Designation: D 6122 – 99

An American National Standard

Standard Practice for Validation of Multivariate Process Infrared Spectrophotometers[1]

This standard is issued under the fixed designation D 6122; 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 covers requirements for the validation of measurements made by on-line, process near- or mid-infrared analyzers, or both, used in the calculation of physical, chemical, or quality parameters of liquid petroleum products. The parameters are calculated from spectroscopic data using multivariate modeling methods. The requirements include verification of adequate instrument performance, verification of the applicability of the calibration model to the spectrum of the sample under test, and verification of equivalence between the result calculated from the infrared measurements and the result produced by the primary method used for the development of the calibration model.

1.2 This practice does not cover procedures for establishing the calibration model used by the analyzer. Calibration procedures are covered in Practices E 1655 and references therein.

1.3 This practice is intended as a review for experienced persons. For novices, this practice will serve as an overview of techniques used to verify instrument performance, to verify model applicability to the spectrum of the sample under test, and to verify equivalence between the parameters calculated from the infrared measurement and the results of the primary method measurement.

1.4 This practice teaches and recommends appropriate statistical tools, outlier detection methods, for determining whether the spectrum of the sample under test is a member of the population of spectra used for the analyzer calibration. The statistical tools are used to determine if the infrared measurement results in a valid property or parameter estimate.

1.5 The outlier detection methods do not define criteria to determine whether the sample, or the instrument is the cause of an outlier measurement. Thus, the operator who is measuring samples on a routine basis will find criteria to determine that a spectral measurement lies outside the calibration, but will not have specific information on the cause of the outlier. This practice does suggest methods by which instrument performance tests can be used to indicate if the outlier methods are responding to changes in the instrument response.

1.6 This practice is not intended as a quantitative performance standard for the comparison of analyzers of different design.

1.7 Although this practice deals primarily with validation of on-line, process infrared analyzers, the procedures and statistical tests described herein are also applicable to at-line and laboratory infrared analyzers which employ multivariate models.

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 consult and establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:*

D 1265 Practice for Sampling Liquefied Petroleum Gases[2]

D 3764 Practice for Validation of Process Stream Analyzers[3]

D 4057 Practice for Manual Sampling of Petroleum and Petroleum Products[3]

D 4177 Practice for Automatic Sampling of Petroleum and Petroleum Products[3]

D 6299 Practice for Applying Statistical Quality Assurance Techniques to Evaluate Analytical Measurement System Performance[4]

E 131 Terminology Relating to Molecular Spectroscopy[5]

E 275 Practice for Describing and Measuring Performance of Ultraviolet, Visible, and Near Infrared Spectrophotometers[5]

E 932 Practice for Describing and Measuring Performance of Dispersive Infrared Spectrophotometers[5]

E 1421 Practice for Describing and Measuring Performance of Fourier Transform Infrared (FT-IR) Spectrometers: Level Zero and Level One Tests[5]

E 1655 Practices for Infrared, Multivariate, Quantitative Analysis[5]

E 1866 Guide for Establishing Spectrophotometer Performance Tests[5]

E 1944 Practice for Describing and Measuring Performance

[1] This practice is under the jurisdiction of ASTM Committee D-2 on Petroleum Products and Lubricants and is the direct responsibility of Subcommittee D02.25 on Validation of Process Analyzers and Statistical Quality Assurance of Measurement Processes for Petroleum and Petroleum Products.

Current edition approved Dec. 10, 1999. Published December 1999. Originally published as D 6122–97. Last previous edition D 6122–97.

[2] *Annual Book of ASTM Standards*, Vol 05.01.

[3] *Annual Book of ASTM Standards*, Vol 05.02.

[4] *Annual Book of ASTM Standards*, Vol 05.04.

[5] *Annual Book of ASTM Standards*, Vol 03.06.

of Fourier Transform Near-Infrared (FT-NIR) Spectrometers: Level Zero and Level One Tests[5]

3. Terminology

3.1 *Definitions:*

3.2 For definitions of terms and symbols relating to IR spectroscopy, refer to Terminology E 131.

3.3 For definitions of terms and symbols relating to multivariate calibration, refer to Practices E 1655.

3.4 *Definitions of Terms Specific to This Standard:*

3.4.1 *action limit*, *n*—the limiting value from an instrument performance test, beyond which the analyzer is expected to produce potentially invalid results.

3.4.2 *analyzer* , *n*—all piping, hardware, computer, software, instrumentation and calibration model required to automatically perform analysis of a process or product stream.

3.4.3 *analyzer calibration*, *n*—see *multivariate calibration*.

3.4.4 *analyzer intermediate precision*, *n*— a statistical measure of the expected long-term variability of analyzer results for samples whose spectra are neither outliers, nor nearest neighbor inliers.

3.4.5 *analyzer model*, *n*—see *multivariate model*.

3.4.6 *analyzer repeatability*, *n*—a statistical measure of the expected short-term variability of results produced by the analyzer for samples whose spectra are neither outliers nor nearest neighbor inliers.

3.4.7 *analyzer result*, *n*—the numerical estimate of a physical, chemical, or quality parameter produced by applying the calibration model to the spectral data collected by the analyzer.

3.4.8 *analyzer validation test,*, *n*—see *validation test.*

3.4.9 *calibration transfer*, *n*— a method of applying a multivariate calibration developed on one analyzer to a different analyzer by mathematically modifying the calibration model or by instrument standardization.

3.4.10 *check sample*, *n*—a single, pure liquid hydrocarbon compound, or a known, reproducible mixture of liquid hydrocarbon compounds whose spectrum is constant over time such that it can be used in a performance test.

3.4.11 *control limits*, *n*—limits on a control chart which are used as criteria for signaling the need for action, or for judging whether a set of data does or does not indicate a state of statistical control. **E 456**

3.4.12 *exponentially weighted moving average control chart*, *n*—a control chart based on the exponentially weighted average of individual observations from a system; the observations may be the differences between the analyzer result, and the result from the primary method.

