Designation: E2554 – 13

An American National Standard

Standard Practice for Estimating and Monitoring the Uncertainty of Test Results of a Test Method Using Control Chart Techniques[1]

This standard is issued under the fixed designation E2554; 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 describes techniques for a laboratory to estimate the uncertainty of a test result using data from test results on a control sample. This standard provides one method for a laboratory to estimate Measurement Uncertainty in accordance with Section A22.3 in *Form and Style of ASTM Standards*.

1.2 Uncertainty as defined by this practice applies to the capabilities of a single laboratory. Any estimate of uncertainty determined through the use of this practice applies only to the individual laboratory for which the data are presented.

1.3 The laboratory uses a well defined and established test method in determining a series of test results. The uncertainty estimated using this practice only applies when the same test method is followed. The uncertainty only applies for the material types represented by the control samples, and multiple control samples may be needed, especially if the method has different precision for different sample types or response levels.

1.4 The uncertainty estimate determined by this practice represents the intermediate precision of test results. This estimate seeks to quantify the total variation expected within a single laboratory using a single established test method while incorporating as many known sources of variation as possible.

1.5 This practice does not establish error estimates (error budget) attributed to individual factors that could influence uncertainty.

1.6 This practice describes the use of control charts to evaluate the data obtained and presents a special type of control chart to monitor the estimate of uncertainty.

1.7 The system of units for this standard is not specified. Dimensional quantities in the standard are presented only as illustrations of calculation methods. The examples are not binding on products or test methods treated.

1.8 *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]

D5184 Test Methods for Determination of Aluminum and Silicon in Fuel Oils by Ashing, Fusion, Inductively Coupled Plasma Atomic Emission Spectrometry, and Atomic Absorption Spectrometry

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

E456 Terminology Relating to Quality and Statistics

E2282 Guide for Defining the Test Result of a Test Method

E2587 Practice for Use of Control Charts in Statistical Process Control

ISO/ASTM 51707 Guide for Estimating Uncertainties in Dosimetry for Radiation Processing

2.2 *ASTM Publications:*[2]

Form and Style for ASTM Standards

Manual on Presentation of Data and Control Chart Analysis 7th Edition

2.3 *ISO Standard:*[3]

ISO/IEC 17025 General Requirements for the Competence of Testing and Calibration Laboratories

3. Terminology

3.1 *Definitions*—The terminology of Terminology E456 applies to this practice except as modified herein.

3.1.1 *control sample, n*—sample taken from a stable, homogeneous material for the purposes of monitoring the performance of a test method in a laboratory.

3.1.1.1 *Discussion*—The control sample material is representative of the product typically tested in the laboratory by the

[1] This practice is under the jurisdiction of ASTM Committee D05 on Coal and Coke and is the direct responsibility of Subcommittee D05.07 on Physical Characteristics of Coal.

Current edition approved Oct. 1, 2007. Published May 2013. Originally approved in 2007. Last previous edition approved in 2007 as E2554 – 07. DOI: 10.1520/E2554-13.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036, http://www.ansi.org.

given test method. A control sample is run periodically using the complete test method protocol to develop a test result. Such test results may be statistically evaluated to monitor test method performance over time. It is not necessary to have an accepted reference value assigned to the control sample material. When the current material is nearly consumed, a replacement material should be run in parallel with the current material to ensure continuity in the control sample program.

3.1.2 *check sample, n*—see *control sample*.

3.1.3 *intermediate precision, n*—the closeness of agreement between test results obtained under specified intermediate precision conditions. **E177**

3.1.3.1 *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."

3.1.3.2 *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.

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

3.1.5 *repeatability, n*—precision under repeatability conditions. **E177**

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *uncertainty control chart, n*—control chart that includes control limits based on the variation attributed to the uncertainty of the test method.

4. Significance and Use

4.1 This practice provides one way for a laboratory to develop data-based Type A estimates of uncertainty as referred to in Section A22 in *Form and Style of ASTM Standards*.

