Designation: C 670 – 96

Standard Practice for Preparing Precision and Bias Statements for Test Methods for Construction Materials[1]

This standard is issued under the fixed designation C 670; 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.

This standard has been approved for use by agencies of the Department of Defense.

1. Scope

1.1 This practice supplements Practice E 177, in order to provide guidance in preparing precision and bias statements for ASTM test methods pertaining to certain construction materials (Note 1). Recommended forms for precision and bias statements are included. A discussion of the purpose and significance of these statements for the users of those test methods is also provided.

NOTE 1—Although under the jurisdiction of Committee C-9, this practice was developed jointly by Committees C-1, D-4, and C-9, and has been endorsed by all three committees. It has subsequently been adopted for use by Committee D-18.

2. Referenced Documents

2.1 *ASTM Standards:*

C 109/C 109M Test Method for Compressive Strength of Hydraulic Cement Mortars (Using 2-in. or 50-mm Cube Specimens)[2]

C 802 Practice for Conducting an Interlaboratory Test Program to Determine the Precision of Test Methods for Construction Materials[3]

E 177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods[4]

3. Terminology

3.1 *Definitions of Terms Specific to This Standard:*

3.2 *one-sigma limit (1s)*—the fundamental statistic underlying all indexes of precision is the standard deviation of the population of measurements characteristic of the test method when the latter is applied under specifically prescribed conditions (a given system of causes). The terminology "one-sigma limit" (abbreviated (1s)) is used in Practice E 177 to denote the estimate of the standard deviation or sigma that is characteristic of the total statistical population. The one-sigma limit is an indication of the variability (as measured by the deviations above and below the average) of a large group of individual test results obtained under similar conditions.

3.2.1 *single-operator one-sigma limit*—the one-sigma limit for single-operator precision is a quantitative estimate of the variability of a large group of individual test results when the tests have been made on the same material by a single operator using the same apparatus in the same laboratory over a relatively short period of time. This statistic is the basic one used to calculate the single-operator index of precision given in the precision statement for guidance of the operator.

3.2.2 *multilaboratory one-sigma limit*—the one-sigma limit for multilaboratory precision is a quantitative estimate of the variability of a large group of individual test results when each test has been made in a different laboratory and every effort has been made to make the test portions of the material as nearly identical as possible. Under normal circumstances the estimates of one-sigma limit for multilaboratory precision are larger than those for single-operator precision, because different operators and different apparatus are being used in different laboratories for which the environment may be different.

3.2.3 *one-sigma limit in percent (1s%)*—in some cases the coefficient of variation is used in place of the standard deviation as the fundamental statistic. This statistic is termed the "one-sigma limit in percent" (abbreviated (1s%)) and is the appropriate standard deviation (1s) divided by the average of the measurements and expressed as a percent. When it is appropriate to use (1s%) in place of (1s) is discussed in Section 6.

3.3 *Acceptable Range of Results*:

3.3.1 *acceptable difference between two results*—the "difference two-sigma limit (d2s)" or "difference two-sigma limit in percent (d2s%)," as defined in Practice E 177, has been selected as the appropriate index of precision in most precision statements. These indexes indicate a maximum acceptable difference between two results obtained on test portions of the same material under the applicable system of causes described in 4.1.1 and 4.1.2 (or whatever other system of causes is appropriate). The (d2s) index is the difference between two individual test results that would be equaled or exceeded in the long run in only 1 case in 20 in the normal and correct operation of the method. The (d2s%) index is the difference between two individual test results expressed as a percent of their average that meets the same requirements. These indexes

[1] This practice is under the jurisdiction of ASTM Committee C-9 on Concrete and Concrete Aggregatesand is the direct responsibility of Subcommittee C09.94on Evaluation of Data.

Current edition approved May 10, 1996. Published July 1996. Originally published as C 670 – 71 T. Last previous edition C 670 – 95.

[2] *Annual Book of ASTM Standards*, Vol 04.01.

[3] *Annual Book of ASTM Standards*, Vol 04.02.

[4] *Annual Book of ASTM Standards*, Vol 14.02.

are calculated by multiplying the appropriate standard deviation (1s) or coefficient of variation (1s%) by the factor $2\sqrt{2}$ (equal to 2.83).

3.3.2 *acceptable range of more than two results*—in cases where the test method calls for more than two test results to be obtained, the range (difference between highest and lowest) of the group of test results must be compared to a maximum acceptable range for the applicable system of causes and number of test results. The range for different numbers of test results including two that would be equaled or exceeded in only 1 case in 20 is obtained by multiplying the appropriate standard deviation (1s) or coefficient of variation (1s%) by the appropriate factor from the second column of Table 1 (Note 2):

NOTE 2—It is important to note that when more than two test results are obtained, an index of precision for the difference between two results can not be used as a criterion for judging acceptability of the range of the group or for other pairs of results selected from the group.