3.4.13 *individual observation control chart*, *n*—a control chart of individual observations from a system; the observations may be the differences between the analyzer result and the result from the primary method.

3.4.14 *inlier*, *n*—see *nearest neighbor distance inlier*.

3.4.15 *inlier detection methods*, *n*—statistical tests which are conducted to determine if a spectrum resides within a region of the multivariate calibration space which is sparsely populated.

3.4.16 *in-line probe*, *n*—a spectrophotometer cell installed in a process pipe or slip stream loop and connected to the analyzer by optical fibers.

3.4.17 *instrument*, *n*—spectrophotometer, associated electronics and computer, spectrophotometer cell and, if utilized, transfer optics.

3.4.18 *instrument standardization*, *n*—a procedure for standardizing the response of multiple instruments such that a common multivariate model is applicable for measurements conducted by these instruments, the standardization being accomplished by way of adjustment of the spectrophotometer hardware or by way of mathematical treatment of the collected spectra.

3.4.19 *line sample*, *n*—a process or product sample which is withdrawn from a sample port in accordance with Practices D 1265, D 4057, or D 4177, whichever is applicable, during a period when the material flowing through the analyzer is of uniform quality and the analyzer result is essentially constant.

3.4.20 *moving range of two control chart*, *n*— a control chart that monitors the change in the absolute value of the difference between two successive differences of the analyzer result minus the result from the primary method.

3.4.21 *multivariate calibration*, *n*—an analyzer calibration that relates the spectrum at multiple wavelengths or frequencies to the physical, chemical, or quality parameters.

3.4.22 *multivariate model*, *n*—a multivariate, mathematical rule or formula used to calculate physical, chemical, or quality parameters from the measured infrared spectrum.

3.4.23 *nearest neighbor distance inlier*, *n*— a spectrum residing within a gap in the multivariate calibration space, the result for which is subject to possible interpolation error.

3.4.24 *optical background*, *n*—the spectrum of radiation incident on a sample under test, typically obtained by measuring the radiation transmitted through the spectrophotometer cell when no sample is present, or when an optically thin or nonabsorbing liquid is present.

3.4.25 *optical reference filter*, *n*—an optical filter or other device which can be inserted into the optical path in the spectrophotometer or probe producing an absorption spectrum which is known to be constant over time, such that it can be used in place of a check or test sample in a performance test.

3.4.26 *outlier detection limits*, *n*—the limiting value for application of an outlier detection method to a spectrum, beyond which the spectrum represents an extrapolation of the calibration model.

3.4.27 *outlier detection methods*, *n*—statistical tests which are conducted to determine if the analysis of a spectrum using a multivariate model represents an interpolation of the model.

3.4.28 *outlier spectrum*, *n*—a spectrum whose analysis by a multivariate model represents an extrapolation of the model.

3.4.29 *performance test*, *n*—a test that verifies that the performance of the instrument is consistent with historical data and adequate to produce valid results.

3.4.30 *physical correction*, *n*— a type of pos processing where the correction made to the numerical value produced by the multivariate model is based on a separate physical measurement of, for example, sample density, sample path length, or particulate scattering.

3.4.31 *post-processing*, *v*—performing a mathematical operation on an intermediate analyzer result to produce the final

result, including correcting for temperature effects, adding a mean property value of the analyzer calibration, and converting into appropriate units for reporting purposes.

3.4.32 *pre-processing*, *v*—performing mathematical operations on raw spectral data prior to multivariate analysis or model development, such as selecting wave length regions, correcting for baseline, smoothing, mean centering, and assigning weights to certain spectral positions.

3.4.33 *primary method*, *n*—the analytical procedure used to generate the reference values against which the analyzer is both calibrated and validated; Practices E 1655 uses the term reference method in place of the term primary method.

3.4.34 *process analyzer system*, *n*—see *analyzer*.

3.4.35 *process analyzer validation samples*, *n*—see *validation samples*.

3.4.36 *spectrophotometer cell*, *n*— an apparatus which allows a liquid hydrocarbon to flow between two optical surfaces which are separated by a fixed distance, the sample pathlength, while simultaneously allowing light to pass through the liquid.

3.4.37 *test sample*, *n*—a process or product sample, or a mixture of process or product samples, which has a constant spectrum for a finite time period, and which can be used in a performance test; test samples and their spectra are generally not reproducible in the long term.

3.4.38 *transfer optics*, *n*—a device which allows movement of light from the spectrophotometer to a remote spectrophotometer cell and back to the spectrophotometer; transfer optics include optical fibers or other optical light pipes.

3.4.39 *validation samples*, *n*—samples that are used to compare the analyzer results to the primary method results through the use of control charts and statistical tests; validation samples used in the initial validation may be line and test samples, whereas validation samples used in the periodic validation are line samples.

3.4.40 *validated result*, *n*—a result produced by the analyzer for a sample whose spectrum is neither an outlier nor a nearest neighbor inlier that is equivalent, within control limits to the result expected from the primary method, so that the result can be used instead of the direct measurement of the sample by the primary method.

3.4.41 *validation test*, *n*—a test performed on a validation sample that demonstrates that the result produced by the analyzer and the result produced by the primary method are equivalent to within control limits.

4. Summary of Practice

4.1 This section describes, in summary form, the steps involved in the validation of an infrared analyzer over the long term. Before this practice may be undertaken, certain preconditions shall be satisfied. The preconditions are described in Section 7. This practice consists of four major procedures.

4.2 Each time a spectrum of a process sample is collected, statistical tests are performed to verify that the multivariate model is applicable to the spectrum. Only spectra whose analysis represents interpolation of the multivariate modeland which are sufficiently close to spectra in the calibration may be used in the analyzer validation.

