



Designation: D6543 – 17

# Standard Guide to the Evaluation of Measurements Made by Online Coal Analyzers<sup>1</sup>

This standard is issued under the fixed designation D6543; 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 ( $\epsilon$ ) indicates an editorial change since the last revision or reapproval.

## 1. Scope

1.1 This guide provides techniques to be used for the evaluation of the measurement performance of online coal analyzers.

1.2 *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.*

1.3 *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:*<sup>2</sup>

[D121 Terminology of Coal and Coke](#)

[D2013 Practice for Preparing Coal Samples for Analysis](#)

[D2234/D2234M Practice for Collection of a Gross Sample of Coal](#)

[D6518 Practice for Bias Testing a Mechanical Coal Sampling System \(Withdrawn 2008\)](#)<sup>3</sup>

[D7430 Practice for Mechanical Sampling of Coal](#)

[E178 Practice for Dealing With Outlying Observations](#)

## 3. Terminology

3.1 *Definitions*—For additional definitions of terms used in this standard, refer to Terminology [D121](#).

3.2 *Definitions of Terms Specific to This Standard:*

<sup>1</sup> This guide is under the jurisdiction of ASTM Committee [D05](#) on Coal and Coke and is the direct responsibility of Subcommittee [D05.23](#) on Sampling.

Current edition approved June 1, 2017. Published June 2017. Originally approved in 2000. Last previous edition approved in 2015 as D6543-15. DOI: 10.1520/D6543-17.

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

<sup>3</sup> The last approved version of this historical standard is referenced on [www.astm.org](http://www.astm.org).

3.2.1 *analyzer system, n*—a coal quality measurement system which includes an online coal analyzer and which may include one or more stages of a coal-sampling system.

3.2.2 *calibration, n*—mathematical modeling of analyzer and comparative coal sampling and analysis data. Factors from the model are used in the online analyzer control software.

3.2.3 *full-stream analyzer, n*—an analyzer system that interrogates the coal on a process belt.

3.2.4 *Latent Variable Model, n*—a mathematical model that can estimate each system's precision, when the analyzer is compared to two independent reference systems.

3.2.5 *online analyzer, n*—an analytical tool consisting of an instrument and systems, which together provide measurements, or estimates, or both, of coal quality parameters.

3.2.6 *outlier, n*—an extreme value that statistical tests indicate to be far enough from other results in a population under consideration to cause suspicion that the value is not a member of the population.

3.2.7 *reference material, n*—material of stable composition that may be used to generate static analyzer measurements.

3.2.8 *reference system, n*—a measurement system used to measure the characteristics of a lot of coal that are also measured by an online analyzer, and against which the online analyzer measurements are compared.

3.2.9 *sample stream analyzer, n*—an analyzer system that is fed a save or reject stream from a sampling system.

3.2.10 *standardization, n*—calibration of an instrument to a reference material using static stability measurements.

3.2.11 *static stability, n*—an estimate of the measurement precision of an instrument obtained on material that is not moving. The estimate normally is expressed as the standard deviation and average of the measurements for a given period of time.

3.2.12 *synchronization error, n*—an error that occurs from comparing measurements made by an online analyzer and a reference system that are not measuring exactly the same lot because of temporal and/or spatial offsets.

## 4. Summary of Guide

4.1 This guide describes how to measure performance of an online analyzer using comparative measurements. The performance evaluation consists of a paired comparison of analyses from a reference method using ASTM sampling, sample preparation, and analysis methods for several lots of coal with the analyses from the online analyzer for the same lots of coal. The data resulting from the comparative test may be evaluated using graphical and statistical techniques outlined below.

4.2 Various techniques are recommended by online analyzer manufacturers for standardization or static testing. These techniques are useful for establishing a benchmark before conducting a comparative test. These techniques may also be used as diagnostic tests in accordance with methods recommended by online analyzer manufacturers and graphical and statistical techniques included in this guide.

## 5. Significance and Use

5.1 Online analyzers are used to provide quality data on lots of coal. The resulting quality data are used as a production tool or for some contractual application. This guide provides the means of evaluating the analyzer system and the data produced.

5.2 Become familiar with the document's terminology and layout. The section on test design and data collection will provide the means by which all the analysis data will be gathered. The test should be carefully designed to ensure the user's requirements are met.

5.3 The procedures defined in this guide can be used to estimate the accuracy and precision of an online analyzer, (1) to conduct acceptance testing following installation and (2) to monitor the accuracy and precision (a) during routine use (quality control), (b) when significant changes are made to the analyzer, and (c) when a significant change in the coal being analyzed occurs (e.g., a different seam at a mine, or a new coal source at a power plant). These procedures can also be used for calibration purposes.

## 6. Selection and Conduct of Performance Evaluations

### 6.1 Introduction:

6.1.1 Several techniques can be used to evaluate the performance of an online analyzer. These techniques provide data that can be evaluated by using the graphical and statistical methods described in Section 7 of this guide.

6.1.2 At the time of installation, all of the graphical and numerical methods outlined in this guide may prove useful. On a routine basis, conducting any of the instrument stability checks and comparative evaluations that do not disrupt normal operations may prove useful. Control charts may be applied to all the performance measures that are gathered on a routine basis, including mean analysis value of reference material, RMSD, etc.

