Designation: D4678 – 15a

Standard Practice for Rubber—Preparation, Testing, Acceptance, Documentation, and Use of Reference Materials[1]

This standard is issued under the fixed designation D4678; 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 covers materials used on an industry-wide basis as reference materials, which are vitally important to conduct product, specification, and development testing in the rubber industry. This practice describes the steps necessary to ensure that any candidate material, that has a perceived need, can become a Reference Material. The practice sets forth the recommendations on the preparation steps for these materials, on the testing that shall be conducted to permit acceptance of any candidate material, and on how the documentation needed for the acceptance shall be recorded for future use and review.

1.2 This practice shall be administered by ASTM Committee D11.

1.2.1 Important sections of this practice are as follows:

	Section
Significance and Use	3
Preparation of Industry Reference Materials	4
Overview of Industry Reference Material Testing	5
Chemical and Physical Specifications for IRM	6
Reference Material Documentation	7
Typical Reference Material Use	8
Recommended Package Size for IRM	Annex A1
Recommended Sampling Plans for Homogeneity Testing of an IRM	Annex A2
Test Plan and Analysis for Homogeneity of an IRM	Annex A3
Test Plan and Analysis to Evaluate an Accepted Reference Value	Annex A4
Statistical Model(s) for IRM Testing	Annex A5
Example of Annex Calculations for a Typical IRM	Appendix X1
Two-Way Analysis of Variance for Calculating *Sr*	Appendix X2

1.3 *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 and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:*[2]

D4483 Practice for Evaluating Precision for Test Method Standards in the Rubber and Carbon Black Manufacturing Industries

D5900 Specification for Physical and Chemical Properties of Industry Reference Materials (IRM)

E122 Practice for Calculating Sample Size to Estimate, With Specified Precision, the Average for a Characteristic of a Lot or Process

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

E826 Practice for Testing Homogeneity of a Metal Lot or Batch in Solid Form by Spark Atomic Emission Spectrometry

3. Significance and Use

3.1 Reference materials are vitally important in product and specification testing, in research and development work, in technical service work, and in quality control operations in the rubber industry. They are especially valuable for referee purposes.

3.2 *Categories, Classes, and Types of Reference Materials (RM):*

3.2.1 Reference materials are divided into two categories:

3.2.1.1 *Industry Reference Materials (IRM)*—Materials that have been prepared according to a specified production process to generate a uniform lot; the parameters that define the quality of the lot are evaluated by a specified measurement program.

3.2.1.2 *Common-Source Reference Materials (CRM)*—Materials that have been prepared to be as uniform as possible but do not have established property (parameter) values; the knowledge of a common or single source is sufficient for certain less critical applications.

[1] This practice is under the jurisdiction of ASTM Committee D11 on Rubber and is the direct responsibility of Subcommittee D11.20 on Compounding Materials and Procedures.

Current edition approved July 1, 2015. Published August 2015. Originally approved in 1987. Last previous edition approved in 2015 as D4678 – 15. DOI: 10.1520/D4678-15A.

[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.

Copyright © ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959. United States

Documents
Document Sharing Hub
Full document details: ASTM D4678-15a

3.2.2 Industry reference materials (IRMs) are divided into additional classes and types according to the method of evaluating the lot parameters and according to the production process for generating the lot material. These are explained more fully (refer to Annex A3 and Annex A4 for more details on the discussion in Section 3).

3.2.3 The following lot parameters are important for reference material use:

3.2.3.1 *Accepted Reference Value (AR Value)*—An average IRM property or parameter value established by way of a specified test program.

3.2.3.2 *Test Lot Limits (TL Limits)*—These are limits defined as ±3 times the standard deviation of individual IRM test results across the entire lot for the property or parameter(s) that defines lot quality; the measurements are conducted in the laboratory of the organization producing the IRM.

3.2.3.3 Although the limits as defined in 3.2.3.2 are given in terms of ±3 times the standard deviation, the rejection of individual portions of the lot as being outlier or non-typical portions in assessing the homogeneity of the lot is done on the basis of ±2 times the appropriate standard deviation, that is, on the basis of a 95 % confidence interval. See Annex A3 and Annex A4 for more information and the evaluation procedures.

3.2.4 All IRMs have an AR value and TL limits; however the AR value may be obtained in one of two ways to produce one of two classes of AR values:

3.2.4.1 *Global AR Value*—This AR value is obtained from an interlaboratory test program where the word "global" indicates an average value across many laboratories.

3.2.4.2 *Local AR Value*—This is an AR value obtained in one laboratory or at one location, usually the laboratory responsible for preparation of the homogeneous lot.