4.2 Laboratories accredited under ISO/IEC 17025 are required to present uncertainty estimates for their test results. This practice provides procedures that use test results to develop uncertainty estimates for an individual laboratory.

4.3 Generally, these test results will be from a single sample of stable and homogeneous material known as a control or check sample.

4.4 The true value of the characteristic(s) of the control sample being measured will ordinarily be unknown. However, this methodology may also be used if the control sample is a reference material, in which case the test method bias may also be estimated and incorporated into the uncertainty estimate. Many test methods do not have true reference materials available to provide traceable chains of uncertainty estimation.

4.5 This practice also allows for ongoing monitoring of the laboratory uncertainty. As estimates of the level of uncertainty change, possibly as contributions to uncertainty are identified and minimized, revision to the laboratory uncertainty will be possible.

5. General Considerations

5.1 *Materials to be Used:*

5.1.1 This methodology requires a quantity of stable and homogeneous material which will serve as the source of control samples (sometimes called check samples). The material shall be similar in composition to the samples of material routinely analyzed by this test method in this laboratory. By stable it is assumed that the test results obtained from this material should be consistent over the time interval that this material will be used. By homogeneous it is assumed that samples taken from the material source will not have a significant variation in the characteristic measured by the test method.

5.1.2 For destructive testing of control sample materials, provision shall be made for depletion and replacement of the control sample material.

5.1.2.1 In some cases, the test method may be nondestructive and the same material may be reused indefinitely.

5.1.2.2 In other cases, the material may be used up, deteriorate, or otherwise gradually change.

5.1.3 The test method should describe the best practices for preparing and storing the control material and taking the control samples.

5.2 *Test Conditions:*

5.2.1 An uncertainty estimation program should be designed to include all known sources of variation, such as operators (analysts), equipment, reagents, and so forth, and these should be deliberately incorporated into the design of the program. In general, these sources of variation will be defined (including acceptable tolerances) by the test method.

5.2.2 In cases in which control over such variations is not possible or undefined, at least 30 to 50 sampling periods shall be evaluated to permit environmental and other factors to be incorporated in the overall estimate.

6. Overall Procedure—Control Charting Methods

6.1 General concepts of control charts are described elsewhere. For more information, see Practice E2587 as well as Manual 7A.[4]

6.2 The general procedure involves two major phases: Preliminary and Monitoring.

6.2.1 *Preliminary Phase:*

6.2.1.1 This phase begins with an initial collection of test results.

6.2.1.2 Preliminary control charts are then prepared and examined. These charts are evaluated to determine if the process is in a state of statistical control. The usual principles of control charting utilize short-term variability to estimate the limits within which samples of test results should vary. For control sample programs this short-term variability is equivalent to repeatability precision. It is expected, however, that

[4] *Manual on Presentation of Data and Control Chart Analysis: 7th Edition*, ASTM International, West Conshohocken, PA, 2001.

additional contributions to variation will be present over time and therefore additional variation, equivalent to intermediate precision, will be encountered.

6.2.1.3 An estimate of uncertainty standard deviation is developed.

6.2.1.4 An uncertainty control chart is then prepared to monitor future sample results.

6.2.2 *Monitoring Phase:*

6.2.2.1 The proposed uncertainty control chart is used to provide evidence that the estimate of uncertainty is not exceeding the estimated value.

6.2.2.2 The estimate of uncertainty should be periodically re-evaluated.

6.2.2.3 Where appropriate, it is recommended that a standard control chart also be maintained to determine whether the variation over time has been reduced to the level of short-term variation (repeatability).

6.3 Two types of control charting methods are recommended to develop estimates of uncertainty. These include:

6.3.1 Mean (*X*bar) and range or standard deviation charts are used when multiple test results are conducted in each time period.

6.3.2 Individual charts (Ind*X*) are used when single test results are obtained in each time period.

6.4 *Variation Estimates:*

6.4.1 Either a range chart or a standard deviation chart may be used to estimate the short-term variability when multiple assays are conducted under repeatability conditions per time period. An estimate from the control chart data can be compared to other estimates of repeatability (within laboratory, short-term variation) if available.

