bias, and day bias.

4. Summary of Guide

4.1 Data supporting a statement of single-operator repeatability is the entrance requirement for any candidate method to be considered for standardization by Committee D22. The task group to which a candidate method is assigned will review it for adequacy in this respect, and conduct further tests as necessary to evaluate its precision and bias, as technically feasible. A method may be accepted as a proposed method, provided the repeatability is known or has been ascertained and provided all other criteria for acceptance have been met. Independent tests by at least three laboratories shall be required to substantiate the repeatability of a method before it attains the status of a standard method. Collaborative testing by at least five laboratories to estimate the interlaboratory bias and, if applicable to evaluate the method's inherent bias with respect to the "true" value is needed for all standard methods and must be accomplished within 5 years of its initial issuance as a standard, if such testing has not already been done. Failure to subject such methods to appropriate collaborative testing, constitutes valid grounds for disallowing its reapproval as a standard.

4.2 Procedures that may be used in collecting the required data are given with particular emphasis upon the applicability to analysis of atmospheres. Documentation requirements are established. Terms that are useful in expressing statements of precision and bias are presented.

[4] Youden, W. J. and Steiner, G. H., *Statistical Manual of the Association of Official Analytical Chemists*, AOAC International, 481 North Frederick Ave., Suite 500, Gaithersburg, MD 20877-2417, 1987.

5. Significance and Use

5.1 The objective of this standard is to provide guidelines to Committee D22 for the evaluation of the precision and bias, or both, of ASTM standard methods and practices at the time of their development. Such an evaluation is necessary to assure that a cross section of interested laboratories could perform the test and achieve satisfactory results, using the method as written. It also provides guidance to the user as to what levels of precision and accuracy may be expected in such usage.

5.2 The write-up of the method describes the media for which the test method is believed to be appropriate. The collaborative test corroborates the write-up within the limitations of the test design. A collaborative test can only use representative media so that universal applicability cannot be implied from the results.

5.3 The fundamental assumption of the collaborative test is that the media tested, the concentrations used, and the participating laboratories are representative and provide a fair evaluation of the scope and applicability of the test method as written.

6. General Policy

6.1 This section describes the general policy to be followed by Committee D22, its subcommittees, and task groups in the development of ASTM standard methods and practices. The objective of Committee D22 is to develop fully evaluated standard methods and practices as far as possible. In cases where this is not expedient, proposed methods, as defined in 6.2, may be developed. In each case, an appropriate task group shall have the responsibility to critically examine the method or practice, conduct evaluation tests by round robins or other techniques including ruggedness tests, and to recommend it, if meritorious, for subcommittee balloting. No method or practice shall be released and recommended for balloting unless the precision or accuracy requirements, or both, as set forth in the following, have been satisfied.

6.1.1 Collaborative testing by D22 is the preferred method of validation. Data obtained by collaborative testing by others may be used in lieu of D22 testing, provided that such testing was equivalent to ASTM approved procedures. In either case, a copy of the test procedures and data must be filed in a research file maintained at ASTM for such purposes.

6.2 *Proposed Method*—A proposed method is one that has found favorable usage in a specific laboratory, or has been used by several laboratories, but has not yet been standardized. In each case, the test method is submitted by its proponents to Committee D22 for standardization.

6.2.1 The minimum requirement for balloting of a proposed method shall be the inclusion in it of a single laboratory's statement of single-operator precision, together with supporting experimental data. Test methods meeting this requirement will be referred to a Task Group, following procedures established by Committee D22.

6.2.2 The experimental data needed to support a proposal must reflect a test of the method as a whole, that is, sampling, apparatus, reagents and, calibration, and must use a procedure that is essentially identical to that described in the proposal.

Any significant deviations between the procedure used to gather the data and the proposed procedure shall be clearly identified.

6.2.3 If such data are missing or inadequate, but the method itself is considered by consensus of Committee D22 to be worthy of further study, a task group may be assigned to conduct experimental studies or enlist the services of at least one competent laboratory to obtain the data upon which to base a statement of single-operator precision.

6.3 *Standard Method–Initial Acceptance*—A method that has found favorable acceptance and for which the within-laboratory repeatability has been verified by a multilaboratory test program, shall be examined by the task group for compliance with the following requirements.

