Documentation Sheet
Sulfur Testing Reference Materials (STRM)
For Use with ASTM D1619, Carbon Black—Sulfur Content
(Evaluated per ASTM D4483-99)
Approved by D24.61: December 4, 2018[1]
Supersedes: None

Introduction

During the December 2017 D24 meeting, members were informed that one of the reference materials designated as "Standard B" was nearly depleted and in need of replacement. After a discussion of the history and the fact that little is known about these materials (see Appendix), it was agreed that it is unlikely that the Standard B material could be replaced with a similar material. The committee decided that it would be best to introduce a completely new set of sulfur testing reference materials or STRM's.

The committee selected five commercially available materials with nominal sulfur values ranging from near zero to 1.8% with values at approximately evenly spaced intervals of 0.4% to make up the new set of reference materials. These materials would be designated as STRM-A to STRM-E in increasing sulfur level order.

A quantity of each material was blended to improve homogeneity. An ITP was conducted in August 2018 to establish the reference values for each material and test method. The results of the analysis of the ITP data is shown in the tables below for Methods A and B. Only two laboratories reported results for Method C. Therefore, a precision table was not prepared for Method C.

Table 1 Precision Parameters for D1619, Sulfur Content of Carbon Black, (Type 1 Precision), Method A: Combustion w/IR Detection

		Units	Percent						
Material	Testing Period	Number of Laboratories (M/H/L)*	Mean Level	Sr	**r**	(r)	SR	**R**	(R)
STRM-A	Aug 2018	13(1/1/0)	0.010	0.002	**0.006**	57.767	0.008	**0.021**	212.250
STRM-B	Aug 2018	13(1/0/3)	0.362	0.007	**0.021**	5.666	0.023	**0.064**	17.649
STRM-C	Aug 2018	13(0/0/0)	0.790	0.011	**0.030**	3.817	0.042	**0.120**	15.198
STRM-D	Aug 2018	13(0/0/2)	1.236	0.012	**0.034**	2.770	0.048	**0.136**	11.020
STRM-E	Aug 2018	13(0/0/0)	1.867	0.027	**0.077**	4.131	0.074	**0.209**	11.190
			Preferred precision shown in bold text.						

*M = number of outliers for Mean; H = number of outliers for High variation; L = number of outliers for Low variation.

[1] The current version of this document is available from Laboratory Standards and Technologies, Inc., 227 Somerset Street, Borger, TX, 79007, www.carbonstandard.com.

Table 2 Precision Parameters for D1619, Sulfur Content of Carbon Black, (Type 1 Precision), Method B: X-ray Fluorescence									
		Units	Percent						
Material	Testing Period	Number of Laboratories (M/H/L)*	Mean Level	Sr	**r**	(r)	SR	**R**	(R)
STRM-A	Aug 2018	17(0/1/0)	0.016	0.005	**0.015**	98.105	0.017	**0.048**	310.799
STRM-B	Aug 2018	17(0/1/0)	0.346	0.005	**0.015**	4.219	0.030	**0.086**	24.908
STRM-C	Aug 2018	18(0/1/1)	0.754	0.012	**0.034**	4.515	0.067	**0.188**	24.986
STRM-D	Aug 2018	18(0/1/2)	1.204	0.014	**0.040**	3.324	0.054	**0.153**	12.685
STRM-E	Aug 2018	17(0/1/1)	1.838	0.020	**0.057**	3.088	0.057	**0.161**	8.764
			Preferred precision shown in bold text.						
*M = number of outliers for Mean; H = number of outliers for High variation; L = number of outliers for Low variation.									

Statistically, relative precision is the preferred precision for this test method because it is the least correlated with the mean level. However, in cases such as with STRM-A, the relative precision values can become very large as the mean level approaches zero. This makes the precision of that material appear to be out of alignment with the other values in the table. Also, since the unit for relative precision is percent, confusion can result when calculating and communicating precision information on a property whose unit of measurement is also percent, such as is the case with sulfur content in carbon black. Therefore, it is recommended that when evaluating the difference between two single test results, the absolute precision values r (within laboratory) and R (between laboratories) be used.

It is possible that the STRM set will last 20 years or more before it is necessary to produce a new set.

Shelf Life

Per ASTM D6915, Standard Practice for Carbon Black—Evaluation of Standard Reference Blacks, the shelf life of the Standard Reference Black (SRB) carbon blacks is indefinite when properly stored in a manner that protects it from exposure to sources of moisture, such as precipitation, other sources of liquid water, or high humidity environments. Since the STRM materials were selected from the SRBs, it is expected that their shelf life is also indefinite when properly stored as stated above.