3.3.3 *variations for single operators*—the system of causes designated for obtaining the quantitative guide to acceptable performance by an operator as stated in 4.1.1 leads to single-operator precision, using the system of modifiers given in Practice E 177 (Note 3). When two results by the same operator differ by more than (d2s) or (d2s%) or the range of more than two results exceeds that obtained by the method described in 3.2.2 there is a significantly large probability that an error has occurred and retests should be made as directed in Note 4.

NOTE 3—Single-operator precision is often referred to as "repeatability," and multilaboratory precision is often referred to as "reproducibility."

NOTE 4—It is beyond the scope of this practice to describe in detail what action should be taken in all cases when results occur that differ by more than the (d2s) limits or by more than the maximum allowable range. Such an occurrence is a warning that there may have been some error in the test procedure, or some departure from the prescribed conditions of the test on which the limits appearing in the test method are based; for example, faulty or misadjusted apparatus, improper conditions in the laboratory, etc. In judging whether or not results are in error, information other than the difference between two test results is needed. Often a review of the circumstances under which the test results in question were obtained will reveal some reason for a departure. In this case the data should be discarded and new test results obtained and evaluated separately. If no physical reason for a departure is found, retests should still be made, but the original tests should not be completely ignored. If the second set of results also differs by more than the applicable limit, the evidence is very strong that something is wrong or that a real difference exists between the two samples tested. If the second set produces a result within the limit, it may be taken as a valid test, but the operator or laboratory may then be suspected of producing erratic results, and a closer examination of the procedures would be in order. If knowledge about the test method in question indicates that certain actions may be appropriate in cases where deviant results occur, then such information should be included in the test method, but details of how this should be done will depend upon the particular test method.

TABLE 1 Maximum Acceptable Range

Number of Test Results	Multiplier of (1s) or (1s%) for Maximum Acceptable Range[A]
2	2.8
3	3.3
4	3.6
5	3.9
6	4.0
7	4.2
8	4.3
9	4.4
10	4.5

[A] Values were obtained from Table A7 of "Order Statistics and Their Use in Testing and Estimation," Vol 1, by Leon Harter, Aerospace Research Laboratories, United States Air Force.

3.3.4 *variations between laboratories*—the system of causes designated for obtaining the quantitative guide for acceptance of results by different laboratories as given in 4.1.2 is multilaboratory precision, using the system of modifiers given in Practice E 177 (Note 3). When results differ by more than (d2s) there is a significantly large probability that one or both laboratories are in error or that a difference exists in the portions of material being used for the tests. In such cases, retests should be made. When possible, newly drawn test samples should be used for such retests as directed in Note 4.

3.4 *Number of Tests*:

3.4.1 *single test results*—the number of tests run must be taken into account when evaluating testing variations. Usually, the statistics used in evaluating precision and the indexes of precision based on them are based on the population distribution of single test results. When this is the case, the index of precision may be used in comparing single tests results only, not averages of two or more tests.

3.4.2 *test results based on averages*—if the precision statement is based on test results that are averages of two or more measurements, then the number of measurements averaged must be stated, and in using the index of precision, averages of exactly that number of measurements must be used. In some cases a test result is defined in the method as the average of two or more individual measurements. In such cases the index of precision for a test result applies to a test result as so defined, although indexes of precision for ranges of individual measurements within a laboratory may also be included as described in 3.3.3.

3.4.3 *precision of individual measurements averaged to obtain a test result*—when two or more measurements are averaged to obtain a test result, the range of the individual measurements may be examined to determine whether the latter meet the criterion of being valid individual measurements under the conditions of the test method. The maximum acceptable range for individual measurements is obtained by multiplying the appropriate standard deviation (1s) or, coefficient of variation (1s%) obtained from averages by the appropriate factor from the second column of Table 2 (Note 5). The maximum acceptable range for individual measurements obtained by this method may be included in the precision statement as an index of precision for individual measurements in the same laboratory as described in Example 8.

NOTE 5—This procedure is only valid if the individual measurements are subject to the same sources of variation as the test result. For example, the single-operator precision of Test Method C 109/C 109M mortar cubes is calculated from test results that include a contribution from variation among batches of mortar. Variation among individual cubes from a single batch does not contain this component of variation. Therefore, differences among individual cubes from a single batch cannot be inferred from the single-operator standard deviation given in Test Method C 109/C 109M and the values in Table 2.

3.4.4 *multilaboratory precision expressed as a maximum*

TABLE 2 Maximum Acceptable Range of Individual Measurements

Number of Measurements Averaged to Obtain a Test Result	Multiplier of (1s) or (1s%) for Averages to Obtain Maximum Acceptable Range of Individual Measurements[A]
2	3.9
3	5.7
4	7.3
5	8.6
6	9.9
7	11.0
8	12.1
9	13.2
10	14.1

[A] Values were calculated from Table 1.

allowable difference between two averages—when the test method calls for the reporting of more than one test result, multi-laboratory precision may be expressed as a maximum allowable difference between averages of such groups, one from each laboratory, and both the (d2s) or (d2s%) limit for individual results and this maximum allowable difference of two averages may be included in the multilaboratory precision statement (Note 6). The maximum allowable difference for averages of a given number of test results, *n*, is obtained by dividing the appropriate (d2s) or (d2s%) limit by the square root of *n*.