3.2.5 An additional parameter is of importance for IRMs that have a global AR value:

3.2.5.1 *Between-Laboratory Limits (BL)*—The group of laboratories that conduct interlaboratory testing to establish an AR-value are not equivalent to a system or population typical of industrial production operations that use the usual ±3 standard deviation limits. Such production operations are systems that have been purged of all assignable causes of variation and are in a state of 'statistical control' with only random variations that cannot be removed. Thus, the recommended limits on all IRMs are the ±2 standard deviation limits that pertain to a 95 % confidence level. If for serious reasons that can be totally justified, ±3 standard deviation limits are required, these may be used provided that full and complete documentation is supplied to justify the limits.

3.2.6 The homogeneity or uniformity of the lot, which determines the magnitude of the TL limits, may be designated as one of two different levels of uniformity. The key factor that determines the level of uniformity is the capability of blending the IRM portions or parts that constitute the lot, to ensure a high degree of uniformity from the blending process. IRMs that cannot be blended will have an extra residual amount of variation (portion to portion) that lowers the level of uniformity.

3.2.6.1 *Uniformity Level 1 (UL-1)*—This is the most uniform or highest level of homogeneity that can be attained by the use of *a specified test* for measuring the parameter that defines lot quality; it is obtained by the use of a blended material and is referred to as a Type *B* (*B* = blended) IRM.

3.2.6.2 *Uniformity Level 2 (UL-2)*—This is the lesser degree of uniformity that is attained by the use of*a specified test* for measuring the parameter that defines lot quality; it is normally obtained for non-blended materials and is referred to as a Type *NB* (not blended) IRM.

3.3 IRMs have a number of use applications in the technical areas, as cited in 3.1.

3.3.1 *Single Laboratory Self Evaluation*—The IRM may be used in a given laboratory (or with a given test system) to compare the test results within the laboratory to the accepted reference value for the IRM. An IRM can also be used for internal statistical quality control (SQC) operations.

3.3.2 *Multi-Laboratory Evaluation*—The IRM may be used between two or more laboratories to determine if the test systems in the laboratories are operating within selected control limits.

3.3.3 One or more IRMs may be used in the preparation of compounds to be used for evaluating non-reference materials in compound testing and performance.

3.3.4 Reference liquid IRMs may be used for immersion testing of various candidate or other reference compounds. Such immersion testing is important due to the deleterious influences of immersion liquids on rubber compounds.

3.3.5 IRMs may also be used to eliminate interlaboratory testing variation known as "test bias:" a difference between two (or more) laboratories that is essentially constant between the laboratories for a given test property level, irrespective of the time of the test comparisons. In such applications a differential test measurement value, (IRM – experimental material), becomes a corrected test result; this corrected value is used as the measure of performance rather than the "as-measured" test value on the experimental material of interest.

3.4 Average values play an important role in various operations and decisions in this practice. For this practice, "average" is defined as the arithmetic mean.

3.5 The various characteristics of IRMs and CRMs (categories, classes, types) are listed in summary form in Table 1.

3.6 This practice and the IRM program it describes was developed to replace a standardization program conducted by

TABLE 1 Categories of Reference Materials[A]

	IRM				CRM
AR Value	Global		Local		None
Homogeneity (TL Limits)	Type *B* (UL-1)	Type *NB* (UL-2) or (UL-1)	Type *B* (UL-1)	Type *NB* (UL-2) or (UL-1)	Single Source Material

[A] AR value = accepted reference value.
TL limits = test lot limits.
Global = AR value obtained from an interlaboratory test program.
Local = AR value obtained from one laboratory.
Type-B = IRM that has been blended to ensure high uniformity.
Type-NB = IRM that cannot be blended.
UL-1 and UL-2 = levels of uniformity in the IRM lot; UL-1 is higher uniformity than UL-2.
See Annex A3 and Annex A4 for more information.

AVA Documents
Document Sharing Hub

Full document details: ASTM D4678-15a

the National Institute of Standards and Technology (NIST) that began in 1948 and has been phased out.

3.7 It is not feasible to write into this practice all the necessary specifications, modes of preparation, sampling, and testing protocols, for the wide variety of materials that will eventually become IRM. Therefore this practice is published to give general guidelines for IRMs.

3.8 A permanent IRM Steering Committee within Subcommittee D11.20 shall be constituted by Subcommittee D11.90 to assist in the utilization of this practice and to make technical and, where required, policy decisions regarding the preparation and administration of IRM.

4. Preparation of Industry Reference Materials

4.1 *Basic Preparation Steps:*

4.1.1 An IRM should be prepared in a way that ensures that the entire quantity or lot of the material is as homogeneous, in composition and vital performance properties, as is possible.

4.1.2 For particulate and liquid materials this implies a thorough physical blending operation during or after the manufacturing steps, or both.

4.1.3 For materials not easily blended after manufacture, two options to ensure homogeneity are recommended:

4.1.3.1 Use highly homogeneous components or other materials that are required in the manufacturing steps or conduct certain blending operations at intermediate manufacturing steps to ensure maximum homogeneity.