Properties for the STRM Set

The testing conducted as part of the ITP and evaluation using ASTM D4483-99, generated accepted reference values (means), AR-values, for the D1619 test methods as shown below and 2 and 3 sigma limits on these values or on individual daily values as obtained by any laboratory using the STRM set. The 2 and 3 sigma limits apply to a single measurement of the listed test properties. Two times the 2 or 3 sigma limit equals the total 4 or 6 sigma range, respectively.

ASTM D4483-99 uses a one-sided k test to identify outliers having high variability. The ITP evaluation used a two-sided k test to identify outlier laboratories with variation that is statistically too high or too low when compared to the variation within the ITP data set. This approach is thought to better represent expected variability in real-world testing and helps to offset memory-bias from an individual's repeated testing of the same material(s). Mandel's h and

k statistics were used to identify outliers. Replacement values were calculated and substituted for outlier values.

'Accepted Reference Value' or AR-value; this is the average (mean), for the D1619 test methods and reference materials listed below in Tables 3 and 4, obtained in an ITP for a large group of typical laboratories using samples taken from the various material lots. See page 5 for more details on the ITP.

'Within Typical Laboratory' 2 and 3 sigma value; this is the within laboratory ± 2 and ± 3 standard deviation (Sr) value (for single measurements) on the STRM set AR-values for the D1619 test methods, as obtained from the same group of typical ITP laboratories.

'Between Typical Laboratory' 2 and 3 sigma value); this is the between laboratory ± 2 and ± 3 standard deviation (SR) value (for single measurements) on the STRM set AR-values for the D1619 test methods, as obtained from the same group of typical ITP laboratories.

2 sigma versus 3 sigma use considerations: Most carbon black test properties (with the exception of pellet hardness maximum) have an acceptable approximation to a normal distribution. With a normal distribution, 95.5% of all the test values are expected to fall within the limits of mean ± 2 sigma and 99.7% will fall within the limits of mean ± 3 sigma. This means that with only random variation present, approximately 1 in 20 results will fall outside the 2 sigma limits and 3 in 1000 will fall outside the 3 sigma limits. This means that when using 2 sigma limits the laboratory will be looking for a problem 1 in 20 test results when there is no problem to be found. This is a waste of valuable resources. On the other hand, when using 3 sigma limits the laboratory will be looking for a problem when there is not a problem only 3 in 1000 test results. However, if the consequences of allowing a problem to go undetected for a long time are too high, using 3 sigma limits may not give adequate warning in sufficient time to implement timely corrective action. Using 2 sigma limits will give an earlier warning of the presence of a problem. It is up to the user to balance the costs of untimely warnings versus the costs of searching for problems that do not exist.

Special consideration for bias: When no absolute reference material exists, such as is the case with carbon black testing, a laboratory's bias can be defined as the difference between its results and the mean result from an ITP involving many laboratories. Every laboratory can be expected to have some level of bias due to the unique combination of testing conditions (equipment, materials, manpower, environment, etc.) that exists within a given laboratory. The level of bias for a given laboratory may or may not be critical. A laboratory that did not participate in the ITP may find that it cannot maintain control within the control limits due to factors unique to that laboratory causing bias in its values, increased variation, or both. The laboratory should conduct an investigation to identify the presence and cause(s) of the bias and variation and eliminate them so that it is aligned with the ITP data. Participation in a multi-laboratory precision study, such as D24's LPRS program, may help to identify the unique sources of bias and variation. The STRM materials can be used to assist a laboratory in determining the presence and magnitude of bias and variation using the values given in Tables 3 and 4 below.

Using the STRM Set – When using any of the D1619 test methods, it is strongly recommended that laboratories determine if they are operating in an "in control" manner, by the use of the ± 2

or ± 3 sigma within-laboratory limits as the laboratory may choose to use. Despite rigorous analysis of the ITP data for the AR-value(s) and associated standard deviation(s), the group of laboratories in this (and any) ITP do not represent a typical "in statistical control system" to which the usual 6 sigma limits are applied. All the assignable causes of variation that are typically eliminated to attain 'statistical control' have not and cannot be, eliminated for the AR testing.

Normalization of the sulfur in carbon black test results is not recommended.