6.1.3 Whenever there is a major change to the operating parameters, the configuration, the calibration, the processes, or the hardware associated with the analyzer or the reference system, the user may wish to perform comparative tests. In

addition to comparative tests, standardization or static tests, or both, as recommended by the online analyzer manufacturer, may be helpful.

6.1.4 Changes in coal characteristics may also impact analyzer performance. Particle size, source of coal, mining techniques, and degree of preparation, which if changed from previous test periods and which are not in the analyzer calibration database, may affect analyzer precision and accuracy.

6.1.5 Additional changes which could merit performance testing include a change in the material or width of the conveyor, or coal flow rates (in the case of full-flow analyzers).

6.2 *Static Stability Measurements for Baseline Assessment*—A reference material may be used to provide a baseline assessment of static measurement precision. The reference material may be used to compare current mean and standard deviation values with mean and standard deviation values, previously collected in the same manner. The resulting comparative data may help determine whether any apparent decline in analyzer dynamic performance may be attributed to a change in the operating characteristics of the analyzer, in the absence of the influence of sampling, preparation, and analysis.

6.2.1 The results of this evaluation can indicate whether analyzer precision has significantly degraded or whether a bias may have occurred. If so, it may be possible to adjust the analyzer to restore initial performance. If the user wishes only to measure current static stability, any available coal may be used in the analysis zone of the analyzer. Note, however, that the actual standard deviation in static stability tests might be influenced by the composition or mass of the coal being examined or analyzer factors, such as the strength of the radioactive sources used by the analyzer or condition of analyzer electronic components.

6.2.2 It is essential that the length of the analysis period be defined, (for example, one minute, two minutes, or five minutes) and be constant in the static stability test. The standard deviation resulting from the static stability test decreases as the length of the analysis period increases.

### 6.3 Comparison of Analyzer System to Reference System Measurements:

6.3.1 Once an analyzer installation has been completed and calibration adjustments have been made, the analyzer owner may require acceptance testing. Also, the analyzer owner may decide to relocate the analyzer. In these cases, comparison tests—that is, to compare the analyzer system results to conventional sampling and analysis techniques—will provide the user calibration verification data and/or data that could be used for recalibration of the analyzer.

6.3.2 Since performance evaluations usually consist of 30 or more comparisons, with each of these lasting from 30 minutes to three hours, these comparisons may be conducted in a batch over several hours or days, or continuously throughout the operational life of the analyzer system.

6.3.3 If two independent conventional coal-sampling and laboratory analysis measurements can be made from each of a series of batches of coal interrogated by the analyzer, the Latent Variable Model (LVM) can be used to provide unbiased

estimates of the measurement precision of the analyzer and of the conventional sampling and analysis systems.

6.3.4 Any two series of measurements are independent if their measurement errors are uncorrelated. Correlation of measurement errors can be avoided and independence assured by use of a true random selection of physical increments or samples of material or by using different schemes and equipment for collection, preparation, and laboratory analysis of the samples, or both. A complete treatment of the subject of independence of measurements and the various means of assuring independence is beyond the scope of this guide.

6.4 Sampling Considerations:

6.4.1 Selection of Appropriate Sampling and Sample Preparation Methods—Decisions regarding sample collection should be governed by Practices D2234/D2234M and D7430 as appropriate. The method to be used for sample preparation should be determined before the beginning of increment collection. Sample preparation techniques should remain consistent (see Practice D2013). Before installation of the analyzer, consideration should be given to the ability to obtain representative samples for comparison to analyzer measurements and the regimen for sample handling and analysis. For the use of mechanical sampling systems, inspection (see Practice D7430, Part C) and bias testing (see Practice D7430, Part D) are good methods for evaluation of the system’s suitability for the test.

6.4.2 Selection of Reference Sample Point(s):

6.4.2.1 Comparative tests can be used to evaluate the performance of either the analyzer itself or the “analyzer system” (consisting of the analyzer and the sampling system that feeds it). The comparative evaluations of the analyzer system can be used to determine the ability of the analyzer system to measure the characteristics of the main coal stream. The selection of the reference system sample point(s) determines whether the comparative tests will assess the analyzer or the analyzer system.

6.4.2.2 The most direct and practical two-instrument test, when the analyzer is fed the secondary reject of a mechanical sampling system, uses the final save to compare directly to the analyzer. A manual or mechanical sample collected from the analyzer discharge may provide an independent sample, which may be used to assess the performance of the analyzer (see Figs. 1-3).