4.3 When the analyzer is initially installed, or after major maintenance is concluded, performance tests are conducted to verify that the instrument is functioning properly. The intent of these tests is to provide a rapid indication of the state of the instrument. These tests are necessary but not sufficient to demonstrate valid analyzer results.

4.4 After the initial performance test is successfully completed, an initial validation test is conducted to verify that the results produced by the analyzer are in statistical agreement with results for the primary method. Once this initial validation is completed, the analyzer results are considered valid for samples whose spectra are neither outliers or nearest neighbor inliers.

4.5 During routine operation of the analyzer, validation tests are conducted on a regular, periodic basis to demonstrate that the analyzer results remain in statistical agreement with results for the primary method. Between validation tests, performance tests are conducted to verify that the instrument is performing in a consistent fashion.

5. Significance and Use

5.1 The primary purpose of this practice is to permit the user to validate numerical values produced by a multivariate, infrared or near-infrared, on-line, process analyzer calibrated to measure a specific chemical concentration, chemical property, or physical property. *The validated analyzer results are expected to be equivalent, over diverse samples whose spectra are neither outliers or nearest neighbor inliers, to those produced by the primary method to within control limits established by control charts for the prespecified statistical confidence level.*

5.2 Procedures are described for verifying that the instrument, the model, and the analyzer system are stable and properly operating.

5.3 A multivariate analyzer system inherently utilizes a multivariate calibration model. In practice the model both implicitly and explicitly spans some subset of the population of all possible samples that could be in the complete multivariate sample space. The model is applicable only to samples that fall within the subset population used in the model construction. A sample measurement cannot be validated unless applicability is established. Applicability cannot be assumed.

5.3.1 Outlier detection methods are used to demonstrate applicability of the calibration model for the analysis of the process sample spectrum. The outlier detection limits are based on historical as well as theoretical criteria. The outlier detection methods are used to establish whether the results obtained by an analyzer are potentially valid. The validation procedures are based on mathematical test criteria that indicate whether the process sample spectrum is within the range spanned by the analyzer system calibration model. If the sample spectrum is an outlier, the analyzer result is invalid. If the sample spectrum is not an outlier, then the analyzer result is valid providing that all other requirements for validity are met. Additional, optional tests may be performed to determine if the process sample spectrum falls in a sparsely populated region of the multivariate space covered by the calibration set, too far from neighboring calibration spectra toensure good interpolation. For example, such nearest neighbor tests are recommended if the calibration sample spectra are highly clustered.

5.3.2 This practice does not define mathematical criteria to

determine from a spectroscopic measurement of a sample whether the sample, the model, or the instrument is the cause of an outlier measurement. Thus the operator who is measuring samples on a routine basis will find criteria in the outlier detection method to determine whether a sample measurement lies within the expected calibration space, but will not have specific information as to the cause of the outlier without additional testing.

6. Apparatus and Considerations for Quantitative On-Line Process IR Measurements

6.1 *Infrared or Near-Infrared Spectrophotometer:*

6.1.1 The analyzer covered by this practice is based on an infrared spectrophotometer, double-beam or single-beam, suitable for recording accurate measurements in the near-infrared (780 to 2500 nm, 12820.5 to 4000 cm^{-1}) or mid-infrared (4000–400 cm^{-1}) regions, or both. The spectral range measured by the analyzer shall be the same as that of the instrument used in collecting the spectral data upon which the multivariate calibration model is based. Complete descriptions of the instrumentation and procedures that are required for quantitative on-line process IR measurements are beyond the scope of this practice. Some general guidelines are given in Annex A1. (**Warning**—There are inherent dangers associated with the use of electrical instrumentation, on-line processes, and hydrocarbon materials. The users of this practice should have a practical knowledge of these hazards and employ appropriate safeguards.)

6.1.2 *In developing spectroscopic methods, it is the responsibility of the user to describe the instrumentation and the performance required to achieve the desired repeatability, reproducibility, and accuracy for the application.*

6.2 *Process Analyzer System:*

6.2.1 The process analyzer system typically includes the spectrophotometer, transfer optics, the hardware for sample handling, the hardware for introduction of reference standards and solvents, the computer for controlling the spectrophotometer and calculating results, and the multivariate model. The system configuration should be compatible with the mid-infrared or near-infrared IR measurement and this practice.

6.3 *Collection of Line Samples:*

6.3.1 Withdraw line samples in accordance with accepted sampling methods as given by Practices D 1265, D 4057, or D 4177, whichever is applicable. Flush the entire sample loop with the process stream sample prior to withdrawal of the line sample.

6.3.2 The intent of this practice is to collect samples that correspond directly to the spectra being collected by the analyzer. Collect the sample at a port close to the optical probe and at a time correlated with the collection of the sample spectrum. This practice requires that parameters that can impact the result also be recorded at the time of sample collection and the effect of these parameters be properly accounted for when comparing the results with the primary method result. For a more detailed discussion of the various lag times that can influence the correspondence between the analyzer measurement and collection of line samples, see Practice D 3764.

6.3.3 Sample storage for extended time periods is not recommended if there is a likelihood that samples degrade with time. Chemical changes occurring during storage will cause changes in the spectrum, as well as changes in the property or quality parameter measured by the primary method.

6.3.4 If possible, at the time of line sample withdrawal, collect sufficient quantity of sample material to allow for multiple measurements of the property or quality parameter by the primary method, should such measurements be required.

7. Preconditions

7.1 Certain preconditions shall be met before this practice can be applied.

7.1.1 Install the analyzer in accordance with manufacturer's instructions.

7.1.2 Develop and validate the multivariate calibration model used on the process analyzer using methods described in Practices E 1655. If a calibration transfer method is used to transfer the model from one analyzer to another, verify the transferred model as described in Practices E 1655.

7.1.3 A quality assurance program for the primary method is required in order to determine the usability of values generated by the primary method in the validation of analyzer performance using this practice (see Section 8).

