Designation: E1655 – 05 (Reapproved 2012)

Standard Practices for Infrared Multivariate Quantitative Analysis[1]

This standard is issued under the fixed designation E1655; 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 These practices cover a guide for the multivariate calibration of infrared spectrometers used in determining the physical or chemical characteristics of materials. These practices are applicable to analyses conducted in the near infrared (NIR) spectral region (roughly 780 to 2500 nm) through the mid infrared (MIR) spectral region (roughly 4000 to 400 cm^{-1}).

NOTE 1—While the practices described herein deal specifically with mid- and near-infrared analysis, much of the mathematical and procedural detail contained herein is also applicable for multivariate quantitative analysis done using other forms of spectroscopy. The user is cautioned that typical and best practices for multivariate quantitative analysis using other forms of spectroscopy may differ from practices described herein for mid- and near-infrared spectroscopies.

1.2 Procedures for collecting and treating data for developing IR calibrations are outlined. Definitions, terms, and calibration techniques are described. Criteria for validating the performance of the calibration model are described.

1.3 The implementation of these practices require that the IR spectrometer has been installed in compliance with the manufacturer's specifications. In addition, it assumes that, at the times of calibration and of validation, the analyzer is operating at the conditions specified by the manufacturer.

1.4 These practices cover techniques that are routinely applied in the near and mid infrared spectral regions for quantitative analysis. The practices outlined cover the general cases for coarse solids, fine ground solids, and liquids. All techniques covered require the use of a computer for data collection and analysis.

1.5 These practices provide a questionnaire against which multivariate calibrations can be examined to determine if they conform to the requirements defined herein.

1.6 For some multivariate spectroscopic analyses, interferences and matrix effects are sufficiently small that it is possible to calibrate using mixtures that contain substantially fewer chemical components than the samples that will ultimately be analyzed. While these surrogate methods generally make use of the multivariate mathematics described herein, they do not conform to procedures described herein, specifically with respect to the handling of outliers. Surrogate methods may indicate that they make use of the mathematics described herein, but they should not claim to follow the procedures described herein.

1.7 The values stated in SI units are to be regarded as standard. No other units of measurement are included in this standard.

1.8 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:*[2]

D1265 Practice for Sampling Liquefied Petroleum (LP) Gases, Manual Method

D4057 Practice for Manual Sampling of Petroleum and Petroleum Products

D4177 Practice for Automatic Sampling of Petroleum and Petroleum Products

D4855 Practice for Comparing Test Methods (Withdrawn 2008)[3]

D6122 Practice for Validation of the Performance of Multivariate Online, At-Line, and Laboratory Infrared Spectrophotometer Based Analyzer Systems

D6299 Practice for Applying Statistical Quality Assurance and Control Charting Techniques to Evaluate Analytical Measurement System Performance

D6300 Practice for Determination of Precision and Bias Data for Use in Test Methods for Petroleum Products and Lubricants

E131 Terminology Relating to Molecular Spectroscopy

[1] These practices are under the jurisdiction of ASTM Committee E13 on Molecular Spectroscopy and Separation Science and are the direct responsibility of Subcommittee E13.11 on Multivariate Analysis.

Current edition approved April 1, 2012. Published May 2012. Originally approved in 1997. Last previous edition approved in 2005 as E1655 – 05. DOI: 10.1520/E1655-05R12.

[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] The last approved version of this historical standard is referenced on www.astm.org.

E168 Practices for General Techniques of Infrared Quantitative Analysis (Withdrawn 2015)[3]
E275 Practice for Describing and Measuring Performance of Ultraviolet and Visible Spectrophotometers
E334 Practice for General Techniques of Infrared Microanalysis
E456 Terminology Relating to Quality and Statistics
E691 Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method
E932 Practice for Describing and Measuring Performance of Dispersive Infrared Spectrometers
E1421 Practice for Describing and Measuring Performance of Fourier Transform Mid-Infrared (FT-MIR) Spectrometers: Level Zero and Level One Tests
E1866 Guide for Establishing Spectrophotometer Performance Tests
E1944 Practice for Describing and Measuring Performance of Laboratory Fourier Transform Near-Infrared (FT-NIR) Spectrometers: Level Zero and Level One Tests

3. Terminology

3.1 *Definitions*—For terminology related to molecular spectroscopic methods, refer to Terminology E131. For terminology relating to quality and statistics, refer to Terminology E456.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *analysis, n—in the context of this practice*, the process of applying the calibration model to a spectrum, preprocessed as required, so as to estimate a component concentration value or property.