6.4.2 Sample averages are examined and may provide estimates of variation caused by other factors. Such factors may include environmental effects, operator factors, reagents, or instruments.

6.5 *Systematic Procedures:*

6.5.1 Specifically designed experiments can be used to ensure all known sources of variation, such as operators (analysts), equipment, reagents, or instruments are incorporated in the general study.

6.5.2 The data generated from this program is available for additional uses, such as control charting to evaluate trends, stratification by analysis, or stratification by equipment to identify training or maintenance needs or both.

7. Specific Procedures

7.1 *Multiple Test Results Generated Per Time Period:*

7.1.1 A specified number of independent test results are taken during each time period. Generally this number is 5 or less. It is preferred that at least 25 sets of test results be obtained before developing the charts.

7.1.2 Either a range chart or a standard deviation chart is prepared. This is examined for special cause variation. If the variability appears random then an estimate of repeatability is computed. This may be done by pooling the sums of squares, using the average standard deviation, or using the average range.

NOTE 1—If the ranges or standard deviations are zero in most of the samples, then this estimate of repeatability standard deviation is suspect and probably unusable. This is usually the result of insufficient resolution of the measurement system in use or severe rounding. An estimate based on the minimum interval size should be substituted for the zeros. As a rule of thumb, consider replacing the zeros when more than about ⅓ are zeros.

7.1.3 A means chart is used to examine variation among time periods. Limits on this chart permit comparison of variation between time periods using repeatability as the estimate of error.

7.1.3.1 If the control chart shows a state of statistical control then the uncertainty will be assumed approximately equivalent to the repeatability standard deviation.

7.1.3.2 In most cases it will be expected that the variability between means will show an “out of control” condition indicating that there are “special” causes of variation in addition to repeatability. The between means variation and within means repeatability estimates are then used to compute an estimate of uncertainty standard deviation.

7.1.4 Using the estimate of uncertainty standard deviation an Uncertainty Control Chart is prepared for future monitoring of the uncertainty. This chart may include control limits for means as a possible lower set of control limits along with the uncertainty control limits based on the estimate of uncertainty.

7.2 *Individual Tests:*

7.2.1 Single tests are generated at each time period. Variation among these results is evaluated.

7.2.2 In some cases, it is possible to incorporate external estimates of repeatability obtained from prior or concurrent studies.

8. Multiple Readings Per Time Period

8.1 *Example 1—Absorbance of Radiochromic Dosimeters:*

8.1.1 Over a period of several days, different sets of three dosimeters were irradiated to the same nominal dose. The irradiation was conducted under standard conditions at a single irradiator facility. Possible sources of random errors could include intrinsic variation in dosimeter response and day-to-day variations in the physical environment, for example, temperature, positioning of dosimeters within the irradiator, and shielding. The data was presented in Guide ISO/ASTM 51707.

8.2 Table 1 consists of three dosimeters irradiated and measured on a single day. Nine time periods are shown. The averages and standard deviations are computed for each time period.

8.3 Prepare a standard deviation control chart.

NOTE 2—Ranges could have been computed instead of standard deviations and a range control chart would be prepared.

8.3.1 Compute the average of the standard deviations (p test periods):

$$\bar{s} = \sum s/p = 0.04489/9 = 0.0050 \quad (1)$$

NOTE 3—If the standard deviations in many of the samples were zero, then we recommend replacing the values of 0 with a value calculated as: Half-interval/$\sqrt{3}$. In this case the intervals are 0.001 and the half-interval would be 0.0005. Then the estimate of s in place of zero would be 0.0005/1.732 = 0.00028.