6.3.1 An initial minimum requirement for establishing a standard method is a statement of within-laboratory precision based on data from three laboratories similar to that described in 6.2.1 – 6.2.3.

6.3.2 If the method purports to measure the concentration of a substance, an investigation of the bias of the method by comparison with a standard must be made by at least one laboratory and the results included in an accuracy statement.

6.3.3 A standard can only be carried under the provisions of 6.3 for five years. Conditions for reapproval are specified in 6.4.

6.4 *Standard Method–Reapproval*—A standard method may be retained if it has found extensive use and between-laboratory precision data have been obtained. Before doing a collaborative study, a ruggedness test should be performed by at least one laboratory (see Guide E1169).

6.4.1 The minimum requirement for retaining a standard method shall be a statement of the between-laboratory precision of the method as established in a collaborative test including at least five participants.

6.4.2 If a bias statement is appropriate for the method, the data supporting the statement should be obtained by at least two laboratories. At least one such test shall include the introduction of potential interferences.

6.5 In all testing, the minimum number of participants should be exceeded to the extent possible. The statistical power of collaborative testing is greatly enhanced as such numbers are increased. The possibility of invalidation of a test by outliers or missing data is also minimized.

7. Sample Requirements

7.1 The precision and bias of test methods are typically evaluated by the data obtained in the measurement of test samples. The extent to which such measurements can be made is dependent upon the availability of test samples of adequate stability and homogeneity. The scope of interest of Committee D22 is wide, ranging from contaminants at the parts-per-billion level up to several percent. Particulate concentrations exist at similar concentration ranges and measurements of radioactivity extend the level even lower. The variety of substances of interest range from simple inorganic constituents to complex organic molecules. Accordingly, it is not possible to set forth rigid sample specifications, but only to delineate guidelines for test sample preparation. Each method should be tested with actual samples for which it is applicable, or as close a simulation as possible. The degree of evaluation will, of course, depend on the simulation achieved, and the statements of precision and accuracy must define the test conditions.

7.2 The ideal test sample is the actual atmosphere for which the method is intended. However the use of such offers complications because the composition may not be known at the moment of test and furthermore may undergo change during the tests. Because actual atmospheric samples cannot be collected and stabilized for long periods of time, two procedures are acceptable. Reproducibility and repeatability may be evaluated by simultaneous measurement by participating laboratories sampling the same atmosphere at substantially the same time. Alternatively, comparison of a candidate method with a standard method of known precision and bias will constitute an acceptable technique for evaluation of precision and accuracy. Such measurements made by several laboratories may be statistically treated to evaluate the reproducibility of the candidate method. In this latter case, the measurements need not be made at the same place and time by the collaborating laboratories.

7.3 A test sample or series of test samples that are stable during the period required to perform a limited series of measurements are adequate for evaluation of single-operator precision to satisfy the requirements for consideration as a proposed method. Three levels of concentration are recommended, with such levels sufficiently well established to determine whether, and to what extent, the repeatability is dependent or independent of concentration level.

7.4 A series of test samples of at least three concentration levels, and available in sufficient number, is required for use by collaborating laboratories to evaluate the repeatability and reproducibility of a candidate method. The samples should be stable during the entire test period, which should include a reasonable time following the collaborative test to permit resolution of any discrepancies encountered during the evaluation procedures. The compositions of the test samples do not need to be known accurately, but the samples furnished to each collaborator must be sufficiently similar to permit evaluation of measurement errors in excess of compositional inhomogeneity. The test samples for repeatability and reproducibility should closely simulate actual source or atmospheric air compositions, including the presence of any known interferents. The statistical statements must reflect the type of test sample for which the precision or bias, or both, are specified. The statement should include the concentration levels studied and the number of laboratories participating.

7.5 Accuracy tests to determine the inherent bias of an analytical method are preferably made under rigorously controlled laboratory conditions utilizing standards of known composition.

7.6 In the absence of samples of known composition, the use of the spiking technique in which standard additions of known constituents are made by established techniques will be acceptable for evaluating the bias of candidate methods. In