ASTM INTERNATIONAL

Designation: E2554 – 18ε1

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.

ε1 NOTE—Editorial corrections were made throughout in September 2019.

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 practice provides one method for a laboratory to estimate Measurement Uncertainty in accordance with Section A22.3 in *Form and Style for 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, health, and environmental practices and determine the applicability of regulatory limitations prior to use.*

1.9 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

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

E92 Test Methods for Vickers Hardness and Knoop Hardness of Metallic Materials

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

E384 Test Method for Microindentation Hardness of Materials

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

2.2 *ASTM Publications:*[2]

Form and Style for ASTM Standards

Manual 7A 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

---

[1] This practice is under the jurisdiction of ASTM Committee E11 on Quality and Statistics and is the direct responsibility of Subcommittee E11.50 on Metrology.
Current edition approved April 1, 2018. Published April 2018. Originally approved in 2007. Last previous edition approved in 2013 as E2554 – 13. DOI: 10.1520/E2554-18E01.

[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.

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 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 for 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.

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

2

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 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—Knoop Hardness Testing:*

8.1.1 A Knoop test is an indentation microhardness test which uses a diamond tip indenter with very light loads in accordance with Test Methods E92 and E384. The tip is ground to have two diagonal lengths. An indentation is performed at a designated load and the longer diagonal length is measured using a measuring microscope. The length is then converted mathematically to a Knoop Hardness, HK.

8.1.2 Periodic verification, or checking of the instrument, is to be conducted on the particular testing machine by taking several measurements on a reference test block. The test block is certified by a reference laboratory using a set of tests conducted on a reference machine.

8.1.3 A set of several hundred samples, first of size 5, and later of size 3, were conducted on a single instrument over a multi-year period.

8.2 Because control charting is being conducted, the initial study should consist of at least 30 or more sample readings, or subgroups. In Table 1 the first 30 subgroups conducted with the instrument and test block are listed. To remain consistent throughout only the first 3 readings are included for each sample or subgroup.

8.3 Prepare a range control chart.

TABLE 1 Initial Length Readings of Knoop Test (Long Diagonal Dimension)

Sample	d_1	d_2	d_3	Range	Mean
1	114.1	113.5	114.0	0.6	113.87
2	115.7	114.8	115.1	0.9	115.20
3	115.3	114.6	114.5	0.8	114.80
4	115.0	114.5	114.7	0.5	114.73
5	116.2	116.1	116.0	0.2	116.10
6	114.2	114.0	114.5	0.5	114.23
7	115.4	114.5	114.9	0.9	114.93
8	116.1	114.2	115.2	1.9	115.17
9	114.6	114.2	114.7	0.5	114.50
10	115.7	113.9	115.0	1.8	114.87
11	115.1	115.2	114.8	0.4	115.03
12	114.6	114.3	115.1	0.8	114.67
13	115.4	114.8	113.9	1.5	114.70
14	115.9	114.1	114.5	1.8	114.83
15	115.4	114.6	114.5	0.9	114.83
16	115.9	115.0	115.3	0.9	115.40
17	115.3	114.6	114.7	0.7	114.87
18	114.5	114.7	114.3	0.4	114.50
19	116.6	115.5	114.3	2.3	115.47
20	115.2	115.5	116.0	0.8	115.57
21	115.7	114.6	115.7	1.1	115.33
22	115.7	114.4	114.7	1.3	114.93
23	115.3	115.3	115.6	0.3	115.40
24	116.1	114.5	114.7	1.6	115.10
25	115.1	115.2	115.1	0.1	115.13
26	115.5	114.4	114.9	1.1	114.93
27	116.2	115.0	115.4	1.2	115.53
28	116.1	114.9	115.1	1.2	115.37
29	115.3	116.0	115.9	0.7	115.73
30	115.2	114.1	114.3	1.1	114.53

Note 2—Standard deviations could have been computed instead of ranges of each sample, and a standard deviation control chart would be prepared.

8.3.1 Compute $\bar{R}$, the average of the ranges (where p is the number of subgroups = 30):

$$\bar{R} = \frac{\Sigma R}{p} = 28.8/30 = 0.96 \quad (1)$$

8.3.2 The upper and lower control limits for the range control chart are found as:

$$\text{UCL}_R = D_4 \times \bar{R} = 2.575 \times 0.96 = 2.47 \quad (2)$$

and:

$$\text{LCL}_R = D_3 \times \bar{R} = 0 \times 0.96 = 0 \quad (3)$$

The control chart factors D_3 and D_4 for sample sizes up to n = 6 are found in Table 2. For larger sample sizes see Manual 7A.

Note 3—Observe that for samples up to size 6 the lower limit for ranges will always be zero. Thus, for some samples we will not be surprised if all the readings were identical. However, if a large percent of

the sample ranges are zero then the resolution of the measurement system is probably insufficient to observe the natural variation.

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 range control chart is examined for unusual values. No readings are outside the limits.

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

8.4.1 The sample means are averaged. The grand average of the 30 means is:

$$\bar{\bar{X}} = \frac{\Sigma \text{means}}{p} = 3450.25/30 = 115.01 \quad (4)$$

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

$$\text{UCL}_{\text{means}} = \bar{\bar{X}} + A_2\bar{R} = 115.01 + 1.023 \times 0.96 = 115.99 \quad (5)$$

$$\text{LCL}_{\text{means}} = \bar{\bar{X}} - A_2\bar{R} = 115.01 - 1.023 \times 0.96 = 114.03 \quad (6)$$

TABLE 2 Factors for Computing Control Chart Central Lines and Limits

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		
	A_2	A_3	C_4	B_3	B_4	d_2	D_3	D_3	
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	