NOTE 6—Note that this is not the same as the situation where a test result is defined as the average of two or more individual measurements. A given test method may include both features. It is important to bear in mind, however, that when more than one result is obtained in one or both laboratories, the (d2s) or (d2s%) limit may not be used as a criterion for judging the differences between selected pairs of results from the two laboratories.

3.5 *field versus laboratory tests*—precision indexes for ASTM test methods are normally based on results obtained in laboratories by competent operators using well-controlled equipment on test portions of materials for which precautions have been taken to ensure that they are as nearly alike as possible. Such precautions and the same level of competence may not be practicable for the usual quality control or routine acceptance testing. Therefore, the normal testing variation among laboratories engaged in quality control and acceptance testing of commercial materials may be larger than indicated by the relationship derived from the one-sigma limit for multilaboratory precision. In this case it is recommended that studies be made to determine the one-sigma limit for tests made under field conditions and realistic adjustments in specification tolerances be made accordingly.

4. General Concepts

4.1 A precision statement meeting the requirements of this practice normally contains two main elements described as follows:

4.1.1 *Single-Operator Precision*—A measure of the greatest difference between two results that would be considered acceptable when properly conducted repetitive determinations are made on the same material by a competent operator.

4.1.2 *Multilaboratory Precision*—A measure of the greatest difference between two test results that would be considered acceptable when properly conducted determinations are made by two different operators in different laboratories on portions of a material that are intended to be identical, or as nearly identical as possible.

4.2 *Other Measures of Precision*—The two elements described in 4.1.1 and 4.1.2 involve the main systems of causes of interest to users of test methods involving construction materials. In cases where other systems of causes apply, the appropriate statistics for those systems should be used and the appropriate combination of modifiers given in Practice E 177 should be used to describe those statistics.

4.3 *Use of Indexes of Precision in Specifications*—The indexes of precision described in this practice are to be used as guides to determine (with a prescribed degree of certainty) whether a given series of results can be considered as valid tests under the conditions assumed in the test method. Comparisons of test results with specification limits should be made only after there is reasonable assurance that the determinations are adequate. Writers of specifications have the responsibility of recognizing the variability of results characteristic of a given test method in setting specification limits, but indexes of precision of the test method should never be added to specification limits by the users of those specifications for the purpose of judging acceptance or rejection of materials.

4.4 *Use of Indexes of Precision for Qualifying an Operator*—Indexes of single-operator precision are sometimes used as a basis for qualifying an operator. The assumption is that results that do not differ by more than the stated index are indicative of proper performance of the test. However, this assumption is not necessarily correct. Uniform misunderstanding of instructions or maladjustments of equipment may produce consistent but erroneous test results. Thus, tests conducted for the purpose of qualifying an operator should be made on materials for which the measured characteristic is known, whenever possible, so that accuracy as well as precision can be evaluated. (See Practice E 177 for a discussion of the terms precision and accuracy.)

5. Basis for Precision Statement

5.1 In order to be valid the indexes of precision to be included in the precision statement as guides for the operator must be based on estimates of the precision of the test method obtained from a statistically designed interlaboratory series of tests. This series of tests must involve a sufficient number of laboratories, materials, and replicate measurements so that the results obtained provide reliable estimates of the true precision characteristic of the test method (Note 7). The procedures described in this practice are based on the assumption that the proper estimates of precision have already been obtained. Practice C 802 is a companion document to this one and describes techniques for conducting an interlaboratory study to obtain the needed estimates of precision. In the case where an approved standard test method is revised, the subcommittee having responsibility over the test method should determine whether the change(s) affect the validity of the existing precision statement in the standard; and if so, should also revise the precision statement accordingly.

NOTE 7—The requirement of “reliable estimates of the true precision”

presupposes an estimate obtained from a properly designed and executed interlaboratory series of tests involving at least 30 degrees of freedom for single-operator precision and at least 10 laboratories.

5.2 For many of the tests under the jurisdiction of Committees C-1, C-9, D-4 and D-18, there is an extensive backlog of interlaboratory test data in the reference sample program of the Cement and Concrete Reference Laboratory (CCRL) and the AASHTO Materials Reference Laboratory (AMRL). Where such data are available, a precision statement can be prepared for each test method based upon a much larger population of data than can normally be assembled in a round-robin program by merely carrying out the mathematical analysis like that illustrated in Appendix X1.

5.3 The Form and Style for ASTM Standards requires that data and details of the experiments used to determine precision and bias be filed as a research report at ASTM Headquarters.

6. Form of Precision Statement

6.1 *Preface Information*—The Form and Style for ASTM Standards requires that the precision and bias statement include the reference numbers of the research report (paragraph 5.3) and a brief description of the experiments that will permit the user of the test method to judge the reliability of the data. Many precision and bias statements are based on non-SI data that have been converted to SI units. The following examples provide recommended wording for the preface to the precision and bias statement.