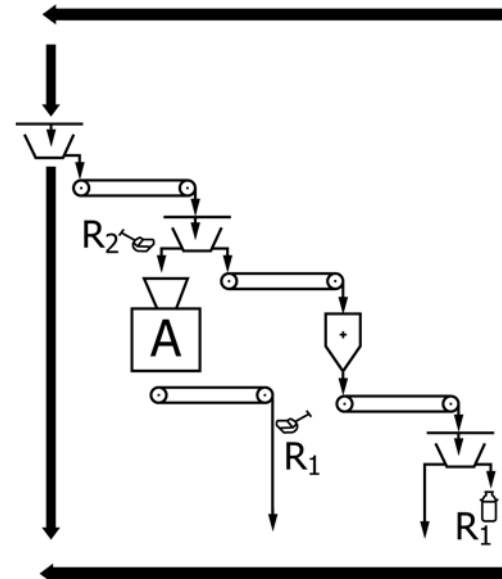


FIG. 2 Analyzer on Secondary Reject—Three-Stage Mechanical Sampling

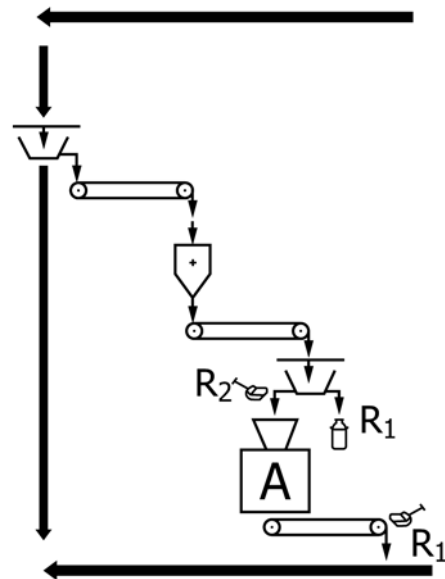


FIG. 3 Analyzer on Secondary Reject—Two-Stage Mechanical Sampling

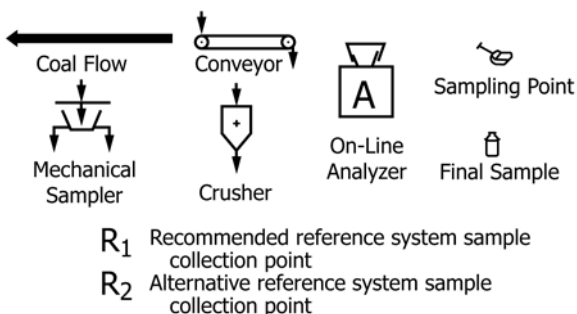


FIG. 1 Key to Schematics

6.4.2.3 There are instances when there is no save sample (see Fig. 4) associated with the system feeding the analyzer. In this case, it is possible to construct a test with several comparisons by collecting separate samples from the analyzer feed and discharge.

6.4.2.4 In some instances, the discharge of the analyzer may be fed to further stages of mechanical sampling. A single stage of sampling downstream of the analyzer is most common (see Fig. 5). In this case, the secondary save will provide a convenient comparison sample. A test of such a system that requires more than one set of comparative data might incorporate a series of samples (stopped belt or full stream cut) collected from the analyzer discharge or the secondary reject stream, as well. In the relatively rare circumstances in which

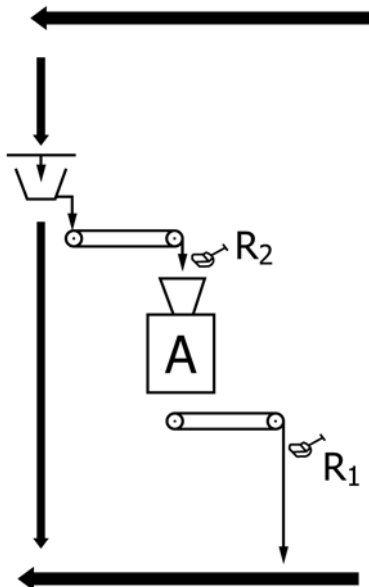


FIG. 4 Single-Stage Mechanical Sampling

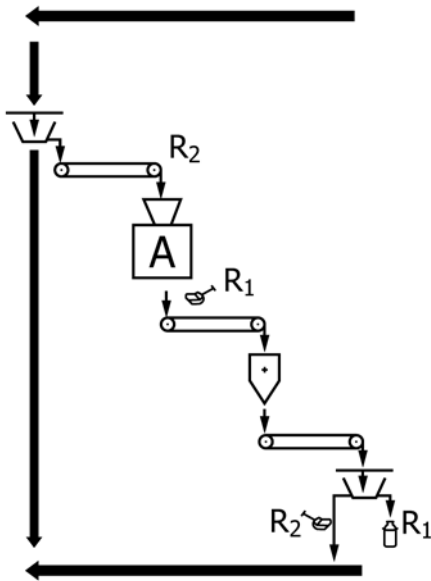


FIG. 5 Analyzer on Primary Save—Two-Stage Mechanical Sampling

the analyzer discharge feeds two additional stages of sampling, the tertiary save is recommended for two-instrument and bias testing, and a stopped belt or full stream cut of the secondary or tertiary rejects of the analyzer discharge may be used as a third instrument. Practical considerations of increment collection at the secondary reject should be balanced with consider-

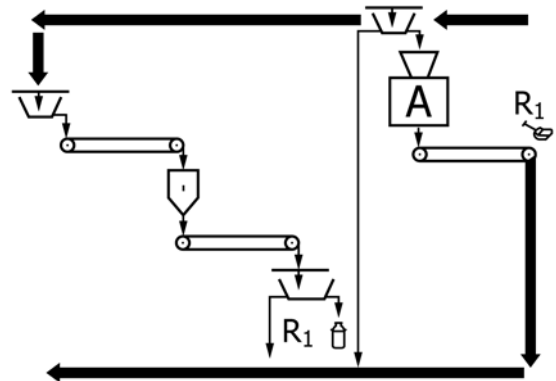


FIG. 6 Analyzer Distinct from Reference

ations of sampling variances introduced by crushing and tertiary sampling. For two instruments to be independent of each other, one or both instruments must interrogate the stream of interest without changing the characteristics of the stream. This may be true in some through-belt noncontacting configurations. In the case of flow-through analyzers that require a sample, the independence of systems is obtained in a case in which the primary coal stream is sampled by one instrument before being sampled by another instrument (see Fig. 6). In this case, the two systems may be evaluated by comparing the analyzer values to the final save of the mechanical system. The analyzer itself may be evaluated by comparison to samples collected at its discharge.