TABLE 1 Multiple Dosimeters Irradiated on Each Day (Data From Guide ISO/ASTM 51707)

Test No.	Rep 1	Rep 2	Rep 3	Average	*s*
1	0.282	0.274	0.276	0.2773	0.0042
2	0.294	0.274	0.284	0.2840	0.0100
3	0.300	0.284	0.292	0.2920	0.0080
4	0.290	0.300	0.292	0.2940	0.0053
5	0.296	0.294	0.297	0.2957	0.0015
6	0.290	0.278	0.284	0.2840	0.0060
7	0.290	0.290	0.290	0.2900	0.0000
8	0.278	0.288	0.286	0.2840	0.0053
9	0.284	0.292	0.292	0.2893	0.0046

8.3.2 The control limits for the standard deviation control chart are found as:

$$\text{UCLs} = B_4 * \bar{s} = 2.568 * 0.0050 = 0.0128 \quad (2)$$

and

$$\text{LCLs} = B_3 * \bar{s} = 0 * 0.0050 = 0 \quad (3)$$

8.3.2.1 The control chart factors B_3 and B_4 for sample sizes up to n=6 can be found in Table 2. For larger sample sizes see Manual 7A.

8.3.3 The control chart is prepared to evaluate the within-sample or time period variation. Control limits as computed are displayed. See Fig. 1.

8.3.4 The standard deviation chart is examined for unusual values. No readings appear to be unusual.

8.4 A control chart for means is prepared by plotting the means.

8.4.1 The sample means are averaged. The grand average, $\bar{\bar{X}}$ is 0.2878.

8.4.2 The control limits for the means control chart are found as:

$$\text{UCL} = \bar{\bar{X}} + A_3\,\bar{s} = 0.2878 + 1.954 * 0.0050 = 0.2976 \quad (4)$$

$$\text{LCL} = \bar{\bar{X}} - A_3\,\bar{s} = 0.2878 - 1.954 * 0.0050 = 0.2781 \quad (5)$$

8.4.2.1 The control chart factors A_3 for sample sizes up to n=6 can be found in Table 2. For larger sample sizes, see Manual 7A.

8.4.3 The control chart limits are plotted as presented in Fig. 2.

8.4.3.1 Examination of the means control chart is conducted to determine whether variation between periods appears to be greater than expected from within sample variation. In this example, there are samples just at, and even beyond, the control limits, which is an indicator that the variation over time is much greater than would be expected based only on within-sample repeatability.

8.5 Estimate the within sample standard deviation. This is an estimate of a single laboratory repeatability standard deviation.

8.5.1 A direct estimate of single laboratory standard deviation is calculated based on the "pooled" variances. This is found by: calculating the squares of each standard deviation; summing the squares; dividing by the number of samples; and taking the square root. In this example:

$$\text{Sum of squares of standard deviations} = \sum s^2 = 0.000297 \quad (6)$$

s_r = estimate of single laboratory repeatability standard deviation

$$= \sqrt{\frac{\sum s^2}{p}} = \sqrt{\frac{0.000297}{9}} = 0.0057$$

8.5.2 An alternative estimate of single laboratory repeatability standard deviation can be computed from the average *s* as:

$$s_r = \bar{s}/c_4 = 0.0050/0.8862 = 0.0056 \quad (7)$$

NOTE 4—When ranges are used instead of standard deviations, an estimate of s_r is found from the average range. In this example, the average range would be found as 0.0097 and the estimate of standard deviation is then found:

$$s_r = \bar{R}/d_2 = 0.0097/1.693 = 0.0057 \quad (8)$$

where:

$\bar{R}$ = average range.

8.5.2.1 Factors c_4 and d_2 for sample sizes up to n=6 can be found in Table 2. For larger sample sizes, see Manual 7A.

8.6 *Between Time or Sample Variation:*

8.6.1 Since there is a between sample or between time variation, an estimate of the between time standard deviation is

TABLE 2 Factors for Computing Control Chart Lines

NOTE 1—These values are extracted from Table 49 of ASTM Manual on Presentation of Data and Control Chart Analysis.