4.1.3.2 Use intensive statistical quality control procedures to ensure a specified degree of homogeneity among the packets, bales, or other discrete units of the material.

4.1.4 Examples, as cited in 4.1.3.1, are such materials as accelerators, antioxidants, sulfur, and reference test (liquid) fuels.

4.1.5 Examples, as cited in 4.1.3.2, are various synthetic rubbers.

4.2 *Packaging of Industry Reference Materials:*

4.2.1 Industry reference materials should be packaged preferably in small quantities or packages. The packages shall be consecutively numbered as they are filled. Nominally the size should be the smallest amount that the average user of the material would require for normal volume testing. High volume users could therefore order multiple package lots. The use of such minimum volume (mass) packages will of course vary, but Annex A1 gives recommended masses or volumes.

4.2.2 Industry reference materials shall be suitably packaged to prevent or retard the change of IRM values with the passage of time or inadvertent exposure to heat, light, moisture, or combinations thereof, in normal storage. The stringency of this requirement varies with the type of IRM. All precautions shall be taken to make IRMs as stable as possible.

4.2.3 Packages shall be dispensed by the manufacturing or distribution organization with a document that shall furnish the following general information:

4.2.3.1 Name and number of the IRM,

4.2.3.2 Name of the manufacturer,

4.2.3.3 Date of manufacture or preparation,

4.2.3.4 Storage conditions, and

4.2.3.5 Reference to ASTM research report for documentation of testing.

4.2.4 For each test property measured to assess lot quality report the following:

4.2.4.1 Accepted reference value,

4.2.4.2 Test lot limits, and

4.2.4.3 Between-laboratory limits.

4.3 *Packaging of Common–Source Reference Materials:*

4.3.1 CRMs shall be packaged and dispensed in the same manner as for IRMs. Each CRM package shall be furnished with a documentation sheet with the following information:

4.3.1.1 Name and number of the CRM,

4.3.1.2 Name of manufacturer,

4.3.1.3 Date of manufacture or preparation,

4.3.1.4 Storage conditions, and

4.3.1.5 Reference to ASTM research report.

5. Overview of Industry Reference Material Testing

5.1 Testing is conducted to *(1)* demonstrate the uniformity of the IRM lot to some selected limits and evaluate the test lot limits, and *(2)* to establish an accepted reference value for the lot and as a secondary goal to evaluate the between-laboratory limits for interlaboratory testing of the IRM where this is applicable.

5.2 *Testing for Homogeneity:*

5.2.1 Homogeneity testing is ideally conducted in one highly qualified laboratory, which is usually the laboratory of the organization that produces the IRM. The lot size is determined and samples are drawn from the lot. Guidance for the size and number of samples is given in Annex A2. The samples taken from the lot are tested according to the instructions given in Annex A3. This latter annex also addresses the concept of different uniformity levels for an IRM and the importance of this in IRM development and use.

5.2.2 It is important that each sample represents a fraction or portion of the total lot that can be physically separated from the remainder of the lot, in the event that the portion represented by the sample is judged to be significantly different from the remainder of the lot and is therefore rejected.

5.2.3 Those portions of the lot that are shown to be significantly different from the remainder or bulk of the lot shall be rejected.

5.2.4 If, in the statistical analysis of Annex A3, a substantial fraction (25 to 30 %) of the lot is declared to be not acceptable for lack of homogeneity, retesting may be permitted. This retesting shall include all suspected portions and a number of accepted homogeneous portions or parts equal in number to the suspect portions. The retest shall be conducted according to Annex A3.

5.2.5 If on retesting and analysis of the newly generated data these same portions are again found to be unequal in property value to the accepted portions by standard statistical tests, they shall be rejected. If the suspected portions are found to be equal to the accepted portions in property values, they may be accepted as part of the lot.

AVA Documents
Document Sharing Hub

Full document details: ASTM D4678-15a

5.3 *Testing for an Accepted Reference Value*—Testing for an accepted reference value may be undertaken once a homogeneous lot has been achieved. The detailed instructions for conducting the interlaboratory program and analyzing the data of the program for an accepted reference value are given in Annex A4. This annex also gives instructions for evaluating the between-laboratory test limits where this is applicable.

5.4 *Additional Testing Background Information:*

5.4.1 To provide some theoretical background for the analyses conducted in Annex A3 and Annex A4, a discussion on statistical model development is given in Annex A5. This permits a more comprehensive understanding of the rationale for the analysis of the IRM test data and for the use of IRMs in various laboratory applications. See also Section 8 for a detailed discussion of how IRMs may be used in laboratory applications.

5.4.2 Appendix X1 gives an example of a complete set of calculations for homogeneity and accepted reference value testing according to the instructions of Annex A2 – Annex A4.