Table 3 Mean (AR-value) and Limit Values for D1619, Method A, Combustion w/IR Detection

Units	Percent	Within Laboratory			Between Laboratories		
Material	Mean (AR-value)	Sr	2 x Sr	3 x Sr	SR	2 x SR	3 x SR
STRM-A	0.010	0.002	0.004	0.006	0.008	0.015	0.023
STRM-B	0.362	0.007	0.014	0.022	0.023	0.045	0.068
STRM-C	0.790	0.011	0.021	0.032	0.042	0.085	0.127
STRM-D	1.236	0.012	0.024	0.036	0.048	0.096	0.144
STRM-E	1.867	0.027	0.055	0.082	0.074	0.148	0.221

Table 4 Mean (AR-value) and Limit Values for D1619, Method B, X-ray Fluorescence

Units	Percent	Within Laboratory			Between Laboratories		
Material	Mean (AR-value)	Sr	2 x Sr	3 x Sr	SR	2 x SR	3 x SR
STRM-A	0.016	0.005	0.011	0.016	0.017	0.034	0.051
STRM-B	0.346	0.005	0.010	0.015	0.030	0.061	0.091
STRM-C	0.754	0.012	0.024	0.036	0.067	0.133	0.200
STRM-D	1.204	0.014	0.028	0.042	0.054	0.108	0.162
STRM-E	1.838	0.020	0.040	0.060	0.057	0.114	0.171

Background and Interlaboratory Test Program Details:
STRM Set

Background - Standard Reference Blacks (SRBs), used for a number of test methods under the jurisdiction of ASTM Committee D24, are prepared according to D6915, "Evaluation of Standard Reference Blacks" with statistical analysis per D4483-99, “Standard Practice for Evaluating Precision for Test Method Standards in the Rubber and Carbon Black Manufacturing Industries”. Five SRBs materials were selected for use as the STRMs.

Evaluation of the STRM Set – The selection of the carbon blacks for the STRMs was based on their nominal sulfur results as known by the D24 members. Test results from all three D1619 methods were included in the ITP data, which was analyzed by test method and used to get the reference values for each D1619 test method. Only two laboratories reported results for Method C. Therefore, precision information is not presented for Method C.

The values listed in Tables 1 and 2 were obtained through the ITP and were analyzed per D4483-99. The values in Tables 3 and 4 are an expansion of values in Tables 1 and 2 to show the 2 standard deviation (sigma) and 3 standard deviation (sigma) limits to be used if a laboratory wishes to do control charting of the D1619 test methods.

For the STRM-A material, the mean value is near zero. Also, a zero value is the lower physical limit for a test result since a content parameter cannot have a negative value. Thus the expected normal distribution curve is truncated at the physical limit of zero. With the mean near zero and the curve truncated at zero, a test result at or near zero actually has the highest probability of being reported for all test results. Therefore, reporting all test results with the same value at or near zero would be expected. It would be inappropriate to exclude those laboratories that reported the same value at or near zero for all of their test results. To exclude them would mis-represent the test precision. Therefore, the variability data for those laboratories who were flagged as an outlier by the 2-sided k test lower limit test was retained in the STRM-A data for both test methods.

Interlaboratory Test Program (ITP) – Twenty pounds of each material were blended to improve material homogeneity. Samples were distributed to 32 laboratories and tested in August 2018. The data analyses were completed in September 2018 and used to prepare the included tables. D24 members approved the STRMs for use on December 4, 2018. Tables 1 and 2 show the number of participating laboratories for each material and test method as well as the number of outliers removed for mean (M), high variability (H), and low variability (L).

To report corrections or request changes to this document, contact Laboratory Standards and Technologies, the chairman of ASTM subcommittee D24.61, or the chairman of ASTM subcommittee D24.66.

Appendix

There are no known existing records concerning the original ASTM sulfur reference materials. Discussions with the persons involved in the selection of this initial set provides no information on the materials except that the selection probably occurred about 1995-1996. Therefore, little is known about when these materials were produced, who produced them, where they were produced, or their carbon black grades except the material designated as “Standard A”, which is known to be an N990 carbon black.

An examination of the ASTM Volume 09.01 books from 1997 and 1998 shows that values for this initial set were first published in the 1998 book in D1619-97. In 1997, revisions to D1619 were approved that 1) removed Oxygen Bomb Calorimeter as Method A, 2) removed High-Temperature Combustion with Iodimetric Detection as Method B, 3) redesignated High-Temperature Combustion with Infrared Absorption Detection as Method A, 4) introduced X-ray Fluorescence as Method B, and 5) included the initial set of reference materials as part of Method B. This initial set consisted of four grades of carbon black and had no designation as a reference material as presented in ASTM D7849, which was developed much later.

Nothing is known about how the values given in D1619 were developed, such as an Interlaboratory Test Program (ITP), how many laboratories participated, or how the values were calculated other than the typical statistical equations for mean and standard deviation were probably used. There probably was an ITP because it is known that the values were based on D1619 Method A, High-Temperature Combustion with Infrared Absorption Detection. Method A has a detection limit of 0.01 percent and therefore, the values are reported to two decimal places. The impact this has on using ASTM sulfur standards for x-ray fluorescence testing is that, while the equipment may have a lower detection limit, the calibration is only good to two decimal places. Test results can only be reported to two decimal places regardless of what the instrument may display.