Observations in Sample, n	Chart for Averages		Chart for Standard Deviations			Chart for Ranges		
	Factors for Control Limits		Factors for Central Line	Factors for Control Limits		Factors for Central Line	Factors for Control Limits	
	A2	A3	C4	B3	B4	d2	D3	D4
2	1.880	2.659	0.7979	0	3.267	1.128	0	3.267
3	1.023	1.954	0.8862	0	2.568	1.693	0	2.575
4	0.729	1.628	0.9213	0	2.266	2.059	0	2.282
5	0.577	1.427	0.9400	0	2.089	2.326	0	2.114
6	0.483	1.287	0.9515	0.030	1.970	2.534	0	2.004

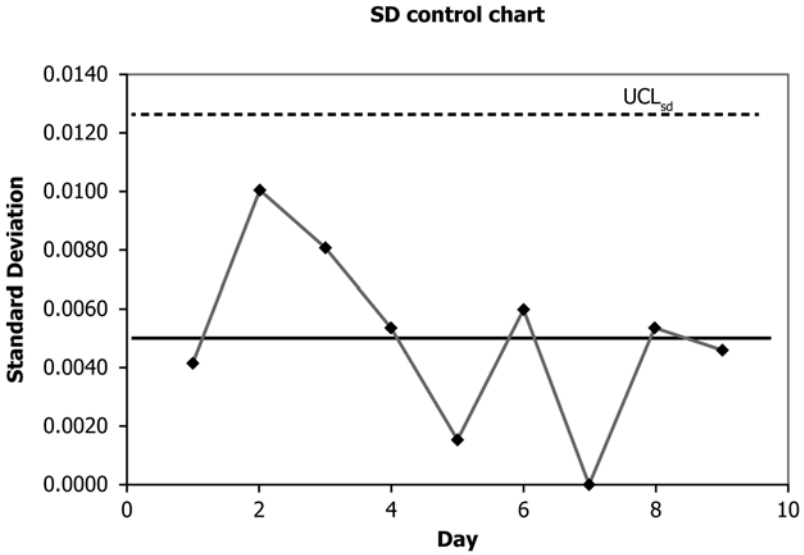

FIG. 1 Standard Deviation Control Chart

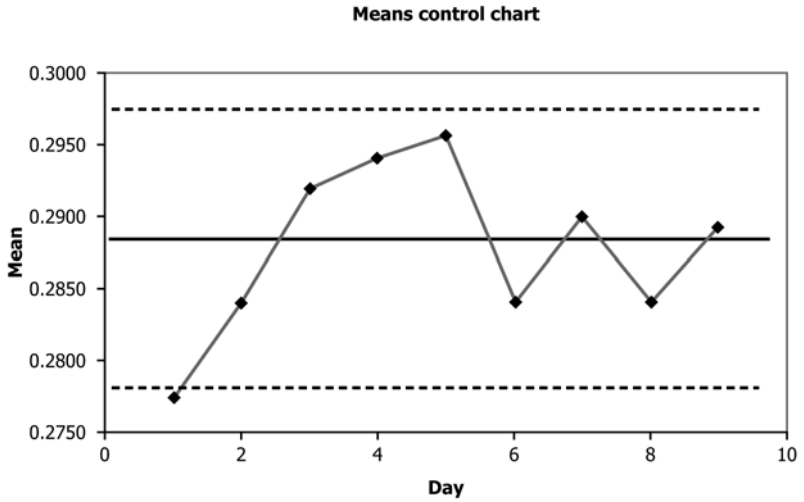

FIG. 2 Control Chart for Means

then computed. First the standard deviation among the sample averages is found. This was computed as 0.00590.

8.6.2 The s_{time} is then computed as:

$$s_{time} = \sqrt{s_{\bar{x}}^2 - \frac{s_{within}^2}{n_{within}}} = \sqrt{s_{\bar{x}}^2 - \frac{s_r^2}{n_{within}}} = \sqrt{0.00590^2 - \frac{0.0057^2}{3}} = 0.0049 \quad (9)$$

where:

$s_{\bar{x}}$ = the standard deviation of the averages,
n_{within} = number of repeats (3),
s_{within} = standard deviation within groups and is equivalent to s_r = single laboratory repeatability standard deviation.