6. Chemical and Physical Specifications for IRM

6.1 Since the chemical and physical specifications for each IRM will vary in kind and degree among the various candidate materials, the details on such information are to be referred to the IRM Steering Committee. As experience is gained this practice may be amended to include more specific guidelines and test protocols. See Specification D5900 for information on all the current IRMs and their specifications.

7. Documentation for Reference Materials

7.1 *Industry Reference Materials (IRM):*

7.1.1 A full report shall be given for each IRM. This shall contain the following information:

7.1.1.1 Name of the material,

7.1.1.2 IRM number,

7.1.1.3 Organization preparing the IRM,

7.1.1.4 Date of preparation or manufacture and testing,

7.1.1.5 Any special preparation or processing steps for the IRM,

7.1.1.6 Raw data and results of the homogeneity and accepted reference value testing,

7.1.1.7 Date of adoption of the IRM,

7.1.1.8 Names of all laboratories in the AR value program,

7.1.1.9 Specific conditions under which the IRM is to be stored while awaiting distribution to laboratories purchasing the IRM, and

7.1.1.10 Any other information of a special nature needed to document special issues not covered in the above list.

7.1.2 All of the information as called for in 7.1.1 shall be prepared in a special report that can be easily interpreted and sent to ASTM International Headquarters. This shall be given a special research report number and kept on file at ASTM.

7.2 *Common-Source Reference Materials (CRM):*

7.2.1 A report shall be prepared for all CRMs with the following information:

7.2.1.1 Name of the CRM,

7.2.1.2 Number of the CRM,

7.2.1.3 Name of organization preparing the CRM,

7.2.1.4 Date of manufacture or preparation,

7.2.1.5 Storage conditions for the material while awaiting shipment, and

7.2.1.6 Any other special information pertinent to the use of the CRM.

7.2.2 All of the information in 7.2.1 shall be provided in a report sent to ASTM and kept on file as a research report.

8. Typical Reference Material Use (Global AR Value)

8.1 *IRM Application—Single Laboratory Self Evaluation:*

8.1.1 A single laboratory can use an IRM to determine how the test or measurement system in the laboratory is performing in relation to the AR value and the limits associated with the AR value. This self-evaluation of a laboratory can be most effectively conducted by setting up a statistical model. Refer to Annex A5 for background and details.

8.1.2 In A5.3.6 of Annex A5, the model for the testing for an AR value is given in Eq A5.9 and reproduced here as Eq 1, with one new added term, $B(g)$. The term y represents the measured test result.

$$y = \mu(0) + B_m + B_L + B(g) + e(g) + e_b(s) + e_w(s) \quad (1)$$

The new term is needed because the entire lot may be comprised of a number of portions or units that have (average) test values that span the (maximum to minimum) range of the lot. This new term, $B(g)$, is the bias component related to the particular portion or unit (of the entire lot) purchased and tested by the user or single laboratory.

8.1.3 In the model represented by Eq A5.9, the B_m and B_L terms were variable, because the system of measurement was the collection of laboratories participating in the ITP to evaluate the AR value. In the single laboratory self-evaluation model of Eq 1, the terms B_m and B_L are fixed; they represent values unique to the single laboratory. Thus there are four fixed or constant terms in the Eq 1 model: μ(0), B_m, B_L, and $B(g)$. The sum of these four terms represents the overall test measurement bias (potential or actual) for the single laboratory.

8.1.4 To determine if the single laboratory measurement system test values agree with the AR value, it is necessary to greatly reduce or eliminate the contribution of the random deviations or (e) terms to the y-value measurement. This is done by making a number of (y-value) measurements over a selected (short-term) period and taking an average of these. As outlined in Annex A5, the random deviations average out to zero in the long run, and thus do not contribute to the measured average y-value. The number of recommended measurements for this purpose is twelve, perhaps one or two per day, for six or twelve consecutive days. On this basis, $e(g) + e_b(s) + e_w(s) \cong 0$. This recommended action demonstrates the power-of-averaging rule. Once the average of twelve is calculated it can be compared to the AR value. Several outcomes are possible for this comparison.

8.1.5 *Potential Outcome 1*—The degree of agreement can be expressed by the difference between the twelve-test average, $y_{(12)}$, and the AR value. If this difference is expressed by Eq 2, where both sides represent absolute values,

$$|y_{(12)} - \text{AR value}| < \text{or} = |TL_{(12)}| \quad (2)$$

AVA Documents
Document Sharing Hub

Full document details: ASTM D4678-15a

then there is good agreement since $y_{(12)}$ falls within the nominal range: AR value ± TL limits. The single laboratory may be said to be operating on target and the sum of all four biases approaches zero. Note that the individual biases may not be zero; their sum is zero.

8.1.6 *Potential Outcome 2*—If the difference between $y_{(12)}$ and the AR value is expressed by Eq 3,

$$\left| y_{(12)} - \text{AR value} \right| >= \left| \textit{TL limits} \right| \quad (3)$$

then the single laboratory is not operating on target: the sum of the four biases is not zero. If the difference ($y_{(12)}$ – AR value) is negative, the laboratory has a negative total bias; if the difference is positive, the total bias is positive.