8. Reference Values and the Quality Assurance Program for the Primary Method

8.1 The property reference value against which analyzer results are compared during validation is established by applying the primary measurement method which was used in the model development to line samples representing the process stream.

8.2 A quality assurance program for the primary method is required for values generated by this method to be used in analyzer validation.

8.2.1 Carefully check the laboratory apparatus used for primary method measurement before these tests are performed to ensure compliance with the requirements of the primary test method.

8.2.2 Test control materials of known composition and quality on a regularly scheduled basis. Plot the primary method results on control charts to ensure the long-term performance of the primary test. Individual values, exponentially weighted moving average, and moving range of two control charts are all recommended for charting the performance of the primary method. Calculate the values for these control charts using equations given in Sections 12 and 13. Plot the differences between the primary method result, and the expected value for the standard sample. Determine the historical precision of the primary method from these regular tests, and compare it to published values for the method to determine if the test is within expected limits. Compare the historical precision to the analyzer precision using statistical tests.

9. Procedure

9.1 A flowchart for the steps involved in this practice is shown in Fig. 1 and Fig. 2.

9.2 *Initial Performance Tests:*

9.2.1 After the multivariate process analyzer has been installed (or reinstalled following major maintenance), check

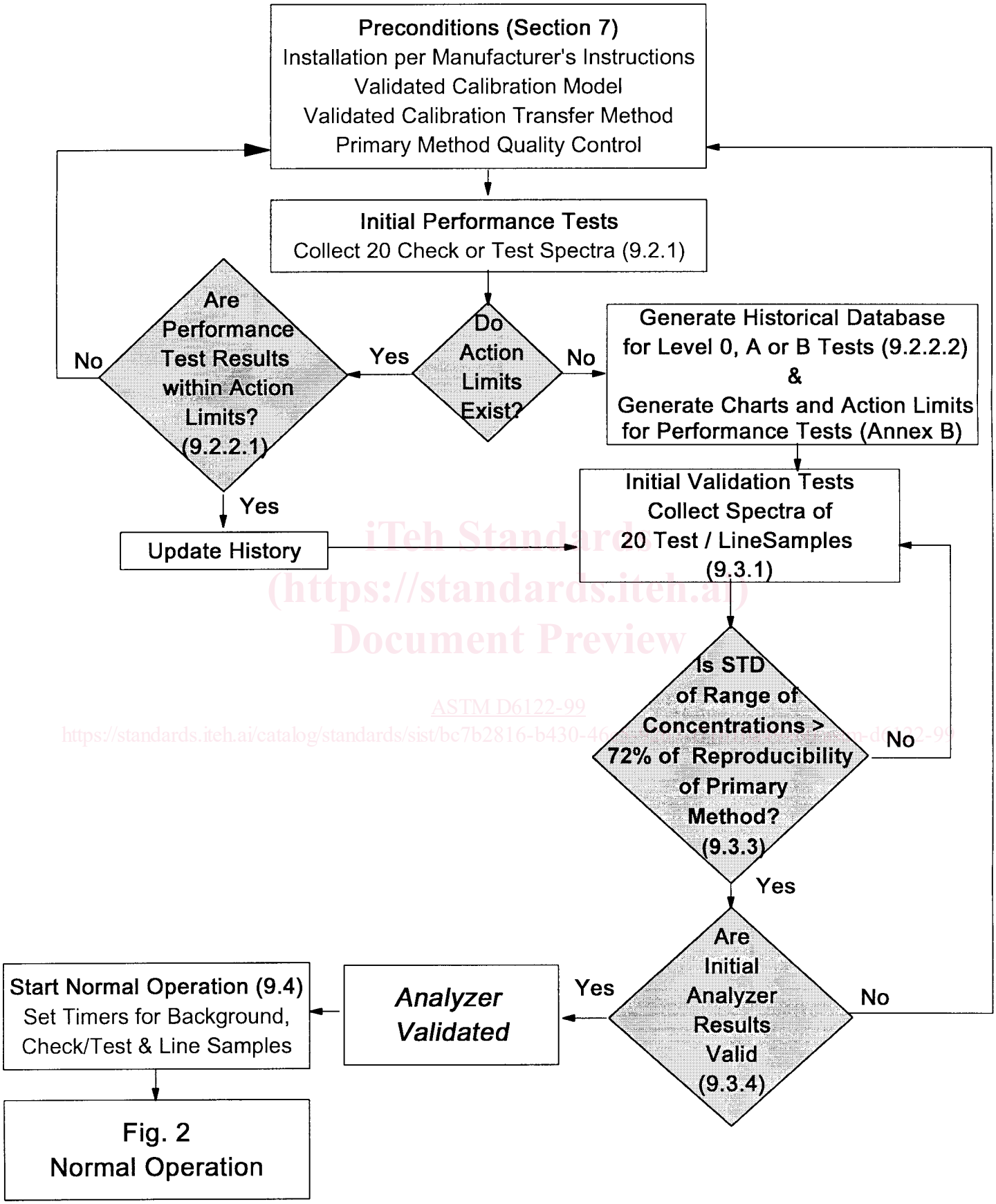

Fig. 2
Normal Operation

FIG. 1 Flowchart of Process Analyzer Validation Practice Initial Startup and Restart after Maintenance