3.2.2 *calibration, n*—a process used to create a model relating two types of measured data. In the context of this practice, a process for creating a model that relates component concentrations or properties to spectra for a set of known reference samples.

3.2.3 *calibration model, n*—the mathematical expression or the set of mathematical operations that relates component concentrations or properties to spectra for a set of reference samples.

3.2.4 *calibration samples, n*—the set of reference samples used for creating a calibration model. Reference component concentration or property values are known (measured by reference method) for the calibration samples and a calibration model is found which relates these values to the spectra during the calibration.

3.2.5 *estimate, n*—the value for a component concentration or property obtained by applying the calibration model for the analysis of an absorption spectrum.

3.2.6 *model validation, n*—the process of testing a calibration model with validation samples to determine bias between the estimates from the model and the reference method, and to test the agreement between estimates made with the model and the reference method.

3.2.7 *multivariate calibration, n*—a process for creating a model that relates component concentrations or properties to the absorbances of a set of known reference samples at more than one wavelength or frequency.

3.2.8 *reference method, n*—the analytical method that is used to estimate the reference component concentration or property value which is used in the calibration and validation procedures.

3.2.9 *reference values, n*—the component concentrations or property values for the calibration or validation samples which are measured by the reference analytical method.

3.2.10 *spectrometer/spectrophotometer qualification, n*—the procedures by which a user demonstrates that the performance of a specific spectrometer/spectrophotometer is adequate to conduct a multivariate analysis so as to obtain precision consistent with that specified in the method.

3.2.11 *surrogate calibration, n*—a multivariate calibration that is developed using a calibration set which consists of mixtures which contain substantially fewer chemical components than the samples which will ultimately be analyzed.

3.2.12 *surrogate method, n*—a standard test method that is based on a surrogate calibration.

3.2.13 *validation samples*—a set of samples used in validating the model. Validation samples are not part of the set of calibration samples. Reference component concentration or property values are known (measured by reference method), and are compared to those estimated using the model.

4. Summary of Practices

4.1 Multivariate mathematics is applied to correlate the spectra measured for a set of calibration samples to reference component concentrations or property values for the set of samples. The resultant multivariate calibration model is applied to the analysis of spectra of unknown samples to provide an estimate of the component concentration or property values for the unknown sample.

4.2 Multilinear regression (MLR), principal components regression (PCR), and partial least squares (PLS) are examples of multivariate mathematical techniques that are commonly used for the development of the calibration model. Other mathematical techniques are also used, but may not detect outliers, and may not be validated by the procedure described in these practices.

4.3 Statistical tests are applied to detect outliers during the development of the calibration model. Outliers include high leverage samples (samples whose spectra contribute a statistically significant fraction of one or more of the spectral variables used in the model), and samples whose reference values are inconsistent with the model.

4.4 Validation of the calibration model is performed by using the model to analyze a set of validation samples and statistically comparing the estimates for the validation samples to reference values measured for these samples, so as to test for bias in the model and for agreement of the model with the reference method.

4.5 Statistical tests are applied to detect when values estimated using the model represent extrapolation of the calibration.

4.6 Statistical expressions for calculating the repeatability of the infrared analysis and the expected agreement between the infrared analysis and the reference method are given.

5. Significance and Use

5.1 These practices can be used to establish the validity of the results obtained by an infrared (IR) spectrometer at the time the calibration is developed. The ongoing validation of estimates produced by analysis of unknown samples using the calibration model should be covered separately (see for example, Practice D6122).

5.2 These practices are intended for all users of infrared spectroscopy. Near-infrared spectroscopy is widely used for quantitative analysis. Many of the general principles described in these practices relate to the common modern practices of near-infrared spectroscopic analysis. While sampling methods and instrumentation may differ, the general calibration methodologies are equally applicable to mid-infrared spectroscopy. New techniques are under study that may enhance those discussed within these practices. Users will find these practices to be applicable to basic aspects of the technique, to include sample selection and preparation, instrument operation, and data interpretation.