8.1.7 *Potential Outcome 3*—If the outcome of the comparison of $y_{(12)}$ versus the AR value is given by Eq 3, the next step is to decide if the laboratory is operating within the between-laboratory limits, which may be considered as current inter-laboratory nominal testing variation (NTV). If the difference between $y_{(12)}$ and the AR value is expressed by Eq 4,

$$\left| y_{(12)} - \text{AR value} \right| <\text{or} = \left| \textit{between} - \textit{laboratory limits} \right| \quad (4)$$

then the single laboratory is operating the test system within the NTV limits of typical laboratories in the industry.

8.1.8 *Potential Outcome 4*—If the difference between $y_{(12)}$ and the AR value is expressed by Eq 5,

$$\left| y_{(12)} - \text{AR value} \right| >= \left| \textit{between} - \textit{laboratory limits} \right| \quad (5)$$

then the single laboratory is not operating within the NTV limits of typical laboratories in the industry.

8.2 *IRM Application—Multi-Laboratory Evaluation:*

8.2.1 One of the most important uses of an IRM is its application to resolve questions and disputes over poor agreement for producer-consumer testing. As demonstrated above, any numerical deviation between two laboratories (when both laboratories have measured the same material) has two types of components: *Type 1*, a combined deviation component due to random measurement variations in both laboratories, and *Type 2*, a deviation component due to inter-laboratory bias. As previously demonstrated, sufficient replicate testing in each of the laboratories will reduce the random component to zero, but will not influence the inter-laboratory bias.

8.2.2 An estimate of the bias in each of the laboratories may be determined by use of an appropriate IRM. One laboratory should supply an IRM sample, taken from one portion or unit of the IRM lot, to both laboratories. Each sample should be large enough to perform at least twelve tests in each laboratory. Each laboratory performs a selected number of tests, from six to twelve, depending upon the importance of the testing dispute, over a selected short-term time period of several days. The average of these tests is defined as y_{avg}.

8.2.3 The overall bias for each laboratory is estimated by means of Eq 6.

$$\text{Estimated Overall Bias}\left(\text{Laboratory } i\right) = \left(y_{\text{avg}} - \text{AR value}\right) \quad (6)$$

The overall IRM bias values for each laboratory can be compared and used to make decisions about resolving any potential testing problem.

8.2.4 The algebraic difference between the two laboratory overall biases is the direct bias between the two laboratories.

$$\text{Direct Inter} - \text{laboratory Bias} = \quad (7)$$

$$\Delta\left(\text{Laboratory 2 Bias} - \text{Laboratory 1 Bias}\right)$$

Such information can be potentially used for corrections in test data. The use of such information for correcting interlaboratory test data should be done only on the basis of a mutual agreement between the participating laboratories.

8.2.5 The procedure as outlined in 8.2.2 and 8.2.3, can be extended to any number of laboratories by modifying the procedural steps in an appropriate manner. This operation, as stated, should only be done on the basis of mutual agreement.

9. Keywords

9.1 common source reference material (CRM); industry reference material (IRM); reference material

ANNEXES

(Mandatory Information)

A1. RECOMMENDED PACKAGE SIZES FOR IRM

A1.1 *Lot Size*—A lot size of 250 to 1000 packages is recommended depending upon the anticipated usage rate.

A1.2 For IRM rubber chemicals with a bulk density similar to accelerators, antioxidants, and sulfur, a container of about 2 dm^3 (litres) is recommended, with mass adjusted to the nearest 100 g required to fill the container.

A1.3 For IRM such as carbon black fillers or other materials used in relatively high proportions in rubber, one of two containers is recommended:

A1.3.1 A 20-dm^3 (litre) container (approximately 5 gal) with mass or volume equivalent adjusted to nearest 0.5 kg.

A1.3.2 A 25-kg bag, if the material is particulate. Materials that may potentially absorb moisture or other gases (CO_2) shall be placed inside another outer container to prevent such absorption.

A1.4 For IRM such as rubbers, a package or bale of 34 kg mass (1/30 metric ton) is recommended.

AVA Documents
Document Sharing Hub

Full document details: ASTM D4678-15a

A2. RECOMMENDED SAMPLING PLANS FOR HOMOGENEITY TESTING OF AN IRM

A2.1 *Introduction*—Sampling plans are required to measure the particular property of the lot that has been selected to assess the quality of the lot. Two sampling plans are given. Plan 1 gives instructions to calculate the sample size such that the maximum deviation, *E*, between the measured property average (the estimate of the lot true value) and the actual lot true value (that is, the measured average of all lot packages, items or portions), may be calculated. An advanced knowledge of the measured property standard deviation is required for this plan. Plan 2 is a less rigorous approach that may be used when it is possible to blend the lot material and achieve greater homogeneity prior to sampling. Practice E122 is used as a reference for this annex.