6.1.1 *Case 1*—Precision is stated in terms of percentage, such as coefficient of variation. The precision indices are non-dimensional and there would be no need for dual presentations. In this case, it is only necessary to state that the data were obtained in the inch-pound system.

Example 1:

The data used to develop the precision statement were obtained using the/(an earlier) version of this Test Method.

6.1.2 *Case 2*—For a combined standard in which both systems of units are to be used separately:

Example 2:

A. *Inch-pound (SI)*—The data used to develop the precision statement were obtained using the inch-pound version of this Test Method. The precision indices shown in parentheses are exact conversions of the values in inch-pound units.

B. *SI (inch-pound)*—The data used to develop the precision statement were obtained using the inch-pound version of this Test Method. The precision indices shown in SI units are exact conversions of the values in parentheses.

6.1.3 *Case 3*—For a standard that has been hard converted to SI units as standard and the inch-pound units are shown in parentheses for information only:

Example 3:

The data used to develop the precision statement were obtained using the previous inch-pound version of this Test Method. The precision indices are exact conversions of the values shown in parentheses.

6.1.4 *Case 4*—For a standard that has been converted to an SI standard and the inch-pound units have been dropped.

Example 4:

The data used to develop the precision statement were obtained using the previous inch-pound version of this Test Method. The indicated precision indices are exact conversions of the values obtained originally in inch-pound units.

6.2 *Manner of Expression*—If the test data on which the precision statement is to be based indicate that the standard deviation is essentially the same for all levels of the property being tested for which data are available, the one-sigma limit and the difference two-sigma limit shall be given in the precision statement expressed in the units of the measured property.

6.2.1 If the standard deviation is essentially proportional to the average for different levels of the property in question (that is, the coefficient of variation is essentially constant) then the "one-sigma limit in percent" (1s%) and difference two-sigma limit in percent (d2s%) shall be given. "One-sigma limit in percent" is, for the purposes of this practice, the same as the coefficient of variation. It is determined by dividing the standard deviation by the mean (average) value of available results and multiplying by 100. Similarly, "difference two-sigma limit in percent" is obtained by dividing (d2s) by the mean and multiplying by 100. When neither of these conditions is met, the applicable limits for specific ranges of the property shall be stated together with the specific ranges for which they are appropriate. The abbreviations (1s), (1s%), (d2s), and (d2s%) are given in footnotes as shown in the examples.

6.3 *Recommended Form of the Precision Statement*—When the proper estimates of precision are available (Note 7), the precision statement shall be written in the form of the appropriate example as given below for each available estimate of precision (standard deviation or coefficient of variation) and corresponding system of causes.

NOTE 8—Some of the following examples have been taken from test methods current at the time this practice was written and others are hypothetical. None of the examples should be taken as being quantitatively correct, since, even if taken from actual situations, the figures may have been subsequently revised.

6.3.1 *Form of Statements for Which One Estimate of Precision for Each System of Causes Applies*:

Example 1:

Precision—The multilaboratory standard deviation has been found to be 0.75 %[A]. Therefore, results of two properly conducted tests from two different laboratories on samples of the same cement should not differ by more than 2.1 %.[A]

[A] These numbers represent, respectively, the (1s) and (d2s) limits as described in ASTM Practice C 670, for Preparing Precision Statements for Test Methods for Construction Materials.

Example 2:

Precision—The single-operator standard deviation has been found to be 0.045 %.[A] Therefore, results of two properly conducted tests by the same operator on the same material should not differ by more than 0.13 %.[A]

[A] These numbers represent, respectively, the (1s) and (d2s) limits as described in ASTM Practice C 670, for Preparing Precision Statements for Test Methods for Construction Materials.

Example 3:

Precision—The multilaboratory coefficient of variation has been found to be 5.0 %.[A] Therefore, results of two different laboratories on identical samples of a material should not differ from each other by more than 14 % of their average.[A]

[A] These numbers represent, respectively, the (1s) and (d2s) limits as described in ASTM Practice C 670, for Preparing Precision Statements for Test Methods for Construction Materials.

Example 4:

The single-operator coefficient of variation has been found to be 2.5 %.[A] Therefore, results of two properly conducted tests by the same operator on the same sample using the same viscometer should not differ from each other by more than 7.0 % of their average.[A]

[A] These numbers represent, respectively, the (1s) and (d2s) limits as described in ASTM Practice C 670, for Preparing Precision Statements for Test Methods for Construction Materials.

6.3.2 *Form of Statements for Which the Precision, Measured by Either the Standard Deviation or the Coefficient of Variation, is not Constant over the Range of Values of the Property in Question*:

6.3.2.1 If the precision limit given applies only over a certain range of the property of the material being measured, this shall be indicated by inserting the words "over the range from __ to __ , or "below", or "above" a certain limit after the words "standard deviation" or "coefficient of variation" in the first sentence of the statement. If precision limits have been obtained for more than one range of the property, separate statements shall be written for each range. The applicable range should also be indicated in subparagraph headings if separate subparagraphs are used as follows:

Example 5:

Single-Operator Precision—The single-operator standard deviation has been found to be 1.4°F (0.8°C)[A] for flash points below 220°F (104°C) and 7.1°F (3.9°C)[A] for flash points above 220°F. Therefore, results of two properly conducted tests by the same operator on the same material should not differ from each other by more than 4°F (2.2°C)[A] for flash points below 220°F or by more than 20°F (11.1°C)[A] for flash points above 220°F.