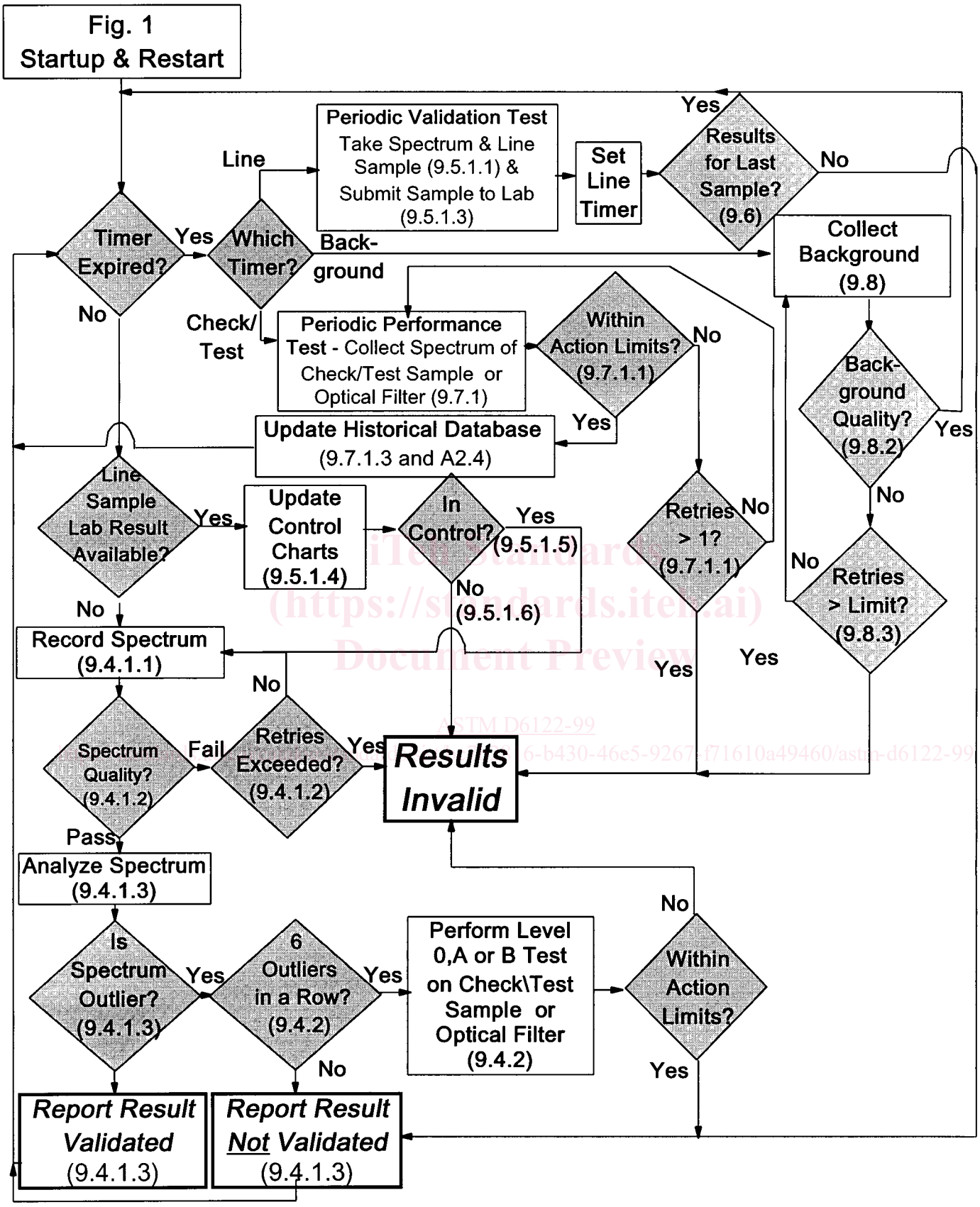

FIG. 2 Flowchart of Process Analyzer Validation Practice Normal Operation

the performance of the instrument. The objective of the check is to determine that current performance of the instrument is

consistent with performance which is known to produce valid analyses. Conduct this initial check out of the instrument within a short period (preferably within 24 h) after installation. Collect spectra of 20 check or test samples and analyze them using one or more of the Level 0, Level A, or Level B performance tests described in Annex A2 and Practice E 1866.

9.2.2 Compare the results for the initial performance tests to performance test action limits. These action limits may be based on historical data for the same tests, on simulations of the effects of performance changes on the analyzer results, or on a combination of historical and simulated data. Methods for establishing action limits are discussed in Annex A2 and Practice E 1866.

9.2.2.1 If the performance test results are within action limits, then the procedure continues with the initial validation tests. If the performance test results are not within action limits, check installation, instrument standardization or calibration transfer, or combination thereof, and correct the cause of the inadequate performance. Repeat the initial performance tests.

9.2.2.2 If action limits for performance tests have not been established, use the results for the initial performance tests to generate an initial historical database against which future tests can be compared, and continue the validation procedure with the steps described in 9.3. In the absence of historical data or performance simulations, the performance of the instrument cannot be verified, but shall be assumed. Should the analyzer fail to validate, inadequate instrument performance could be responsible.

9.3 *Initial Validation* (see Section 12 for details):

9.3.1 Once the initial performance tests are completed, collect spectra of 20 line and test samples and analyze them using the multivariate model. In order for the results to be used in the initial validation test, the spectra of the 20 line or test samples shall not be either outliers or nearest neighbor inliers (see Section 11 and Annex A3). Replace samples whose spectra are outliers or nearest neighbor inliers with other line or test samples.

9.3.2 Withdraw line samples from the process using methods described in Practices D 1265, D 4057, or D 4177, whichever is applicable, and analyze them by the primary method. The line sample shall correspond directly to the spectrum collected in 9.3.1.

9.3.3 Check that the standard deviation of the analyzer results for the 20 validation samples is at least 72 % of the reproducibility of the primary method for each property/component being modeled. If not, collect spectra of additional line or test samples, or both, until the standard deviation is adequate.

9.3.4 Compare values calculated by the analyzer to those obtained by the primary method using statistical tests described in Section 12. If the values are within statistical agreement, then the analyzer results are considered valid, and the analyzer can be used to analyze line samples. If the values are not within statistical agreement, then the installation, instrument standardization or calibration transfer, or combination thereof, are checked and corrected, and the procedure starts over with initial performance tests as described in 9.2.

