Designation: ~~D4821 – 07$^{\varepsilon 1}$~~ D4821 – 14

Standard Guide for Carbon Black—Validation of Test Method Precision and Bias[1]

This standard is issued under the fixed designation D4821; 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.

~~ε^1 NOTE—Editorially corrected Footnote 3 and corresponding references in August 2011.~~

1. Scope

1.1 This guide covers ~~a procedure~~ procedures for using the ASTM Standard Reference Blacks[2] (SRBs) and the HT and INR Iodine Number Standards to continuously monitor the precision of those carbon black test methods for which ~~standard~~reference values have been established. It also offers guidelines for troubleshooting various test methods.

1.2 This guide establishes ~~the x-chart control limits to be used when continuously monitoring~~procedures for the use of x-charts to continuously monitor those tests listed in Section 2~~. Alternatively, these control limits may be used as a basis for comparison to testing precision computed within a laboratory~~. for within-lab precision (repeatability) and between-lab accuracy (reproducibility).

~~1.3 This guide uses statistical control chart methodology as discussed in MNL7[3] to determine if a laboratory's test results differ significantly from the accepted values of the SRBs.~~

1.3 This guide provides a statistical procedure for improving test reproducibility when a laboratory cannot physically calibrate its apparatus to obtain the ~~standard~~reference values of the ASTM ~~SRBs,~~ reference blacks, within the ranges given in ~~the precision statement of the test method.~~this guide.

~~1.5 The values stated in SI units are to be regarded as the standard. The values given in parentheses are for information only.~~

~~1.6 *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:*[3]

D1510 Test Method for Carbon Black—Iodine Adsorption Number

D1513 Test Method for Carbon Black, Pelleted—Pour Density

D1765 Classification System for Carbon Blacks Used in Rubber Products

D2414 Test Method for Carbon Black—Oil Absorption Number (OAN)

~~D3037~~D3265 Test Method for Carbon ~~Black—Surface Area by Nitrogen Adsorption~~Black—Tint Strength ~~(Withdrawn 1999)~~

~~D3191~~D3324 ~~Test Methods for Carbon Black in SBR (Styrene-Butadiene Rubber)—Recipe and Evaluation Procedures~~Practice for Carbon Black—Improving Test Reproducibility Using ASTM Standard Reference Blacks (Withdrawn 2002)[4]

~~D3192 Test Methods for Carbon Black Evaluation in NR (Natural Rubber)~~

~~D3265 Test Method for Carbon Black—Tint Strength~~

D3493 Test Method for Carbon Black—Oil Absorption Number of Compressed Sample (COAN)

~~D3765~~D6556 Test Method for Carbon ~~Black—CTAB (Cetyltrimethylammonium Bromide)~~Black—Total and External Surface Area by Nitrogen Adsorption ~~(Withdrawn 2007)~~

[1] This guide is under the jurisdiction of ASTM Committee D24 on Carbon Black and is the direct responsibility of Subcommittee D24.61 on Carbon Black Sampling and Statistical Analysis.

Current edition approved ~~Oct. 1, 2007~~June 1, 2014. Published ~~November 2007~~August 2014. Originally approved in 1988. Last previous edition approved in ~~2006~~2007 as ~~D4821 – 06.~~D4821 – 07$^{\varepsilon 1}$. DOI: ~~10.1520/D4821-07E01.~~10.1520/D4821-14.

[2] Standard Reference Blacks are available from Laboratory Standards & Technologies, Inc., 227 Somerset St., Borger, TX 79007.

[3] 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.

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

D4820 Test Methods for Carbon Black—Surface Area by Multipoint B.E.T. Nitrogen Adsorption (Withdrawn 2000)[5]
D5816E177 Test Methods for Carbon Black—External Surface Area by Multipoint Nitrogen AdsorptionPractice for Use of the Terms Precision and Bias in ASTM Test Methods (Withdrawn 2000)
D6556E2282 Test Method for Carbon Black—Total and External Surface Area by Nitrogen AdsorptionGuide for Defining the Test Result of a Test Method
E2586 Practice for Calculating and Using Basic Statistics

3. Terminology

3.1 *Definitions:*

3.1.1 *accepted reference value, n*—a value that serves as an agreed-upon reference for comparison, and which is derived as: *(1)* a theoretical or established value, based on scientific principles, *(2)* an assigned or certified value, based on experimental work of some national or international organization, or *(3)* a consensus or certified value, based on collaborative experimental work under the auspices of a scientific or engineering group.