[A] These numbers represent, respectively, the (1s) and (d2s) limits as described in ASTM Practice C 670, for Preparing Precision Statements for Test Methods for Construction Materials.

Example 6:

Multilaboratory Precision—The multilaboratory standard deviation has been found to be 2.1°F (1.2°C)[A] for flash points below 220°F (104°C) and 8.8°F (4.9°C)[A] for flash points above 220°F. Therefore, results of two properly conducted tests on the same material in two different laboratories should not differ from each other by more than 6°F (3.3°C)[A] for flash points below 220°F or by more than 25°F (13.9°C)[A] for flash points above 220°F.

[A] These numbers represent, respectively, the (1s) and (d2s) limits as described in ASTM Practice C 670, for Preparing Precision Statements for Test Methods for Construction Materials.

6.3.2.2 If the precision, whether expressed as an absolute limit or a percent, is not constant over the given range or for all the materials tested, but the limit given is the maximum value of the index of precision, the abbreviation "max" shall be inserted after the closing parenthesis of the abbreviation for the type of limit in the footnote: that is (1s) max, or (1s%) max. Also the word" maximum" shall be inserted in the first sentence of the precision statement. This form should rarely be used, and then only as a last resort. See the Irregular or Nonlinear Relationship Between Standard Deviation, Coefficient of Variation and Average Level section in Practice C 802.

Example 7:

Precision—The maximum single-operator-machine-multibatch coefficient of variation has been found to be 4.25 %.[A] Therefore, results of two properly conducted tests by the same operator of concrete cylinders from two different batches should not differ from each other by more than 12 % of their average.[A]

[A] These numbers represent, respectively, the (1s%) and (d2s%) max limits as described in ASTM Practice C 670, for Preparing Precision Statements for Test Methods for Construction Materials.

6.3.2.3 See Examples 11 and 12 for alternative tabular form of precision statements.

6.4 *Form of Statement for Which a Test Result is Defined as the Average of a Specified Number of Measurements*:

Example 8:

Single-Operator Precision—The single-operator standard deviation of a single test result (where a test result is, as defined in this test method, the average of three separate measurements) has been found to be 2.0 %.[A] Therefore, results of two properly conducted tests (each consisting of the average of three individual measurements) should not differ by more than 5.7 %[A] and the range (difference between highest and lowest) of the three individual measurements used in calculating the average should not exceed 11.4 %.[B]

[A] These numbers represent, respectively, the (1s) and (d2s) limits as described in ASTM Practice C 670, for Preparing Precision Statements for Test Methods for Construction Materials.

[B] Calculated as described in 3.4.3 of Practice C 670.

6.5 *Form of Statements for Which More than One Test Result is Reported*:

Example 9:

Single-Operator Precision—The single-operator standard deviation of a single test result has been found to be 125 psi (861 kPa). Therefore, results of two properly conducted tests by the same operator should not differ by more than 350 psi (2413 kPa).[A] The test method calls for reporting three test results. The range (difference between highest and lowest) of the three test results obtained by the same operator should not exceed 410 psi (2827 kPa).[B]

[A] These numbers represent, respectively, the (1s) and (d2s) limits as described in ASTM Practice C 670, for Preparing Precision Statements for Test Methods for Construction Materials.

[B] Calculated as described in 3.4.3 of Practice C 670.

Example 10:

Multilaboratory Precision—The multilaboratory standard deviation of a single test result has been found to be 225 psi (1551 kPa). Therefore, results of two properly-conducted tests in different laboratories on the same material should not differ by more than 640 psi (4413 kPa).[A] The averages of three test results in two different laboratories should not differ by more than 370 psi (2551 kPa).[B]

[A] These numbers represent, respectively, the (1s) and (d2s) limits as described in ASTM Practice C 670, for Preparing Precision Statements for Test Methods for Construction Materials.

[B] Calculated as described in 3.4.3 of Practice C 670.

6.6 *Alternative Form of the Precision Statement*—In cases where separate statements for a number of different materials or a number of different levels of a property are involved, the form recommended in 6.2 may become cumbersome. In such cases, the statement may be written in table form in accordance with the following examples:

Example 11:

Precision—Criteria for judging the acceptability of solubility test results obtained by this method are given as follows:

NOTE 9—The figures given in Column 2 are the standard deviations that have been found to be appropriate for the materials and conditions of test described in Column 1. The figures given in Column 3 are the limits that should not be exceeded by the difference between the results of two properly conducted tests.