5.3 The calibration procedures define the range over which measurements are valid and demonstrate whether or not the sensitivity and linearity of the analysis outputs are adequate for providing meaningful estimates of the specific physical or chemical characteristics of the types of materials for which the calibration is developed.

6. Overview of Multivariate Calibration

6.1 The practice of infrared multivariate quantitative analysis involves the following steps:

6.1.1 *Selecting the Calibration Set*—This set is also termed the training set or spectral library set. This set is to represent all of the chemical and physical variation normally encountered for routine analysis for the desired application. Selection of the calibration set is discussed in Section 17, after the statistical terms necessary to define the selection criteria have been defined.

6.1.2 *Determination of Concentrations or Properties, or Both, for Calibration Samples*—The chemical or physical properties, or both, of samples in the calibration set must be accurately and precisely measured by the reference method in order to accurately calibrate the infrared model for prediction of the unknown samples. Reference measurements are discussed in Section 9.

6.1.3 *The Collection of Infrared Spectra*—The collection of optical data must be performed with care so as to present calibration samples, validation samples, and prediction (unknown) samples for analysis in an alike manner. Variation in sample presentation technique among calibration, validation, and prediction samples will introduce variation and error which has not been modeled within the calibration. Infrared instrumentation is discussed in Section 7 and infrared spectral measurements in Section 8.

6.1.4 *Calculating the Mathematical Model*—The calculation of mathematical (calibration) models may involve a variety of data treatments and calibration algorithms. The more common linear techniques are discussed in Section 12. A variety of statistical techniques are used to evaluate and optimize the model. These techniques are described in Section 15. Statistics used to detect outliers in the calibration set are covered in Section 16.

6.1.5 *Validation of the Calibration Model*—Validation of the efficacy of a specific calibration model (equation) requires that the model be applied for the analysis of a separate set of test (validation) samples, and that the values predicted for these test samples be statistically compared to values obtained by the reference method. The statistical tests to be applied for validation of the model are discussed in Section 18.

6.1.6 *Application of the Model for the Analysis of Unknowns*—The mathematical model is applied to the spectra of unknown samples to estimate component concentrations or property values, or both, (see Section 13). Outlier statistics are used to detect when the analysis involves extrapolation of the model (see Section 16).

6.1.7 *Routine Analysis and Monitoring*—Once the efficacy of one or more calibration equations is established, the equations must be monitored for continued accuracy and precision. Simultaneously, the instrument performance must be monitored so as to trace any deterioration in performance to either the calibration model itself or to a failure in the instrumentation performance. Procedures for verifying the performance of the analysis are only outlined in Section 22. For petrochemicals, these procedures are covered in detail in Practice D6122. The use of Practice D6122 requires that a quality control procedure be established at the time the model is developed. The QC check sample is discussed in Section 22. For practices to compare reference methods and analyzer methods, refer to Practices D4855.

6.1.8 *Transfer of Calibrations*—Transferable calibrations are equations that can be transferred from the original instrument, where calibration data were collected, to other instruments where the calibrations are to be used to predict samples for routine analysis. In order for a calibration to be transferable it must perform prediction after transfer without a significant decrease in performance, as indicated by established statistical tests. In addition, statistical tests that are used to detect extrapolation of the model must be preserved during the transfer. Bias or slope adjustments, or both, are to be made after transfer only when statistically warranted. Calibration transfer, that is sometimes referred to as instrument standardization, is discussed in Section 21.

7. Infrared Instrumentation

7.1 A complete description of all applicable types of infrared instrumentation is beyond the scope of these practices. Only a general outline is given here.

7.2 The IR instrumentation is comprised of two categories, including instruments that acquire continuous spectral data over wavelength or frequency ranges (spectrophotometers), and those that only examine one or several discrete wavelengths or frequencies (photometers).

7.2.1 Photometers may have one or a series of wavelength filters and a single detector. These filters are mounted on a