6.4.3 *Synchronization*—To evaluate the analyzer’s performance, the lot of coal interrogated by the analyzer must be the same lot of coal measured by the reference system. This evaluation is usually accomplished by sampling the coal stream after it passes the analyzer’s analysis zone except for sample stream analyzers. Care should be taken to ensure that transport times within the coal-handling system of interest, within mechanical sampling systems used to provide comparative measurements, and between sample collection and the analyzer be measured and accounted for in any comparisons. Comparisons in which intervening storage of the coal takes place between the analyzer and reference measurements should be avoided.

6.4.4 *Length of Comparison Period*—The length of time chosen for a comparison period should be commensurate with the period of time—that is, the lot time—during which the analyzer will normally produce a result of interest to the user.

6.4.5 *Number of Comparisons*—Although as few as 15 comparisons may be useful in conventional control charting, as many as 60 comparisons are recommended if one is to obtain useful confidence limits for the Latent Variable Model estimate of precision.

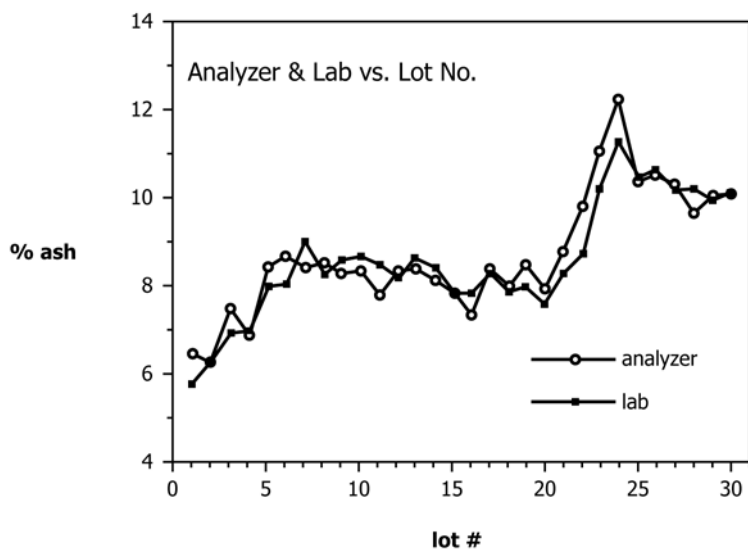


FIG. 7 Trend Plot Showing the Analyzer and Reference System Measurements Versus the Lot Number on the Same Graph

## 7. Evaluation of Data

7.1 *Introduction*—Once comparative data have been collected, it is recommended that the data be evaluated using both graphical and statistical, that is, numerical techniques. The sections that follow give recommendations for the construction and interpretation of various charts or graphs as well as the techniques used to compute and interpret various statistics.

### 7.2 Stability Evaluations:

#### 7.2.1 Graphical Techniques, see 7.3.2.

7.2.2 *Numerical Methods and Their Interpretation*—The two statistics most relevant to static stability evaluations are the mean and standard deviation of the static analyses. To calculate the mean and standard deviation of the static analyses, the individual analyzer static analysis data points (analysis by analysis) are needed. The sample mean is given by the formula:

$$\bar{x} = \left[ \sum_{i=1}^n x_i \right] / n \quad (1)$$

where:

$\bar{x}$  = the mean,

$x_i$  = the  $i$ th measurement, and

$n$  = the number of measurements.

The sample standard deviation  $s_d$  is given by:

$$s_d = \sqrt{\left[ \sum_{i=1}^n (x_i - \bar{x})^2 \right] / (n - 1)} \quad (2)$$

7.3 *Comparative Evaluations*—Comparison of analyzer results to those generated by conventional techniques should be performed and assessed using the techniques described below. These evaluations require a series of comparisons between analyzer measurements and reference system measurements. It is recommended that such comparisons be made routinely.

#### 7.3.1 Outliers:

7.3.1.1 At times, a measurement from an analyzer or reference system may appear to be an aberrant value. Statistical rules leading one to look for causes of outliers when they exist

are given in Practice E178. In applying the methods of Practice E178, the difference between the analyzer and reference value is computed for each pair.

7.3.1.2 Each difference is then divided by the sample standard deviation of the differences, with the absolute value of the quotient designated as  $U_i$  for the  $i$ th pair. It is recommended here that any pair ( $i$ ), whose  $U$  value is found using Table 1 of Practice E178 to exceed the table value for an appropriate upper significance level, that is, 99 %, be treated as an outlier. An investigation of the cause of the outlier should be undertaken, and it may be appropriate to exclude the outlier data from analyzer performance evaluation calculations.

7.3.2 *Graphical Methods*—Regardless of the numerical method(s) used to assess instrument performance, any comparison between analyzer values and reference system values should include graphical displays. At a minimum, a trend plot and a scatter plot should be constructed. Many conclusions can be drawn or hypotheses constructed about the quality of the calibration of the instrument, the synchronization of the data, and the precision of the analyzer or the reference system.