9.4 *Normal Operation:*

9.4.1 Once the initial analyzer system validation is completed, normal operations for analysis of process samples may be conducted. Conduct tests of the performance of the analyzer and of the validity of the analyzer results on a periodic, regularly scheduled basis. When these tests are not scheduled, the normal application of the analyzer for on-line analysis proceeds as follows:

9.4.1.1 Collect a spectrum of the process sample.

9.4.1.2 Optionally, conduct tests on the spectrum in order to determine that the quality of the spectrum is adequate for use in estimating results by way of application of the multivariate model. Spectrum quality tests are generally defined by the instrument manufacturer or model developer, or both. If spectrum quality tests are used, allow a finite number of retries on the spectrum collection before the analyzer is considered inoperative, and the results produced invalid.

9.4.1.3 Analyze the spectrum using the calibration model, to produce one or more results, possibly uncertainties in these results, and statistics which are used to determine if the spectrum is an outlier or nearest neighbor inlier relative to the sample population used in the development of the calibration model (see Section 11 and Annex A3). If the spectrum recorded during normal operation of the analyzer is not an outlier or nearest neighbor inlier, then the calculated property values produced are considered valid as long as the analyzer quality control charts are up to date and the differences between the analyzer results and the primary method results are within control limits. If the spectrum recorded during the normal operation of the analyzer is an outlier or nearest neighbor inlier, then the specific results associated with that spectrum are considered to be invalid.

9.4.2 When six successive spectra recorded during the normal operation of the analyzer are all outliers, conduct performance tests to determine if the instrument performance is within action limits (see 10.3.3).

9.5 *Periodic Validation Tests:*

9.5.1 Conduct periodic analyzer validation tests at regularly scheduled intervals, preferably once a week (see Section 13).

9.5.1.1 Simultaneously, withdraw a line sample from the process and collect a spectrum of the process stream with the process analyzer.

9.5.1.2 Analyze the spectrum using the multivariate model to produce a result, and to produce outlier and nearest neighbor inlier statistics. If the spectrum is an outlier or nearest neighbor inlier, it cannot be used for the validation test, and the procedure starts over with 9.5.1.

9.5.1.3 Analyze the line sample by the primary method used in the development of the calibration.

9.5.1.4 Compare the analyzer and primary method results by plotting their difference on control charts as described in Section 13.

9.5.1.5 If the difference is within control limits, then the predicted result for the analyzer is considered to be valid.

9.5.1.6 If the difference is not within control limits, then the result for the analyzer are invalid. Check the control charts for the primary method (see Section 8) to ensure that the primary method is within control limits. If the primary method is not within control limits, determine and correct the cause of the

error, and repeat the primary method test. If the primary method is within control limits, conduct performance tests to check if the instrument performance is within action limits. If the instrument performance is not within action limits, service to the analyzer may be necessary.

9.5.2 Collect validation samples, analyze them by the primary method, and compare the analyzer and primary method results using control charts on a periodic basis. The exact period between validation samples will depend on the nature of the analyzer application. At minimum, collect and analyze a validation sample at least once within each seven-day period. More frequent validation testing may be appropriate for applications where analyzers are being used to certify products. The period between validation samples should not be less than the typical time required to obtain the reference data by way of the primary method.

9.6 If the laboratory, primary method results for a line sample are not available by the time the next time sample is scheduled to be collected, then the results produced by the analyzer are to be considered invalid until such time as the overdue results become available and the control charts are updated.

9.7 *Performance Tests:*

9.7.1 It is recommended that performance tests be conducted on a regularly scheduled basis, preferably daily, between the periodic analyzer validation tests. The objective of the test is to demonstrate that the analyzer performance is consistent between validation tests. Details on performance tests are given in Section 10, Annex A2, and Practice E 1866.

9.7.1.1 If the results for the performance tests are within action limits, continue operation of the analyzer.

9.7.1.2 If the results of the performance tests are not within action limits, then repeat the test. If the results of the repeat test are not within action limits, then the analyzer results are considered invalid, and the analyzer should be serviced.

9.7.1.3 If action limits have not been established for the performance tests, it is recommended that validation tests be performed more frequently to establish the historical database against which the limits can be set (see Annex A2 and Practice E 1866).

9.8 *Optical Backgrounds:*

9.8.1 Collect new optical backgrounds on a regularly scheduled interval, or when indicated by analyzer performance results.

9.8.2 Tests may be conducted on the collected optical background to determine its quality. Background quality tests are generally defined by the instrument manufacturer or model developer, or both.

9.8.3 If background quality tests are used, allow a finite number of retries on the spectrum collection before the analyzer is considered inoperative, and the results produced invalid.

10. Performance Tests

10.1 Performance tests are conducted to determine whether the performance of the instrument (the spectrophotometer, the optical cell, and all transfer optics in between) is adequate to produce spectra of the quality sufficient for valid analyses. Typically, check or test samples are introduced into the analyzer, the spectra of these samples are analyzed using the appropriate Level 0, Level A, or Level B performance test, and the results are plotted on charts and compared to action limits. For analyzers equipped with in-line probes, it may be impractical to remove the probe to conduct performance tests. For such analyzers, alternative procedures described in Annex A2 and Practice E 1866 may be used to conduct performance tests. Adequacy of the spectra is determined by comparison to a historical database of spectra of sufficient and insufficient quality. Alternatively, simulations of possible changes in instrument performance can be used to define the performance that is adequate for a given application. A description of Level 0, A, and B tests, and of methods for setting action limits for performance tests based on historical data and on simulations, are described in detail in Annex A2 and Practice E 1866.