3.1.1.1 *Discussion*—

A national or international organization, referred to in *(2)*, generally maintains measurement standards to which the reference values obtained are traceable. **E177**

3.1.2 *accuracy, n*—the closeness of agreement between a test result and an accepted reference value.

3.1.2.1 *Discussion*—

The term accuracy, when applied to a set of test results, involves a combination of a random component and of a common systematic error or bias component. **E177**

3.1.3 *ASTM reference blacks, n*—a set of carbon blacks that span the useful range of the test method for which they are reference materials. **D3324**

3.1.4 *bias, n*—the difference between the expectation of the test results and an accepted reference value.

3.1.4.1 *Discussion*—

Bias is the total systematic error as contrasted to random error. There may be one or more systematic error components contributing to the bias. A larger systematic difference from the accepted reference value is reflected by a larger bias value. **E177**

3.1.5 *characteristic, n*—a property of items in a sample or population which, when measured, counted or otherwise observed, helps to distinguish between the items. **E2282**

3.1.6 *coefficient of variation, CV, n*—for a nonnegative characteristic, the ratio of the standard deviation to the mean for a population or sample. **E2586**

3.1.7 *intermediate precision, n*—the closeness of agreement between test results obtained under specified intermediate precision conditions.

3.1.7.1 *Discussion*—

The specific measure and the specific conditions must be specified for each intermediate measure of precision; thus, "standard deviation of test results among operators in a laboratory," or "day-to-day standard deviation within a laboratory for the same operator."

3.1.7.2 *Discussion*—

Because the training of operators, the agreement of different pieces of equipment in the same laboratory and the variation of environmental conditions with longer time intervals all depend on the degree of within-laboratory control, the intermediate measures of precision are likely to vary appreciably from laboratory to laboratory. Thus, intermediate precisions may be more characteristic of individual laboratories than of the test method. **E177**

3.1.8 *intermediate precision conditions, n*—conditions under which test results are obtained with the same test method using test units or test specimens taken at random from a single quantity of material that is as nearly homogeneous as possible, and with changing conditions such as operator, measuring equipment, location within the laboratory, and time. **E177**

3.1.9 *measured value, n*—an observed test results as opposed to a standard value. **D3324**

3.1.10 *normalization, n*—the practice of applying a statistical correction to test measurements to improve accuracy.

3.1.10.1 *Discussion*—

The correction of test data using a straight-line equation (linear regression) where measurements of ASTM reference blacks are analyzed with published accepted reference values to determine a slope and y-intercept. Normalization is a proven technique to improve the accuracy or reproducibility of laboratory data when all other means of calibration do not satisfactorily achieve a desired state of calibration.

3.1.11 *observation, n*—the process of obtaining information regarding the presence or absence of an attribute of a test specimen, or of making a reading on a characteristic or dimension of a test specimen. **E2282**

3.1.12 *observed value, n*—the value obtained by making an observation. **E2282**

3.1.13 *precision, n*—the closeness of agreement between independent test results obtained under stipulated conditions.

3.1.13.1 *Discussion*—

Precision depends on random errors and does not relate to the accepted reference value.

3.1.13.2 *Discussion*—

The measure of precision usually is expressed in terms of imprecision and computed as a standard deviation of the test results. Less precision is reflected by a larger standard deviation.

3.1.13.3 *Discussion*—

"Independent test results" means results obtained in a manner not influenced by any previous result on the same or similar test object. Quantitative measures of precision depend critically on the stipulated conditions. Repeatability and reproducibility conditions are particular sets of extreme stipulated conditions. **E177**

3.1.14 *regression of standard values on measured values, n*—statistical equation derived by the method of least-squares. **D3324**

3.1.15 *repeatability, n*—precision under repeatability conditions.