7.3.2.1 *Trend Plot*—Fig. 7 is an example of a trend plot, showing the analyzer and reference system measurements versus the lot number on the same graph. Provided that product variation is sufficiently large compared to the precision of each of these measurements, one expects to see “tracking.” That is, when the conventional measurement shows a decrease in ash, one expects the analyzer measurement to show a similar decrease. Fig. 7 above appears to be an example of good tracking. A trend plot of the difference between the analyzer and the reference sample result also may be used. Shown in Fig. 8 is a magnification of the Fig. 7 plot for lot numbers 10 to 20. Notice that, on this scale, the “tracking” is less convincing. Because the product variability is quite small during this time, this does not necessarily mean poor performance. Again, tracking is expected only when product variability is large compared to the measurement precision of each system.

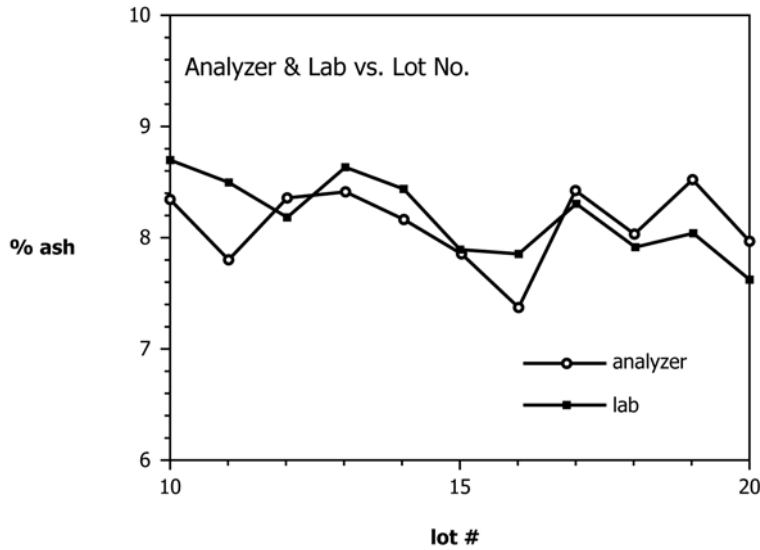


FIG. 8 Magnification of Fig. 7 Plot for Lot Numbers 10 to 20

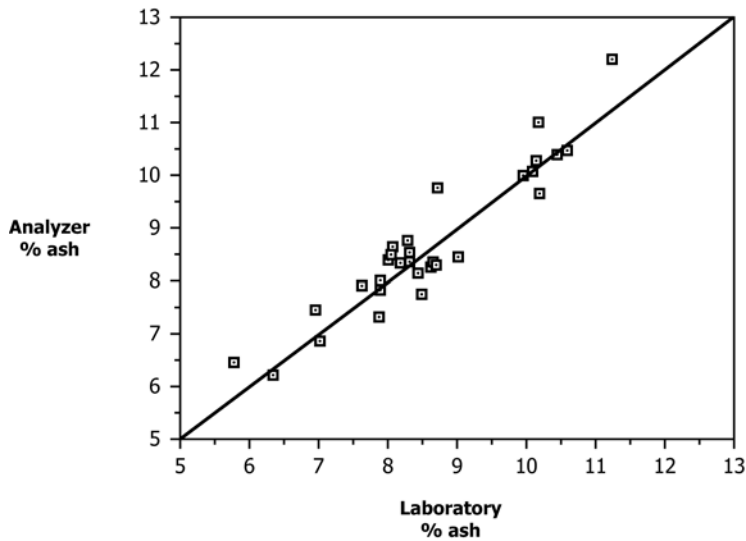


FIG. 9 Scatter Plot or “x-y Plot” of Analyzer Results Versus Reference System Measurements (the Data are the Same as Those from Fig. 7)

7.3.2.2 *Scatter Plot*—Shown in Fig. 9 is an example of a scatter plot or “x-y-plot” of analyzer results versus reference system measurements (the data are the same as those from Fig. 7). For an accurate and well-calibrated analyzer, one expects the points to fall along a line with zero intercept and a slope of one, as indicated in the graph. An improperly calibrated analyzer may show a good linear pattern but with a nonzero intercept or a non-unity slope in the parameter being measured,

while an imprecise analyzer will show a large scatter. Again, unless the product variability is reasonably large compared to instrument precision, comparisons may not be very conclusive.

NOTE 1—Fig. 9 shows a reasonable scatter around the 45° line (a line with zero intercept and a slope of one), with the exception of three potential outliers at the upper right. The numerical methods discussed in this guide would not by themselves indicate a problem of this sort.