A2.2 *Sampling Plans for the IRM*—One of two sampling plans shall be used to sample from the lot. Sampling Plan 1 is preferred.

A2.2.1 *Sampling Plan 1*—Draw from the lot a number of samples, *n*, to satisfy the selected or desired maximum deviation, *E*, between *Xn*, the lot average using all *n* samples and $\bar{X}a$ the lot true value as defined in A2.1. The calculation is performed by using Eq A2.1. Use either the advanced estimate of the property standard deviation, *Se*, or a well-established standard deviation, *S*.

$$n = (3\ Se/E)^2 \tag{A2.1}$$

where:

E = $\bar{X}a - \bar{X}n$, and
Se = advanced estimate of the standard deviation of the measured property that defines lot quality; this may be obtained from measurements on the lot after its manufacture or from process control data during actual production of the lot.

A2.2.2 The quantity *E* is a 3 (standard deviation) limit on $\bar{X}n$. Thus the range, $\bar{X}n \pm E$, will contain the true value of the lot property with a 99.7 % confidence level. If there is no advanced knowledge of *Se*, it shall be estimated from the lot with a minimum of twelve samples equally spaced throughout the lot. The test results from this preliminary estimate of *Se* may be incorporated into the test results from additional lot sampling and testing, should the value of *n* exceed twelve.

A2.2.3 The deviation, *E*, may be expressed in terms of the lot standard deviation *Se* (or *S* if it is known). Table A2.1 gives a series of values of *n* that correspond to values of *E* expressed as a fraction of *Se* (or *S*).

TABLE A2.1 *E* Values for Selected Values of *n*

n	*E*
12	0.87 *Se*
16	0.75 *Se*
20	0.67 *Se*
24	0.61 *Se*
30	0.55 *Se*
36	0.50 *Se*
56	0.40 *Se*
100	0.30 *Se*

A2.2.4 The value of *E* shall be selected based upon general experience of those familiar with the specific testing in consultation with the permanent D11 IRM Steering Committee. The samples shall be taken during production at intervals so as to sample the entire lot in a uniform process (equally distributed sample selection). Two options are offered: sample during the package filling process or sample from finished packages.

A2.2.5 *Size of Samples*—The physical volume or mass taken for a sample will depend on the material being sampled. For rubbers, the sample size shall be 3 to 4 kg. For rubber chemicals, liquids, carbon black, or fillers, an appropriate amount shall be taken to allow for several test portions per sample. This is important for retest operations.

A2.3 *Sampling Plan 2*—If, for certain justified reasons, a sampling plan as described in Sampling Plan 1 cannot be carried out, Sampling Plan 2 shall be conducted. This second sampling plan requires fewer samples and it may be carried out when extensive blending of the IRM has been conducted during the manufacturing or production process or subsequent to its production but prior to the sampling operation. This blending ensures that a high degree of homogeneity exists and decreases the reliance on an extended sampling operation.

A2.3.1 Select at least twelve samples from the lot on a basis that ensures that all portions of the lot from beginning to end of the production process are represented equally among the twelve or more samples. If the samples are selected from a large container, ensure that all zones or locations in the container are equally represented in the samples.

A2.3.2 The sample size for Plan 2 shall be as specified in A2.2.5, with a sufficient amount in each sample for several tests for whatever property is being measured.

AVA Documents
Document Sharing Hub

Full document details: ASTM D4678-15a

A3. TEST PLAN AND ANALYSIS FOR HOMOGENEITY OF AN IRM

A3.1 *Introduction:*

A3.1.1 This annex gives the instructions for evaluating the homogeneity for any candidate IRM. The testing (*1*) establishes the degree of uniformity for the measured properties that define the quality of the lot, and (*2*) generates data to establish the test lot limits (TL limits).

A3.1.2 The homogeneity is established on the basis of a 95 % confidence interval, that is, portions of the lot are rejected as outlier portions based on their measured values exceeding ±2 standard deviation limits. Once a homogeneous lot has been accepted on this basis, the TL limits are defined as the ±3 standard deviation limits of the measured individual test results in the IRM production laboratory. Practice E826 is used as a reference for this annex.

A3.2 *Brief Theory of Homogeneity Analysis:*

A3.2.1 The homogeneity analysis concepts are developed by considering a perfectly homogenous lot of material. If n samples of this lot are taken and k-replicate tests are conducted on each sample, the homogeneity of the lot is demonstrated if the pooled standard deviation or variance among the n sets of k-replicates is statistically equal to the standard deviation or variance (adjusted for the value of k), among the n samples drawn from the lot, that is, the observed variation among the lot samples is not significantly greater than that contributed by the testing itself!