10.2 When conducting the performance tests, operate the instrument in the most stable and reproducible conditions attainable, as defined by the manufacturer. Allow sufficient warm-up time before the commencement of any measurements. If the calibration model was based on spectra of samples held within a specified temperature range, then allow all samples, including check and test samples, to equilibrate to this temperature prior to spectral measurement. If possible, the optical configuration used for measurements of test and check samples should be *identical* to that used for measurement of line samples. If identical optical configurations are not possible due to analyzer design, the user should recognize that the performance tests may not measure the performance of the entire instrument. Data collection and computation conditions should be equivalent to those used in the collection of the spectra used in the calibration model. Introduce fresh reference material into the spectrophotometer cell for each measurement. Flow through the cell during the measurement is not required. Date and time stamp the spectral data used in performance tests, and store the results of the tests in a historical database.

10.3 *Timing of Analyzer Performance Tests:*

10.3.1 Conduct performance tests on a regularly scheduled basis, preferably daily, to test instrument performance consistency between validation tests. Compare the results of the performance tests with action limits for the tests. If a significant change in the performance is observed, conduct a second analysis to verify the change. If the significant change in performance is verified, mark analyzer results *not validated* until the cause and effect of the change can be determined. If the change in performance is not verified, conduct analyses of five additional check or test samples to demonstrate that the first occurrence was an anomaly, before continuing with normal operation.

10.3.1.1 The significance of a change in instrument performance may be unknown in the absence of historical data or simulations. In such case, more frequent validation testing may be required to demonstrate the relationship between analyzer performance and valid analyses. If, after a change in instrument performance is observed, the analyzer results remain in control, the change is not adversely effecting analyzer results. If, however, the analyzer results go out of control relative to the primary method, the change is adversely affecting analyzer results.

10.3.1.2 If historical data or simulations exist to demonstrate that change in performance is sufficient to produce invalid analyses, then service the analyzer to correct the problem. Service of this type is considered major maintenance, and initial performance and validation tests are required before resuming analyzer operation.

10.3.2 When an analyzer is installed, or after major maintenance has been performed, conduct 20 instrument performance tests using the check or test sample over a 24-h period to capture any diurnal performance variations. Compare the performance test results for the 20 samples with performance test action limits to determine if the analyzer performance is adequate. Add the test results for the 20 analyses to the historical database against which future performance tests are compared. Once these performance tests have been successfully completed, initiate the initial validation of the analyzer.

10.3.3 If, during the course of normal operation, the spectra of six successive samples are determined to be spectral outliers (see Section 11 and Annex A3), it is recommended that performance tests be conducted to demonstrate that the outlier diagnostics are responding to chemical changes in the process stream and not to changes in the instrument performance. If the results for the performance tests are outside action limits, then the outlier diagnostics may be responding to instrument performance and the analyzer should be serviced. If the results for the performance tests are within action limits, then the outlier diagnostics are most likely responding to changes in the process which are producing materials outside the range of the current calibration. If the process remains outside the range of the calibration for extended periods, it is recommended that the instrument performance be verified periodically using performance tests, until such time as the process returns to a state where the model is again applicable. If the process has changed so as to be permanently outside the range of the calibration, then a new model should be developed following Practices E 1655. Revalidate the analyzer with the new model following the procedure described herein.

10.3.4 Conduct performance tests if a bias is observed between the analyzer and primary method values to determine if the bias is the result of a change in instrument performance.

10.4 *Reference Materials for Instrument Performance Tests:*

10.4.1 Check samples are generally used for conducting performance tests. Check samples are single, pure, liquid hydrocarbon compounds or mixtures of liquid hydrocarbon compounds of definite composition. An alternate to using a check sample is to use an actual process sample called a test sample. For systems equipped with in-line probes, optical filters may be used as reference materials for instrument performance tests.

NOTE 1—Performance tests conducted on test samples are only intended to check the stability of analyzer performance over time. While the analyzer results for the test sample can be compared to the results for the primary method, such comparisons are not a substitute for the validation tests described in Sections 12 and 13. Analyzer results for test samples can be used in the calculation of the analyzer intermediate precision (see Section 16).

10.4.2 Details on reference materials for instrument performance tests are given in Annex A2 and Practice E 1866.

11. Verification that the Model is Applicable to the Spectrum of the Process Stream Sample – Spectral Outlier and Nearest Neighbor Inlier Detection

11.1 The spectra of the calibration samples define a set of variables that are used in the calibration model. If, when unknown samples are analyzed, the variables calculated from the spectrum of the unknown sample lie within the range of the variables for the calibration, the estimated value for the unknown sample is obtained by interpolation of the model. If the variables for the unknown sample are outside the range of the variables in the calibration model, the estimate represents an extrapolation of the model. Additionally, if the spectrum of the sample under test contains spectral features that were not present in the spectra of the calibration samples, then these features represent variables that were not included in the calibration, and the analysis of the sample spectrum represents an extrapolation of the model.

11.2 For the purpose of this practice, an analyzer result is considered valid only if the analysis involves an interpolation of the multivariate calibration model. Outlier detection methods are used to determine if an analysis represents an interpolation or an extrapolation of the multivariate model. The mathematics involved in outlier detection are described in Practices E 1655 and in Annex A3. The calculation of outlier statistics is by necessity an integral part of the analyzer software since these calculations shall be conducted each time the multivariate model is applied to a spectrum to produce a result. Appropriate limits for outlier tests will generally be set by the calibration model developer based on statistics from the calibration set.

11.2.1 A Mahalanobis Distance or leverage statistic is employed to determine if the spectrum being analyzed represents an interpolation or extrapolation of the variable space defined by the calibration model.

11.2.2 A spectral residuals statistic is employed to detect extrapolation of the calibration model due to spectra features which were not present in the spectra of the calibration set.