3.1.15.1 *Discussion*—

Repeatability is one of the concepts or categories of the precision of a test method.

3.1.15.2 *Discussion*—

Measures of repeatability defined in this compilation are repeatability standard deviation and repeatability limit. **E177**

3.1.16 *repeatability conditions, n*—conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time.

3.1.16.1 *Discussion*—

See precision, the "same operator, same equipment" requirement means that for a particular step in the measurement process, the same combination of operator and equipment is used for every test result. Thus, one operator may prepare the test specimens, a second measure the dimensions and a third measure the mass in a test method for determining density.

3.1.16.2 *Discussion*—

By "in the shortest practical period of time" is meant that the test results, at least for one material, are obtained in a time period not less than in normal testing and not so long as to permit significant change in test material, equipment or environment. **E177**

3.1.17 *repeatability limit (r), n*—the value below which the absolute difference between two individual test results obtained under repeatability conditions may be expected to occur with a probability of approximately 0.95 (95 %).

3.1.17.1 *Discussion*—

The repeatability limit is times the repeatability standard deviation. This multiplier is independent of the size of the interlaboratory study.

3.1.17.2 *Discussion—*

The approximation to 0.95 is reasonably good (say 0.90 to 0.98) when many laboratories (30 or more) are involved, but is likely to be poor when fewer than eight laboratories are studied. **E177**

3.1.18 *repeatability standard deviation (sr), n*—the standard deviation of test results obtained under repeatability conditions.

3.1.18.1 *Discussion—*

It is a measure of the dispersion of the distribution of test results under repeatability conditions.

3.1.18.2 *Discussion—*

Similarly, "repeatability variance" and "repeatability coefficient of variation" could be defined and used as measures of the dispersion of test results under repeatability conditions.—In an interlaboratory study, this is the pooled standard deviation of test results obtained under repeatability conditions.

3.1.18.3 *Discussion—*

The repeatability standard deviation, usually considered a property of the test method, will generally be smaller than the within-laboratory standard deviation. (See **within-laboratory standard deviation**.) **E177**

3.1.19 *reproducibility, n*—precision under reproducibility conditions. **E177**

3.1.20 *reproducibility conditions, n*—conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment.

3.1.20.1 *Discussion—*

Identical material means either the same test units or test specimens are tested by all the laboratories as for a nondestructive test or test units or test specimens are taken at random from a single quantity of material that is as nearly homogeneous as possible. A different laboratory of necessity means a different operator, different equipment, and different location and under different supervisory control. **E177**

3.1.21 *reproducibility limit (R), n*—the value below which the absolute difference between two test results obtained under reproducibility conditions may be expected to occur with a probability of approximately 0.95 (95 %).

3.1.21.1 *Discussion—*

The reproducibility limit is times the reproducibility standard deviation. The multiplier is independent of the size of the interlaboratory study (that is, of the number of laboratories participating).

3.1.21.2 *Discussion—*

The approximation to 0.95 is reasonably good (say 0.90 to 0.98) when many laboratories (30 or more) are involved but is likely to be poor when fewer than eight laboratories are studied. **E177**

3.1.22 *reproducibility standard deviation (sR), n*—the standard deviation of test results obtained under reproducibility conditions.

3.1.22.1 *Discussion—*

Other measures of the dispersion of test results obtained under reproducibility conditions are the "reproducibility variance" and the "reproducibility coefficient of variation."

3.1.22.2 *Discussion—*

The reproducibility standard deviation includes, in addition to between laboratory variability, the repeatability standard deviation and a contribution from the interaction of laboratory factors (that is, differences between operators, equipment and environments) with material factors (that is, the differences between properties of the materials other than that property of interest). **E177**

3.1.23 *standard deviation, n—of a population*, σ, the square root of the average or expected value of the squared deviation of a variable from its mean; *of a sample*, s, the square root of the sum of the squared deviations of the observed values in the sample divided by the sample size minus one. **E2586**

3.1.24 *standard value, n*—the value assigned to a reference black by ASTM Committee D24 on Carbon Black.