NOTE 5—If the difference under the radical sign is negative, meaning the estimate of s_{time}^2 is negative, then this may be interpreted as indicating that the variation associated with time is negligible and the estimate of s_{time} is set to zero.

8.7 The Uncertainty standard deviation is estimated from a single time and a single repeat.

$$S_u = \sqrt{s_{time}^2 + s_r^2} = \sqrt{0.0049^2 + 0.0057^2} = 0.00753 \quad (10)$$

NOTE 6—This value is equivalent to an estimate of intermediate precision based on multiple time periods.

8.8 An uncertainty control chart is then prepared to monitor future samples. All the initial values should show in control state. The Uncertainty Control Limits are established as defined in Eq 12 and 13 and added to the chart (Fig. 3). First we compute the standard deviation for sample averages assuming there is variation due to time and repeatability. This is an estimate of the uncertainty associated with samples, not individual test results, and is found as:

$$s_{u-aves} = \sqrt{s_{time}^2 + \frac{s_r^2}{n}} = 0.00590 \quad (11)$$

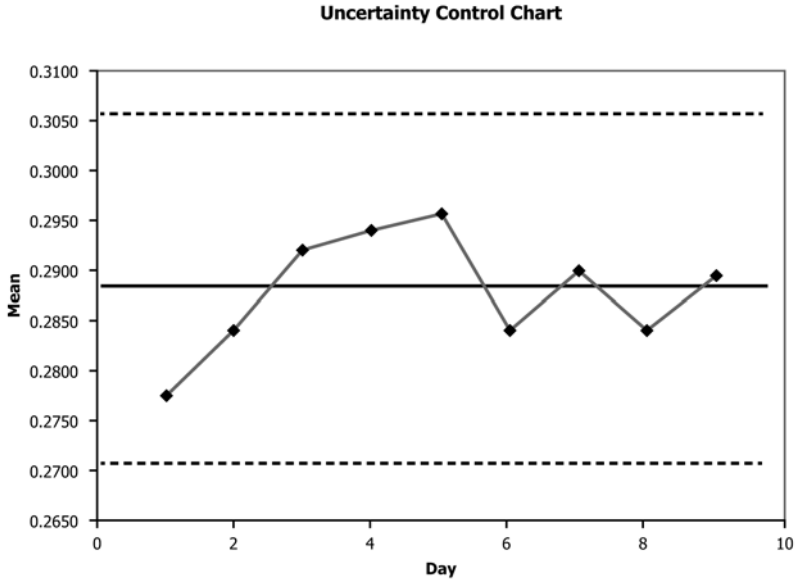

FIG. 3 Uncertainty Control Chart

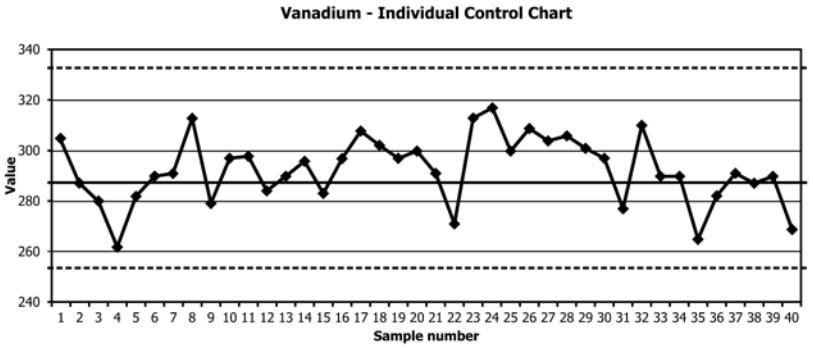

FIG. 4 Values Including the Limits for 40 Samples

$$UCL_{Uncertainty} = \bar{\bar{X}} + 3\ s_{u-aves} = 0.2878 + 3*0.00590 = 0.2878 + 0.0177 = 0.3055 \quad (12)$$

$$LCL_{Uncertainty} = \bar{\bar{X}} - 3\ s_{u-aves} = 0.2878 - 3*0.00590 = 0.2878 - 0.0177 = 0.2701 \quad (13)$$

NOTE 7—The use of the multiplier of 3 is in keeping with traditional control charting practices and are interpreted in a similar manner when used to monitor the process.