11.2.3 Optionally, a Nearest Neighbor Distance statistic can be employed to determine when the spectrum being analyzed falls in a sparsely populated region of the multivariate calibration space. While analyses of such spectra represent interpolation of the model, there may be insufficient information in the model to produce valid analyses for these samples. The use of a Nearest Neighbor Distance statistic is recommended if the calibration samples are highly clustered in the multivariate space. It is the responsibility of the model developer to determine if use of a Nearest Neighbor Distance statistic is appropriate. If a Nearest Neighbor Distance statistic is employed, then the results for any sample whose Nearest Neighbor Distance exceeds the predetermined limit are considered invalid. Such samples are referred to as Nearest Neighbor Inliers.

11.3 Annex A3 discusses available outlier detection methods. Further details on outlier methods and on notations used in their calculations are in Practices E 1655. Users may substitute other outlier detection methods providing they are at least as rigorous as those described in Annex A3 and Practices E 1655. If alternative outlier detection methods are substituted, it is the

user's responsibility to demonstrate that any analyzer results that are marked as invalid by the tests described herein are also marked as invalid by the substituted methods.

11.4 While it is generally preferable that the outlier statistics be generated using the same modeling method that was used to generate the calibration model, this is not required. For instance, MLR models do not provide spectral residual statistics. If an MLR model is used as the calibration model, an additional PCR or PLS model may be used to provide the necessary residuals statistics. If a supplementary model is used to generate outlier statistics, construct the supplementary model using the same set of calibration samples used for the predictive model, and apply the outlier statistics which will be used on the process analyzer system in the validation of the model in accordance with Practices E 1655.

11.4.1 Outlier tests detect differences in the spectrum of the process sample relative to the spectra of the calibration samples. These spectral differences may be due to differences in the chemistries of the samples, or due to differences in the performance of the spectrometer used to collect the spectra. Table 1 discusses inferences that may be drawn from outlier test results. The outlier tests by themselves do not distinguish between the instrument and the sample being the cause of the outlier result. Instrument performance tests may be used to help determine if the outlier test is responding to changes in the process or in the instrument.

12. Analyzer System Initial Validation

12.1 The initial validation of the analyzer is performed by comparing the analyzer and primary method results for a set of at least 20 initial validation samples. The primary method results are regressed against the analyzer results. A statistical test is performed on the regression results. The null hypothesis for the test is that the slope of the regression line is less than or equal to zero, that is, that there is no positive correlation between the two sets of results. If the null hypothesis is rejected, then there is a statistically significant positive correlation between the two sets of results.

12.2 *Initial Validation Samples:*

12.2.1 Initial validation of the analyzer is performed with a minimum of 20 samples. The actual number of samples used in the initial validation is designated by *n*. Spectra of these samples must yield potentially valid results (for example, the spectra must not be outliers) as defined in Section 11. For analyzer validation, select samples with chemical concentrations or physical properties which are interpolations within the range for which the calibration was developed and validated.

12.2.2 Select initial validation samples which exhibit sufficient variation in the property or composition being measured. At a minimum, it is recommended that the standard deviation of the analyzer results among the initial validation samples should be at least 72 % of the reproducibility of the primary method *for each property to be measured.*

NOTE 2—Seventy-two percent of the reproducibility is equivalent to twice the standard deviation of the reproducibility. Strictly speaking, the standard deviation of both the analyzer results and the primary method values are preferably at least 72 % of the reproducibility of the reference method to ensure that there is sufficient variation in the results to perform a meaningful statistical test. However, the primary method values (see Section 8) are not necessarily available at the time the initial validation samples are collected. If the analyzer does not pass the initial validation tests described in 12.1, and if the standard deviation in the reference values is less than 72 % of the reproducibility, the user should consider repeating the initial validation with samples that show a larger variability.

NOTE 3—If the primary method against which the analyzer results is being compared is not an ASTM method, the reproducibility of the method may not be known. The repeatability of the primary method values can be estimated from quality assurance data (see Section 8) and used in place of the reproducibility. The user should be aware that the repeatability will generally be smaller than the reproducibility, and that 72 % of the repeatability will typically represent less variation than 72 % of the producibility. If the analyzer does not pass the initial validation tests described below, the user should consider repeating the initial validation with samples that show a larger variability.

12.2.2.1 Samples in the required property range for validating one property may not be suitable for validating another property derived from the same spectral measurement. (For example, three motor gasoline grades may span five octane range but may have a constant Reid vapor pressure. They would, thus, be suitable for initial validation of an analyzer measuring octane, but not Reid vapor pressure).

12.2.2.2 While line samples are preferable, the process may not exhibit sufficient variation during the period of initial validation to provide the required sample variation. In this case, test samples that were not used for the model development may be included in the set of samples used for initial validation to achieve the required variation. Confirm the integrity of these test samples by appropriate testing prior to use. Preferably, test samples should not make up more than 25 % of the set of initial validation samples.

12.2.2.3 Check samples resembling the process stream may be used in place of test samples providing that their spectra are not outliers.

12.2.3 *Initial Validation Correlation (Slope) Test*—Test the correlation between the analyzer results, and the primary method results for the 20 initial validation samples by the following calculation:

12.2.3.1 Perform a regression of the primary method results, Y_r, versus the analyzer results, Y_a. Calculate the slope of the regression, *m*, as follows:

$$m = \frac{\sum (Y_a - \overline{Y}_a)(Y_r - \overline{Y}_r)}{\sum (Y_a - \overline{Y}_a)^2} \quad (1)$$

TABLE 1 Inferences Related to Outlier Detection or Instrument Failure

Mahalanobis Distance Test	Spectral Residual Test	Inferences	Status of Analyzer Result
Less than limit	less than limit	spectrum within range of calibration spectra	result valid if control charts are current and within control limits
Greater than limit	less than limit	possible instrument malfunction or model extrapolation due to sample component outside range for calibration	invalid result
Less than limit	greater than limit	possible instrument malfunction or model extrapolation due to sample absorption not present in calibration spectra	Invalid result
Greater than limit	geater than limit	possible instrument malfunction or model extrapolation	invalid result