A3.2.2 If it is not known whether the lot is truly homogeneous, and the sample variation is shown to be statistically greater than the pooled k-replicate variation, it is then presumed that the lot is not homogeneous. The next problem is to decide which part or parts of the lot depart from the bulk (or some remainder) of the lot.

A3.2.3 Decisions about homogeneity are made on the basis of ranges. Tests are conducted (k-replicates) on each of the n samples; averages (X), of the k values are calculated and the range among the averages is evaluated. This observed range among all lot samples, w(obs), which is equal to X(max) – X(min), is compared to a critical range, w(crit), that is evaluated on the basis of the expected range, if the only source of variation among the n samples is that due to the k-replicates. If w(obs) is greater than w(crit), then some part or parts of the lot deviate substantially from the bulk or remainder and are the cause of the non-homogeneity.

A3.2.4 If the n sample averages are sorted from low (minimum) to high (maximum), the deviating part or parts are easily identified and can be discarded from the lot. This part by part sequence is repeated until w(obs) is less than w(crit) and a homogeneous lot is obtained.

A3.2.5 Once a homogeneous lot is obtained, the TL limits are calculated from the standard deviation of the remaining packages, items, or portions of the lot (or the entire lot if no portions are rejected).

A3.3 *Conducting the Homogeneity Analysis:*

A3.3.1 *Correcting for Test Machine Drift:*

A3.3.1.1 The first step is to conduct the testing on the n samples drawn from the lot. Each sample is to be tested k number of times. If a large number of samples has been drawn or a large number taken during a production process and the time span to conduct all $n \times k$ tests is more than one day, an evaluation for measurement system drift shall be made. This evaluation is conducted by testing a control material according to a specific plan, depending on the number of samples taken from the lot.

A3.3.1.2 The control material shall be of the same type and have approximately the same property level as the IRM and be as uniform as possible. If any doubts exist on uniformity, blending shall be done if this is possible. The testing is conducted according to a specified plan with terms defined as follows:

NOTE A3.1—The control material may be a part or small fraction of the lot of the IRM that is sampled.

A3.3.1.3 *Test Number*—The samples are to be numbered in consecutive order as they are drawn from the lot during the sampling process. A random testing order for these consecutive samples is recommended. If this is not possible, a notation should be made in the report.

A3.3.1.4 *Test Replication*—The test sequence involving all n samples is conducted k times (k replicates) in the random sample order as indicated under "Test Number." The value for k is ideally 4. If this imposes a burden on the testing program or if the number of samples is large, a value of 3 or, at the very least, 2, may be selected for k.

A3.3.1.5 *Control Testing Frequency*—The control material, C, shall be tested at a frequency that is dictated by the size of the lot (number of samples drawn). Table A3.1 gives the testing frequency. The frequency is defined as the number of IRM samples, NI_i, tested between successive control samples. The control material is always tested first.

A3.3.2 *Analysis for Drift:*

A3.3.2.1 Tabulate the control test values in order of testing; C_1, C_2, C_3, ... C_n, and calculate the differences Δ between immediate successive values of C as follows:

$$\Delta_1 = C_1 - C_2 \quad (A3.1)$$

$$\Delta_2 = C_2 - C_3 \quad (A3.2)$$

$$\Delta_3 = C_3 - C_4, \text{ etc.} \quad (A3.3)$$

TABLE A3.1 Control Material Testing Frequency

Number of samples, m	Testing Order and Frequency Sequence for Control Material[A]
5 or less	Frequency = 2, that is, use [C, NI_1, NI_2, C, NI_3, NI_4, C, NI_5]
5 to 10	Frequency = 3, that is, use [C, NI_1, NI_2, NI_3, C, NI_4, NI_5, NI_6, C, etc.]
11 to 20	Frequency = 4
21 to 50	Frequency = 5

[A] C = control material; NI_i = test frequency = number of IRM samples tested between successive control samples.

AVA Documents
Document Sharing Hub

Full document details: ASTM D4678-15a

A3.3.2.2 Calculate the variance of the C values, called S_1^2, based on successive differences as follows:

$$S_1^2 = \sum \Delta^2/2(m-1) \quad \text{(A3.4)}$$

where:

Δ = difference in immediate successive values of C, and
m = total number of control material samples tested.

A3.3.2.3 Calculate a second estimate of the variance among the C values S_2^2 as follows:

$$S_2^2 = \sum d^2/(m-1) \quad \text{(A3.5)}$$

where:

$d = (Ci - \bar{C})$ = difference of each measured Ci value from the average of all C values, designated $\bar{C}$.

A3.3.2.4 A decision on the occurrence of drift is made as follows. Calculate the ratio of S_1^2 to S_2^2. If the ratio S_1^2/S_2^2 obtained from the C value measurements is larger than the critical ratio values listed in Table A3.2 for the specified value of m, which is the total number of C values measured, then a statement can be made that there is no drift. The confidence level for this statement is 95 %.