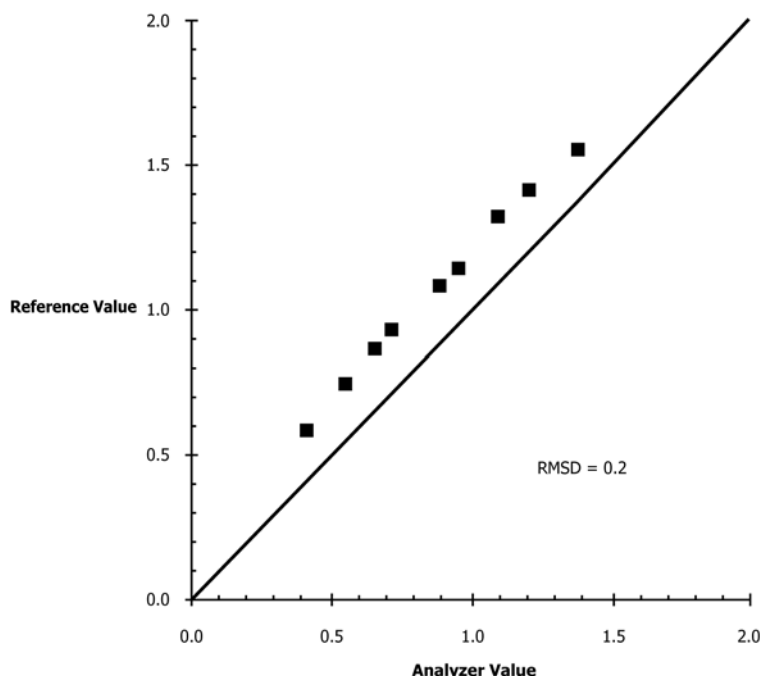


FIG. 10 Example of a Very Precise Instrument Whose RMSD is Large Because of Improper Calibration

7.3.3 Numerical Methods:

**RMSD Accuracy Statistic**

7.3.3.1 The root mean squared difference (RMSD), is as its name implies the square root of the mean of the differences between reference and analyzer values squared.

7.3.3.2 Calculations—By formula, RMSD is calculated as follows:

$$RMSD = \sqrt{\left[ \frac{\sum_{i=1}^n (x_i - y_i)^2}{n} \right]} \quad (3)$$

where:

- $n$  = the number of paired comparisons,
- $x_i$  = the analyzer value in the  $i$ th comparison, and
- $y_i$  = the reference value in the  $i$ th comparison.

7.3.3.3 Interpretation—Note that the RMSD statistic becomes larger (or worse) as any of the following occur: the bias increases, the precision of the instrument worsens, or the precision of the reference system worsens. A good (or small) RMSD statistic, therefore, usually implies a precise analyzer with small bias being evaluated by a fairly precise reference system. On the other hand, a small RMSD may arise strictly because the coal quality values have stayed within a narrow range during the evaluation period. A large RMSD does not necessarily mean the analyzer is performing poorly. The large RMSD may result from a poor calibration that leaves the instrument biased. See Fig. 10 as an example of a very precise instrument whose RMSD is large because of improper calibration. The proper response to this assessment of analyzer performance should be to adjust the calibration, not to reject the instrument as unworthy. Furthermore, a perfect instrument may produce a sizable RMSD if the ability of the reference system to equal the true value of the coal quality measurement

is poor. This situation could arise if the variances of sampling, preparation, or analysis in the reference system are significant.

**Latent Variable Model Estimate of System Precision**

7.3.3.4 After obtaining the analyzer’s measurement and two independent reference measurements for each of several batches of coal, the estimates of measurement precision can be calculated using the latent variable model method.

7.3.3.5 Number of Batches—The number of batches of coal used for comparison should be selected to achieve the desired level of precision for the estimates of precision as described in Rose.<sup>4</sup> Measurements from 30 to 60 batches of coal are typically used.

7.3.3.6 Calculations: Assume the latent variable model:

$$X_i = \beta_{01} + \beta_{11}v_i + e_{ix} \quad (4)$$

$$Y_i = \beta_{02} + \beta_{12}v_i + e_{iy} \quad (5)$$

$$Z_i = v_i + e_{iz} \quad (6)$$

where:

- $X_i$  = the measured value for observation  $i$  from the analyzer (assumed to be possibly biased);
- $Y_i$  = the measurement for observation  $i$  from the first reference system (assumed to be possibly biased);
- $Z_i$  = the reference measurement for observation  $i$  from the second reference system (assumed to be unbiased);
- $V_i$  = the latent coal quality for observation  $i$ ;
- $\beta_{01}$  and  $\beta_{11}$  = the coefficients for bias in the analyzer measurement;

<sup>4</sup> Rose, C. D., “Calibration and Testing of On-Line Coal Analyzers,” *Journal of ASTM International*, Vol 1, No. 4, April 2004.

$\beta_{02}$  and  $\beta_{12}$  = the coefficients for bias in the measurement from reference system 1; and

$e_{ix}$ ,  $e_{iy}$ ,  $e_{iz}$  = the measurement errors associated with the analyzer and two reference systems.

The same estimates of covariance can be calculated as follows:

$$m_{XY} = \frac{1}{(n-1)} \sum_{i=1}^n (Y_i - \bar{Y})(X_i - \bar{X}) \quad (7)$$

$$m_{XZ} = \frac{1}{(n-1)} \sum_{i=1}^n (X_i - \bar{X})(Z_i - \bar{Z}) \quad (8)$$

$$m_{YZ} = \frac{1}{(n-1)} \sum_{i=1}^n (Y_i - \bar{Y})(Z_i - \bar{Z}) \quad (9)$$

$$m_{XX} = \frac{1}{(n-1)} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (10)$$

$$m_{YY} = \frac{1}{(n-1)} \sum_{i=1}^n (Y_i - \bar{Y})^2 \quad (11)$$

$$m_{ZZ} = \frac{1}{(n-1)} \sum_{i=1}^n (Z_i - \bar{Z})^2 \quad (12)$$

Estimates of coefficients for the bias models can then be calculated as follows:

$$\hat{\beta}_{11} = \frac{m_{XY}}{m_{YZ}} \quad \hat{\beta}_{01} = \bar{X} - \hat{\beta}_{11}\bar{Z} \quad (13)$$

$$\hat{\beta}_{12} = \frac{m_{XZ}}{m_{ZZ}} \quad \hat{\beta}_{02} = \bar{Y} - \hat{\beta}_{12}\bar{Z} \quad (14)$$

The user is referred to Rose<sup>4</sup> for calculation of the standard error of the estimates for the fit coefficients, which may be used to test for statistical significance.