Material and Type Index	Standard Deviation[A]	Acceptable Range of Two Results[A]
Single-operator precision:		
Asphalts, solubility more than 99 %[B]	0.035	0.10
Tars, liquid grades[C]	0.11	0.31
Tars, semi-solid[A]	0.17	0.48

Multilaboratory precision:		
Asphalts, solubility more than 99 %[C]	0.090	0.26
Tars, liquid grades[C]	0.22	0.61
Tars, semi-solid[C]	0.83	2.34

[A] These numbers represent, respectively, the (1s) and (d2s) limits as described in ASTM Practice C 670, for Preparing Precision Statements for Test Methods for Construction Materials.
[B] Applicable when either carbon disulfide, carbon tetrachloride, trichloroethylene, or benzene are used.
[C] Applicable when carbon disulfide is used.

Example 12:
Precision—Criteria for judging the acceptability of viscosity test results obtained by this method are given as follows:

NOTE 10—The figures given in Column 2 are the coefficients of variation that have been found to be appropriate for the materials and conditions of test described in Column 1. The figures given in Column 3 are the limits that should not be exceeded by the difference between the results of two properly conducted tests expressed as a percent of their mean.

Material and Type Index	Coefficient of Variation (percent of mean)[A]	Acceptable Range of Two Results (percent of mean)[A]
Single-operator precision:		
Asphalt Cements at 275°F (135°C)	0.64	1.8
Liquid asphalts at 140°F (60°C):		
below 3000 cSt	0.53	1.5
3000 cSt and above	0.71	2.0
Multilaboratory precision:		
Asphalt cements at 275°F (135°C)	3.1	8.8
Liquid asphalts at 140°F (60°C):		
below 3000 cSt	1.06	3.0
3000 cSt and above	3.11	9.0

[A] These numbers represent, respectively, the (1s%) and (d2s%) limits as described in this Practice.

7. Bias Statement

7.1 Bias is a systematic error inherent in the test method that contributes to the difference between a population mean of the measurements or test results and an accepted reference or true value. In any test method, tolerances are placed on the accuracy of measuring equipment. All tests made with a given set of equipment which has an error within the permitted tolerance will produce results with a small consistent bias, but that bias is not inherent in the test method and is not included in the bias statement for the test method. There are two conditions which permit the bias of a test method to be estimated: (1) a standard reference sample of known value has been tested by the test method, and (2) the test method has been applied to a sample which has been compounded in such a manner that the true value of the property being measured is known, such as may be the case, for example, in a test for cement content of concrete. Judgment is required to determine whether a potential reference sample is suitable for the purpose. For example, a metal bar of accurately known physical properties might not be suitable for establishing the bias of a test for the corresponding concrete properties because the level of values may differ by an order of magnitude. When it is possible to examine bias, it is necessary to determine whether there are enough data to determine statistically that the mean of the test results is significantly different from the true value. When it is, an absolute measure of bias cannot be made, but confidence limits may be placed on the bias.

7.2 For most test methods there is no reference value available. In those cases a statement based on one of the following may be used:

Example 1:
Bias—The test method has no bias because the values determined can be defined only in terms of the test method.

Example 2:
Bias—Since there is no accepted reference material suitable for determining the bias in this test method, no statement on bias is made.

Example 3:
Bias—No justifiable statement can be made on the bias of this test method because (insert here the reason).

7.3 Where it is possible to determine if bias exists, proceed as follows.

7.3.1 Form at least 30 pairs of results in which X_1 is the known reference value and X_2 is the experimental value. Form the quotient

$$t = \frac{\bar{X}_2 - \bar{X}_1}{s/\sqrt{N}} \quad (1)$$

where:
$\bar{X}_1$ = the mean of the reference values,
$\bar{X}_2$ = the mean of the experimental values,
s = the standard deviation of the differences $(X_2 - X_1)$, and
N = the number of pairs.

This quotient has a t distribution with N−1 degrees of freedom. Reject the hypothesis that no bias exists if $t < t_{½\,½\,\alpha}$ or $t > t_{1-½\,\alpha}$. Usually α, the level of significance, will be taken as 0.05. For an α of 0.05 and a sample of 30 pairs, the above inequalities reduce to $t < -2.05$ or $t > 2.05$. Thus if the calculated value of t falls between −2.05 and 2.05, it is concluded that there is no bias.

7.3.2 Where the value of t falls in the rejection range, the confidence limits for bias are:

$$\bar{X}_2 - \bar{X}_1 + t_{1/2\,\alpha}\, s\sqrt{1/N} \text{ and } \bar{X}_2 - \bar{X}_1 + t_{1-1/2\,\alpha}\, s\sqrt{1/N} \quad (2)$$

NOTE 11—In the above expression the first value of t is always negative.

7.3.3 In some cases the bias may be a function of level of the quantity being measured. If differences in means for different levels are significantly different from each other, the above procedure may be applied to each such level.

7.3.4 Where a test for bias has been made, a statement based on one of the following may be made:

Example 1:
Bias—When experimental results are compared with accepted reference values (or known values from accurately compounded specimens), the test method is found to have no bias.