A3.3.2.5 If the ratio S_1^2/S_2^2 is less than the critical tabulated value, then a statement can be made that drift has occurred. The confidence level for this statement is also 95 %.

A3.3.3 *Correction for Drift:*

A3.3.3.1 If analysis shows that drift is absent, the measured values of the IRM samples should be used. If this is the case, proceed to the next section. If drift is shown to be present, make a correction of the IRM sample values for the drift.

A3.3.3.2 A correction for drift is made on the basis of a linear drift behavior. The "drift" is corrected by the use of drift correction factors, F_i.

A3.3.3.3 Arrange the data values obtained for the control material in chronological order (C_1, C_2, ... C_n). Compute the drift factors, F_i, as follows:

$$F_1 = (C_1 + C_2)/2C_1 \quad \text{(A3.6)}$$

$$F_2 = (C_2 + C_3)/2C_1 \quad \text{(A3.7)}$$

$$F_3 = (C_3 + C_4)/2C_1 \quad \text{(A3.8)}$$

$$F_4 = (C_4 + C_5)/2C_1 \quad \text{(A3.9)}$$

$$F_5 = (C_5 + C_6)/2C_1 \quad \text{(A3.10)}$$

$$F_6 = (C_6 + C_7)/2C_1 \quad \text{(A3.11)}$$

$$F_7 = (C_7 + C_8)/2C_1 \quad \text{(A3.12)}$$

$$F_8 = (C_8 + C_9)/2C_1 \quad \text{(A3.13)}$$

TABLE A3.2 Critical Values of (S_1^2 /S_2^2) Ratio at 95 % Confidence Level[A]

Number of Control Samples Measured, m	(S_1^2/S_2^2) Ratio
4	0.39
5	0.42
6	0.44
8	0.49
10	0.53
12	0.56
15	0.60
20	0.65
25	0.68
30	0.72
35	0.74
40	0.76
45	0.78
50	0.80

[A] See NBS Special Publication N-63-2, U.S. Government Printing Office, Washington, DC 19630; see also C. A. Bennett, *Industrial and Engineering Chemistry*, 43, 2063 (1951). Ratio values for $m = 30$ to 50 calculated from extrapolation equation; Ratio (S_1^2/S_2^2) = 0.146 + 0.386 × $\log_{10}$ (m).

A3.3.3.4 Divide the measured IRM values by the appropriate drift factor to obtain the corrected value for the IRM samples. The appropriate drift factor for the IRM values is that factor, calculated from the control or C values that brackets the measured materials (IRM samples) within the time or measurement span for the two C values. Apply the factor F_1 to the IRM samples between $\bar{C}_1$ and C_2; apply F_2 to the IRM samples between C_2 and C_3, etc. Therefore, for a Frequency 3 Program:

$$\text{(Corrected) IRM Sample 1, 2, or 3 value} = \frac{\text{Measured IRM Sample 1, 2, or 3 Value}}{F_1} \quad \text{(A3.14)}$$

$$\text{(Corrected) IRM Sample 4, 5, or 6 value} = \frac{\text{Measured IRM Sample 4, 5, 6 Value}}{F_2}, \text{ etc.} \quad \text{(A3.15)}$$

A3.3.3.5 Tabulate the drift corrected IRM values for subsequent analysis and review.

A3.3.4 *IRM Lot Characteristics:*

A3.3.4.1 An IRM may be one of two different types: (*1*) a particulate or liquid material that can be blended to improve uniformity, or (*2*) a material produced in a form that cannot be blended. Thus there are two types of IRM:

(1) a Type *B* IRM, a material that may be blended, and
(2) a Type *NB* IRM, a material that cannot be blended.

A3.3.4.2 The procedure to demonstrate what degree homogeneity exists in the lot, differs depending on what type of IRM is being evaluated. The important issue is the "variation metric" or standard deviation that is used to calculate an expected range for the measured average values of the lot samples. This standard deviation depends on the history of the lot at the time of sampling and on the type of IRM, Type *B* or Type *NB*.

A3.3.5 *Basic Concepts for Evaluating Homogeneity:*

A3.3.5.1 The homogeneity analysis is based on the distribution of the q-statistic, defined by Eq A3.16.

$$q = w/[Sr/(k)^{0.5}] \quad \text{(A3.16)}$$

where:

w = range (that is, maximum-minimum) of n average values,
Sr = residual standard deviation (pooled) of individual test values, and
k = number of individual values used for each average.

A3.3.5.2 Table A3.3 gives the 95 % confidence level (or $p = 0.05$) critical q-values for this distribution, with DF = the degrees of freedom for the pooled Sr and n = the number of sample averages in the calculated range, w. In applying the basic statistical expression of Eq A3.16, the equation is

AVA Documents
Document Sharing Hub

Full document details: ASTM D4678-15a