The latent variable model provides the following estimates of variance:

$$V_c = \frac{m_{XZ}m_{YZ}}{m_{XY}} \quad (15)$$

$$V_z = m_{ZZ} - V_c \quad (16)$$

$$V_x = m_{XX} - \hat{\beta}_{11}^2 V_c \quad (17)$$

$$V_y = m_{YY} - \hat{\beta}_{12}^2 V_c \quad (18)$$

where:

$V_c$  = the estimate of the variance of the coal, batch by batch;

$V_x$  = the estimate of the variance of the errors of batch measurement by the analyzer;

$V_y$  = the estimate of the variance of the errors of batch measurement by the reference 1 method; and

$V_z$  = the estimate of the variance of the errors of batch measurement by the reference 2 method.

The estimate of the index of measurement precision for each of the three systems is then given by:

$$s_z = \sqrt{V_z} \quad (19)$$

$$s_y = \sqrt{V_y}$$

$$s_x = \sqrt{V_x}$$

7.3.3.7 One may wish to determine whether there is sufficient evidence to conclude that the calculated indices of precision are different from a specific value of interest at a particular level of confidence. If this is the case, the user is again referred to Rose.<sup>4</sup>

7.3.3.8 *Interpretation*—The LVM method is used to estimate the precision of three independent measurement systems, given the assumption that there may be a bias between any two reference systems that can be described as a linear function of the latent coal quality. The method assumes one of the reference systems to be unbiased and provides a method to estimate the coefficients for the bias equation for the other measurement systems.

In the limit that the first order fit coefficients ( $\hat{\beta}_{11}$  and  $\hat{\beta}_{12}$ ) can be assumed to be unity, the expected estimates of precision from the latent variable method are identical to those from the Grubbs' estimators of precision.<sup>5</sup>

Use of the LVM method will allow for the simultaneous estimate of analyzer precision as well as an estimate of the optimal calibration coefficients for the instrument. The method provides an estimate for the precision of the analyzer at the time the data to be used in the evaluation was gathered. It does not presume to predict precision in the future.

### Standard Deviation of the Differences

7.3.3.9 The sample standard deviation of the differences between the analyzer and reference measurements is an indicator of the combined indexes of precision of both the measurement systems.

7.3.3.10 *Calculations*—The sample standard deviation of the differences  $SD_d$  is given below. Note that with the reasonable assumption that the analyzer measurement errors are uncorrelated with reference measurement errors.

where  $d_i = x_i - y_i$  and

$$\bar{d} = \left[ \sum_{i=1}^n d_i \right] / n \quad (20)$$

7.3.3.11 *Interpretation*—If the variance of reference measurement errors is small compared to the variance of analyzer measurement errors, and if the number of data sets is 30 or more, the sample standard deviation of the differences may reasonably approximate an upper bound for the analyzer precision.

7.3.3.12 *Test for Bias*—Bias of interest in the application of online analyzers is the bias between analyzer measurements and reference measurements made by conventional coal sampling and analysis. The statistical techniques in the Practice **D6518** may be used to test for bias between the measurements made by an online analyzer and reference measurements made by conventional coal sampling and analysis.

## 8. Keywords

8.1 comparative measurement; online coal analyzer

<sup>5</sup> Grubbs' estimators were described in ASTM D6543 between 2000 and 2014a.



ANNEXES

(Mandatory Information)

A1. EXAMPLE OF THREE-MEASUREMENT TEST

TABLE A1.1 Example of Three-Measurement Test for Precision Estimates

Batch No.	(X)	(Y)	(Z)	X- $\bar{X}$	Y- $\bar{Y}$	Z- $\bar{Z}$
1	18.48	20.51	17.69	2.32	1.39	1.50
2	18.33	22.34	19.10	2.16	3.23	2.92
3	10.97	13.75	11.56	-5.20	-5.37	-4.63
4	18.57	23.17	20.17	2.40	4.06	3.98
5	14.84	20.25	16.06	-1.33	1.14	-0.13
6	11.95	13.08	11.05	-4.22	-6.03	-5.13
7	13.11	14.06	12.75	-3.06	-5.05	-3.43
8	15.12	17.67	14.70	-1.05	-1.44	-1.49
9	20.08	22.17	19.40	3.91	3.06	3.22
10	18.61	24.64	18.93	2.44	5.53	2.75
11	11.75	13.28	11.91	-4.42	-5.83	-4.27
12	19.40	23.59	19.35	3.23	4.47	3.17
13	19.55	22.77	19.61	3.38	3.66	3.43
14	15.07	18.10	14.64	-1.10	-1.01	-1.55
15	17.47	20.57	18.12	1.30	1.46	1.94
16	11.63	12.55	11.18	-4.54	-6.56	-5.01
17	14.22	15.75	14.41	-1.94	-3.36	-1.77
18	19.31	22.77	19.45	3.14	3.66	3.27
19	18.37	20.89	18.61	2.20	1.77	2.42
20	19.97	22.80	19.52	3.80	3.68	3.33
21	16.10	20.43	15.70	-0.07	1.32	-0.48
22	10.68	12.51	10.02	-5.49	-6.60	-6.16
23	17.76	21.80	19.03	1.59	2.69	2.85
24	18.62	23.75	18.91	2.45	4.64	2.73
25	16.74	19.14	17.17	0.58	0.03	0.99
26	18.26	21.10	17.63	2.10	1.99	1.44
27	16.97	21.03	18.01	0.81	1.92	1.82
28	14.21	16.24	13.14	-1.96	-2.87	-3.05
29	16.41	19.59	16.48	0.25	0.48	0.29
30	12.49	13.05	11.23	-3.68	-6.06	-4.96
Mean	16.169	19.112	16.184			