3.1.24.1 *Discussion—*

Usually this value is calculated as the average test result of an interlaboratory testing program. **D3324**

3.1.25 *test determination, n*—the value of a characteristic or dimension of a single test specimen derived from one or more observed values. **E2282**

3.1.26 *test method, n*—a definitive procedure that produces a test result. **E2282**

3.1.27 *test result, n*—the value of a characteristic obtained by carrying out a specified test method. **E2282**

3.1.28 *test sample, n*—the total quantity of material (containing one or more test specimens) needed to obtain a test result as specified in the test method. See **test result**. **E2282**

3.1.29 *test specimen, n*—the portion of a test sample needed to obtain a single test determination. **E2282**

3.1.30 *trueness, n*—the closeness of agreement between the population mean of the measurements or test results and the accepted reference value.

3.1.30.1 *Discussion—*

"Population mean" is, conceptually, the average value of an indefinitely large number of test results. **E177**

3.1.31 *variance,* σ^2, s^2, *n*—square of the standard deviation of the population or sample. **E2586**

3.1.32 *within-laboratory standard deviation, n*—the standard deviation of test results obtained within a laboratory for a single material under conditions that may include such elements as different operators, equipment, and longer time intervals.

3.1.32.1 *Discussion—*

Because the training of operators, the agreement of different pieces of equipment in the same laboratory and the variation of environmental conditions with longer time intervals depend on the degree of within-laboratory control, the within-laboratory standard deviation is likely to vary appreciably from laboratory to laboratory. **E177**

4. Significance and Use

4.1 ~~One of the major causes of poor test precision is the lack of calibration or standardization of instruments, apparatus, reagents, and technique among laboratories. The sum of all sources of testing error is unique for an individual laboratory. A least-squares regression of a laboratory's actual test values for reference materials to the established mean values will result in a unique least-squares regression line (and equation) for that laboratory. Generally, there are two reasons for using the SRBs in testing: (~~This guide recommends the use of statistical x-charts to graphically monitor test data determined for the ASTM reference blacks for~~1) to monitor testing performance (see~~ those test methods given in Section ~~4~~2~~) to ensure that no systematic error or bias is affecting the test results, or (~~. All laboratories are encouraged to utilize ~~2) to establish a statistical calibration (see Section~~statistical x-charts and ASTM ~~5) when the correction of assignable causes (see Section~~reference blacks because this enables a comparison of testing precision within and ~~6) does not yield in-control test results.~~between laboratories. The guide describes practices for the use of repeatability and reproducibility limits and x-charts.

4.2 In addition to the calibration of a test method by physicochemical means, a statistical method for achieving calibration of a test method is ~~presented.~~presented (that is, *normalization*).

4.3 ~~This guide outlines the use of control charts to graphically present calibration test data determined for the ASTM SRBs for those test methods given in Section~~Poor test precision can be the result of poor repeatability or poor reproducibility or both. Causes may include inadequate operator training, improperly maintained equipment or laboratory environment, variation in sample ~~2. All laboratories are encouraged to utilize statistical control charts and the SRBs because this allows a comparison of testing precision within a laboratory to the "industry average" values found in~~preparation or analysis techniques, the lack of calibration or standardization of instrumentation, worn-out apparatus, reagents that do not meet ~~Table 1,~~specifications, ~~Table 2, or~~different sources ~~Table 3.~~of instrumentation or equipment, and material heterogeneity. The sum of all sources of testing error is unique for an individual laboratory.

4.4 The techniques of this guide can be used to continuously monitor testing execution and precision for other tests that are not listed in SectionPrecision data for ASTM Reference Blacks are found in Tables 1-32 or for materials that fall outside the range of the SRBs for those tests listed in Section 2. In these cases, each laboratory will have to establish the applicable mean and control limit values for the "local reference." The monitoring will then consist of a comparison of present test results for the "local reference" to past performance within that laboratory instead of to "industry average" values.These include standard reference blacks (SRB's) Series 8, HT and INR Iodine Standards. The HT or INR Iodine standards are recommended for monitoring iodine testing.

NOTE 1—Preferred precision values are bolded in Tables 1-3.