Example 2:

Bias—When experimental results are compared with accepted reference values (or known values from accurately compounded specimens), the bias of the test method is found with 95 % confidence to lie between 0.0062 and 0.0071.

Example 3:

Bias—When experimental results are compared with accepted reference values (or known values from accurately compounded specimens), the bias of the test method is found with 95 % confidence to lie between − 0.0004 and − 0.0001 in the range of 6 to 10 and between − 0.0006 and − 0.0002 in the range of 10 to 15.

APPENDIX

(Nonmandatory Information)

X1. USE OF CCRL AND AMRL REFERENCE SAMPLE RESULTS FOR PRECISION STATEMENTS

X1.1 *Introduction:*

X1.1.1 Where the test method has been applied to a reference sample distributed by CCRL or AMRL, use the data from Table 1, Summary of Results, analyze them in accordance with Practice C 802, and formulate precision statements conforming to the requirements of this practice.

X1.1.2 The example which follows uses data from Table 1, Summary of Results, of a CCRL reference sample report (shown in Fig. X1.1). Data from tests for AMRL are reported in the same format.

X1.2 *Use of Table I, CCRL Report:*

X1.2.1 The compressive strength data for Samples 35 and 36 are selected for illustrating the use of Table I of the CCRL Reference sample report. They provide four levels of the measured value.

X1.2.2 The procedure consists largely of selecting appropriate values from Table 1of the CCRL report (Fig. X1.1) and placing them in the correct location in the "Approximate Values (Upper 5 % Level) for the Ratio of Highest to Lowest Variance" Table of Practice C 802 (shown in Fig. X1.2).

NOTE X1.1—In this particular example it was necessary to select the most appropriate lines of data to use in CCRL Table I. For the three-day strength results only the complete set of data before elimination of outliers permitted a statistically valid estimate of within-laboratory standard deviation (or random error). Thus, no choice was possible. For the 7 day test, the random error could be calculated both with and without outliers. The data with outliers removed were considered preferable.

X1.2.3 Place the information indicated by the circles numbered 1, 2, 3, and 4 in Table I, in order, in the "Average" column of the "Approximate Values (Upper 5 % Level) for the Ratio of Highest to Lowest Variance" Table. Items are placed

CCRL REFERENCE SAMPLE PROGRAM

CEMENT SAMPLES NUMBER 35 AND NUMBER 36

PHYSICAL DETERMINATIONS, FINAL REPORT, APRIL 9, 1975

TABLE I SUMMARY OF RESULTS -CONTINUED

		NO OF LABS		SAMPLE NO. 35 AVERAGE	SAMPLE NO. 35 STAND DEV	SAMPLE NO. 35 C.V. 0/0	SAMPLE NO. 36 AVERAGE	SAMPLE NO. 36 STAND DEV	SAMPLE NO. 36 C.V. 0/0	RANDOM ERROR R.E.	RANDOM ERROR C.V.(1) 0/0	RANDOM ERROR C.V.(2) 0/0
AIR CONTENT	PRCNT	169		9.2846+00	1.1057+00	11.909	8.2030+00	1.6417+00	20.014			
AIR CONTENT	PRCNT	167	(10)	9.3329+00	1.0185+00	10.913	8.2269+00	1.6367+00	19.895			
AIR CONTENT	PRCNT	166	(11)	9.3518+00	9.9185-01	10.606	8.2337+00	1.6393+00	19.910			
AC MIX WATER	PRCNT	165		7.0581+01	2.3145+00	3.279	7.0788+01	2.3982+00	3.388	9.9974-01	1.416	1.412
AC MIX WATER	PRCNT	160	(12)	7.9402+01	1.8715+00	2.658	7.0651+01	2.0291+00	2.872	8.0402-01	1.271	1.267
AC MIX WATER	PRCNT	159	(13)	7.0360+01	1.7997+00	2.558	7.0610+01	1.9694+00	2.789	8.9764-01	1.276	1.271
AC FLOW	PRCNT	168		8.6762+01 (2)	3.5650+00 (16)	4.109 (1)	8.7446+01 (7)	3.6143+00 (12)	4.133 (5)	2.9680+00 (12)	3.421 (11)	3.394
COMP STR 3D	PSI	178		3.0361+03	2.4328+02	8.013	2.3404+03	2.2370+02	9.556	1.8689+02	8.156	7.995
COMP STR 3D	PSI	172	(14)	3.0471+03	2.1698+02	7.121	2.3194+03	1.8837+02	8.122			
COMP STR 3D	PSI	169	(15)	3.0542+03	2.0460+02	6.699	2.3135+03	1.7642+02	7.626			
COMP STR 7D	PSI	178		4.2346+03	2.8097+02	6.635	3.5026+03	2.7698+02	7.908	1.9084+02	4.507	5.448
COMP STR 7D	PSI	171	(16)	4.2525+03 (4)	2.4581+02 (15)	5.780 (3)	3.4897+03 (9)	2.3564+02 (17)	6.753 (8)	1.4513+02 (14)	3.413 (13)	4.159
COMP STR 7D	PSI	170	(17)	4.2570+03	2.3923+02	5.620	3.4976+03	2.3479+02	6.712	1.3205+02	3.107	3.786
FINENESS AP SQCM/G		172		3.6056+03	8.9951+01	2.495	3.4319+03	8.7340+01	2.545	4.0536+01	1.374	1.443
FINENESS AP SQCM/G		164	(14)	3.6081+03	7.8506+01	2.176	3.4335+03	7.2471+01	2.111	3.9767+01	1.102	1.158
FINENESS WT SQCM/G		80		2.0221+03	1.0426+02	5.156	1.8350+03	9.6496+01	5.259	5.0305+01	2.488	2.741
FINENESS WT SQCM/G		79	(19)	2.0178+03	9.7608+01	4.837	1.8308+03	8.9296+01	4.878	5.0626+01	2.509	2.765
FINENESS WT SQCM/G		78	(20)	2.0212+03	9.3684+01	4.635	1.8342+03	8.4339+01	4.598	5.0943+01	2.529	2.777
NO 325 SIEV PER CT		131		9.0878+01	1.2455+00	1.370	8.7406+01	1.5049+00	1.722			