A1.1 Table A1.1 illustrates hypothetical data from a three-measurement test. Keep in mind that although only  $n = 30$  sets of data are used in the illustration, useful tests may require a number of data sets with  $n$  equal to or greater than 60.

A1.2 For the data of Table A1.1 the covariance matrix can be calculated as:

$$\begin{bmatrix} m_{XX} & m_{XY} & m_{XZ} \\ - & m_{YY} & m_{YZ} \\ - & - & m_{ZZ} \end{bmatrix} = \begin{bmatrix} 8.745 & 11.074 & 9.325 \\ - & 15.335 & 12.304 \\ - & - & 10.427 \end{bmatrix} \quad (A1.1)$$

And the estimates of the coefficients for the bias model are then calculated:

$$\begin{aligned} \hat{\beta}_{11} &= \frac{m_{XY}}{m_{YZ}} = \frac{11.074}{12.304} = 0.900 \\ \hat{\beta}_{01} &= \bar{X} - \hat{\beta}_{11}\bar{Z} = 16.169 - 0.900 \times 16.184 = 1.602 \\ \hat{\beta}_{12} &= \frac{m_{XZ}}{m_{ZZ}} = \frac{9.325}{10.427} = 1.188 \\ \hat{\beta}_{02} &= \bar{Y} - \hat{\beta}_{12}\bar{Z} = 19.112 - 1.188 \times 16.184 = -0.108 \end{aligned} \quad (A1.2)$$

The latent variable model provides the following estimates of variance:

$$\begin{aligned} V_c &= \frac{m_{XZ}m_{YZ}}{m_{XY}} = \frac{9.325 \times 12.304}{11.074} = 10.361 \\ V_x &= m_{XX} - \hat{\beta}_{11}^2 V_c = 8.745 - 0.900^2 \times 10.361 = 0.352 \\ V_y &= m_{YY} - \hat{\beta}_{12}^2 V_c = 15.335 - 1.188^2 \times 10.361 = 0.723 \\ V_z &= m_{ZZ} - V_c = 10.427 - 10.361 = 0.066 \end{aligned} \quad (A1.3)$$

And the estimate of the index of measurement precision for each of the three systems is then:

$$\begin{aligned} s_x &= \sqrt{V_x} = \sqrt{0.352} = 0.59 \\ s_y &= \sqrt{V_y} = \sqrt{0.723} = 0.85 \\ s_z &= \sqrt{V_z} = \sqrt{0.066} = 0.26 \end{aligned} \quad (A1.4)$$

## A2. EXAMPLE CALCULATIONS: RMSD AND $S_d$ CALCULATIONS

A2.1 An example is shown in **Table A2.1**.

**TABLE A2.1 Ash Measurement (Wt., %)**

Comparison Number	Analyzer	Reference <sup>A</sup>	Difference	(Difference) <sup>2</sup>
1	12.75	12.14	0.61	0.372
2	11.35	11.12	0.23	0.053
3	15.92	15.59	0.33	0.109
4	10.48	10.90	(0.42)	0.176
5	12.32	12.20	0.12	0.014
6	13.14	12.83	0.31	0.096
7	15.26	15.77	(0.51)	0.261
8	13.40	14.03	(0.63)	0.397
Total mean			0.005	1.478

<sup>A</sup>May be the mean of two reference values:

$$\begin{aligned}
 \text{RMSD} &= (1.478/8)^{1/2} \\
 &= (0.185)^{1/2} \\
 &= 0.430 \\
 S_d &= 0.460
 \end{aligned}$$

*ASTM International takes no position respecting the validity of any patent rights asserted in connection with any item mentioned in this standard. Users of this standard are expressly advised that determination of the validity of any such patent rights, and the risk of infringement of such rights, are entirely their own responsibility.*

*This standard is subject to revision at any time by the responsible technical committee and must be reviewed every five years and if not revised, either reapproved or withdrawn. Your comments are invited either for revision of this standard or for additional standards and should be addressed to ASTM International Headquarters. Your comments will receive careful consideration at a meeting of the responsible technical committee, which you may attend. If you feel that your comments have not received a fair hearing you should make your views known to the ASTM Committee on Standards, at the address shown below.*

*This standard is copyrighted by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959, United States. Individual reprints (single or multiple copies) of this standard may be obtained by contacting ASTM at the above address or at 610-832-9585 (phone), 610-832-9555 (fax), or service@astm.org (e-mail); or through the ASTM website (www.astm.org). Permission rights to photocopy the standard may also be secured from the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, Tel: (978) 646-2600; http://www.copyright.com/*