4. Procedure for Continuously Monitoring Testing

4.1 For each test of interest, test each SRB listed for that test in Table 1, Table 2, or Table 3 in duplicate. Use the mean value for each SRB to establish the baseline values. A new baseline should be determined whenever the test equipment or conditions change. If a "local reference" is going to be used for test monitoring, it should be tested at the same time and included in the baseline data.

4.2 Select one (or more) SRBs from the SRB 4, SRB 5, SRB 6, SRB 7, or SRB HT materials (see Note 1) or a "local reference" to cover the range of interest. Because of the differing grades in each SRB set and material ages, do not mix materials from different SRB sets. For example, do not use A, B, and C from set 4 with D, E, and F from set 5. This is especially critical for oil absorptometer calibration. An absorptometer calibrated with F5 (or F5A) must be checked with other members of the 5 set. Likewise, an absorptometer calibrated with F6 or F7 must be checked with other members of the 6 or 7 set, respectively.

NOTE 1—The SRB 4 and 5 sets are depleted and not commercially available. Some members of the SRB 6 set are depleted and not commercially available. SRB G5 is used as a member of the SRB 6 and 7 sets. The SRB 4 and SRB 5 sets may still be, and the SRB 6 set is known to be, in use in some laboratories. Because of the known effects of material aging, it is recommended that the most current set of SRBs be used for test monitoring.

4.3 Prepare a control chart for each of the selected SRBs or "local reference" material(s) for each test method as presented by Part 3 of MNL7.[3] It is an accepted practice to control chart one reference material on each day of testing and rotate through each selected reference material on successive days of testing.

4.4 The target values given in Table 1 for SRB6, Table 2 for SRB7, and Table 3 for SRB HT materials were determined during the validation of the respective materials. Values are used as control chart limits (x-chart) plus or minus three single test repeatability standard deviations.[6] The mean and control chart limits (three standard deviations) for use on the x-charts must be determined by each laboratory for any "local reference" material(s).

4.5 Plot the uncorrected values for the selected reference materials (see Note 2). If a control limit is exceeded, perform a retest immediately. If the retest falls outside the control limits, stop testing and begin a search for an assignable cause (see Section 6 for a list of possible assignable causes). Once the cause is corrected and the reference material's values are within the established control limits, testing can resume.

NOTE 2—Selected SRBs from SRB4, SRB5, SRB6, and SRB7 must be plotted on separate charts. Do not plot SRB6 and SRB7, for example D-6 and D-7, on the same chart.

4.6 Examples of typical x-charts are found in Figs. 1 and 2.

4.7 If only one reference material is used to regularly monitor testing performance, additional reference materials must be tested periodically to ensure that no systematic error or bias is affecting the test results. Test one or more of the reference materials not routinely used and compare the test results to the original baseline values to ensure that the testing system is still stable. Deviation from the original baseline values indicates the possibility of systematic testing error. If instability is suspected, all the reference materials should be tested and the results compared to the original baseline values. On a longer time frame basis, all the reference materials should be tested and compared to the original baseline values to determine the long-term testing stability. Initiate corrective action as indicated (see Section 6). If stability cannot be demonstrated, it may be necessary to apply a statistical correction (see Section 5).

4.8 A laboratory can estimate its testing precision relative to the "industry average" by calculating the three standard deviation values from its actual test data and comparing this to the control limit values for those tests given in Table 1, Table 2, or Table 3.

5. Guide to Accepted Normalization Practices for Carbon Black Test Methods

5.1 Accepted normalization practices for test methods found in Classification D1765, Table 1 are described below.

5.1.1 *Test Method D1510, Iodine Number*—Test Method D1510 contains instructions on how to perform a normalization of the test results. The HT or INR Iodine reference materials are recommended for monitoring iodine testing. The SRB HT and INR reference materials are specially prepared carbon blacks that have been shown to have stable iodine number values over a period of many years. If normalization is required, it shall be done using only the SRB HT or INR reference material values as given in Table 2 and Table 3. Typically, this test method does not require normalization unless the HT or INR reference material values are not within the published precision or accuracy limits, or both. The statistical correction described in Section 6 should not be used