(10) FOLLOWING LABS ELIMINATED 179 304

(11) FOLLOWING LABS ELIMINATED 246
179 304

(12) FOLLOWING LABS ELIMINATED 21 59 179 304 375

(13) FOLLOWING LABS ELIMINATED 305
21 59 179 304 375

(14) FOLLOWING LABS ELIMINATED 28 29 45 52 145 250

(15) FOLLOWING LABS ELIMINATED 71 152 390
28 29 45 52 145 250

(16) FOLLOWING LABS ELIMINATED 21 28 45 71 145 304 390

(17) FOLLOWING LABS ELIMINATED 250
21 28 45 71 145 304 390

(18) FOLLOWING LABS ELIMINATED 18 26 30 52 91 94 304

(19) FOLLOWING LABS ELIMINATED 144

(20) FOLLOWING LABS ELIMINATED 28
144

FIG. X1.1 Example of Completed Table I, Summary of Results, of a CCRL Reference Sample Report

TABLE 5 (from ASTM C 802)

Material	Average	Standard Deviation		Coefficient of Variation	
		Within-Laboratory	Between-Laboratory	Within-Laboratory	Between-Laboratory
3 day strength (36)	1-2340	5-187	7-224	11-8.00%	15-9.56%
3 day strength (35)	2-3036		8-243	12-6.16%	16-8.01%
7 day strength (36)	3-3498	6-132	9-235	13-3.79%	17-6.73%
7 day strength (35)	4-4257		10-239	14-3.10%	18-5.62%
Average		160	235		

FIG. X1.2 Placement of Selected Data From Table I of CCRL Reference Sample Report into Table 5 of Practice C 802

in order of increasing magnitude.

X1.2.4 Place Circles 5 and 6 in “Within-Laboratory Standard Deviation” column of the “Approximate Values (Upper 5 % Level) for the Ratio of Highest to Lowest Variance” Table in the order shown.

X1.2.5 Place Circles 7, 8, 9, and 10 in order in the“ Between-Laboratory Standard Deviation” column of the “Approximate Values (Upper 5 % Level) for the Ratio of Highest to Lowest Variance” Table.

X1.2.6 Place Circles 11, 12, 13, and 14 in order in the “Within-Laboratory Coefficient of Variation” column of the “Approximate Values (Upper 5 % Level) for the Ratio of Highest to Lowest Variance” Table.

X1.2.7 Place Circles 15, 16, 17, and 18 in order in the “Between-Laboratory Coefficient of Variation” column of the “Approximate Values (Upper 5 % Level) for the Ratio of Highest to Lowest Variance” Table.

X1.2.8 As discussed in the Determination of Form of Precision Statement section of Practice C 802, examine the“ Approximate Values (Upper 5 % Level) for the Ratio of Highest to Lowest Variance” table to determine whether either the standard deviations or coefficients of variation are independent of level of measurement. Note that in this case the standard deviations are essentially constant.

X1.2.9 Write the precision statement based on standard deviation in accordance with the Estimates of Precision section of this practice, as shown in X1.3. In both cases, the standard deviation is multiplied by 2 $\sqrt{2}$ to obtain the d2s value, where d2s is the “difference two-sigma limit” as defined in Practice E 177.

X1.3 *Sample Precision Statement:*

X1.3.1 The single-operator standard deviation has been found to be 160 psi (1100 kPa) (the “one-sigma” [1s] limit per Practice C 670) throughout the range 2300 to 4300 psi (15 860 to 29 650 kPa). Therefore, results of two properly conducted tests by the same operator on similar batches should not differ by more than 453 psi (3123 kPa) (d2s limit).

X1.3.2 The multilaboratory standard deviation has been found to be 235 psi (1620 kPa) (1s limit) throughout the range 2300 to 4300 psi (15 860 to 29 650 kPa). Therefore, results of two different laboratories on similar batches should not differ from each other by more than 665 psi (4585 kPa) (d2s limit).