NOTE 8—The initial calculation of the uncertainty standard deviation for samples, Eq 11, is mathematically equivalent to the standard deviation of the averages, $s_{\bar{x}}$

8.8.1 Averages should continue to be plotted on the uncertainty control chart. If any points go beyond the limits this should be a signal to investigate for possible causes. An unusual number of points at or beyond the limits may indicate that the estimate of uncertainty is too small and should be recalculated.

9. Individual Tests

9.1 A reference material or a stable and homogenous material representative of the material normally tested can be used.

9.1.1 A minimum of about 25 sample test results should be completed with reasonable time intervals between readings. The standard deviation of this set of individual readings will serve as an estimate of uncertainty.

9.1.2 This estimate of standard deviation shall be used to provide "control limits" for the review of past results and ongoing monitoring of the method. Limits of overall average ±3 standard deviations would be appropriate.

9.1.3 A graph showing the sequence of readings with the limits added allows examination to determine if unusual conditions may be present and contributing to the overall variation.

9.2 *Example 2—Vanadium in Oil:*

9.2.1 Vanadium is an impurity in oil. A critical level is approximately 300 ppm. Test Method D5184 is followed to determine the level of impurities. A quantity of control material was prepared using a typical batch of oil with a level of impurities similar to those encountered in practice (vanadium at approximately 300 ppm). This overall batch was subdivided into multiple containers and individual samples were drawn daily. Each sample was processed and tested along with regular samples in a single laboratory following standard procedures.

9.2.2 Multiple operators, different days, and conditions as generally experienced were incorporated into the program.

9.2.3 Forty consecutive samples were taken and are listed in Table 3 (units of mg/kg). These are used to provide an estimate of the variation of the test.

9.2.4 The average and standard deviation of the 40 values are found as: average = 292.5 and standard deviation (sd) = 13.3.

9.2.5 Based on these values, limits that would include 99.75 % of all readings would be within ±3 sd of the average. These limits would be:

TABLE 3 Forty Vanadium Samples, mg/kg

Sample	Vanadium	Sample	Vanadium
1	305	21	291
2	287	22	271
3	280	23	313
4	262	24	317
5	282	25	300
6	290	26	309
7	291	27	304
8	313	28	306
9	279	29	301
10	297	30	297
11	298	31	277
12	284	32	310
13	290	33	290
14	296	34	290
15	283	35	265
16	297	36	282
17	308	37	291
18	302	38	287
19	297	39	290
20	300	40	269

$$292.5 + 3*13.3 = 332.4 \quad (14)$$

and

$$292.5 - 3*13.3 = 252.7 \quad (15)$$

9.2.6 A graph of the values including the limits for these 40 samples is shown in Fig. 4. No points fall outside the limits.

9.2.7 The value of standard deviation then serves as an estimate for the uncertainty standard deviation.

9.2.8 This uncertainty standard deviation is considered the laboratory's uncertainty for vanadium determinations at about 300 ppm.

10. Ongoing Studies

10.1 Additional factors should be introduced and identified over time. These should include, when appropriate, different operators, recalibration of equipment, weather changes (temperature, humidity), and so forth.

10.2 It is expected that these additional factors will result in an increase in the magnitude of the uncertainty estimate.

10.3 The uncertainty shall be reviewed to determine if there are changes either as improvement or a worsening of the degree of variation within laboratories.

11. Keywords

11.1 control chart; control sample; intermediate precision; test result; uncertainty