Chapter 7
Certification and Approval Mechanisms for Non-criteria Pollutant Monitoring Systems; Calibration and Demonstration of Performance

7.1 Introduction

The purpose of this chapter is to aid in understanding how CEM systems are evaluated and approved for use in regulatory applications. Quality assurance procedures that are used on an on-going basis to maintain CEM data quality within acceptable limits are discussed. Understanding these issues is important because, 1) specific technical requirements in this area may constrain the selection of CEM equipment, 2) the complexity of a monitoring program at a particular facility may be significantly affected, and 3) the specification or choice of initial approval procedures and quality assurance activities may affect the cost and risk associated with a monitoring program.

Many non-criteria pollutant monitoring applications present special and difficult challenges in this area. The lack of appropriate calibration materials or procedures, inadequate performance specifications and quality assurance requirements, and the absence of reference test procedures all constrain the availability of non-criteria pollutant monitoring systems. Some of these issues are being addressed by instrument developers and control agencies and progress is being made in resolving problems. Many approaches that have been successful in the field of criteria pollutant monitoring cannot be transferred to monitoring of non-criteria pollutants. An overview of certification and approval mechanisms is presented in Figure 7-1. As indicated on the chart, initial monitor set-up activities are very important considerations for non-criteria pollutant monitors. For such monitors, the type of procedures and approaches used by the equipment supplier or facility personnel to establish proper calibration and operation of a measurement system are quite diverse and are highly dependent on the specific sampling and analysis system under consideration. For some systems, these initial set-up and calibration procedures necessitate understanding the intricacies of the analytical technique, the limitations of calibration check procedures, and the constraints of available calibration standards. The "black box" empirical approach traditionally used for criteria pollutant CEM systems is not applicable to many non-criteria pollutants. For many of these monitors the constraints which affect the initial set-up of the monitor also limit the methods which may be used to demonstrate the performance of the CEM system. Many of the issues associated with the initial set-up and calibration of non-criteria pollutant CEM systems are discussed in Section 7.2.

Methods and procedures used to demonstrate the performance of non-criteria pollutant CEM systems both initially and on an ongoing basis are shown in Figure 7-1. Monitoring system certification is based on determining whether the analyzer is properly calibrated, and if the sampling system can deliver the gaseous compounds to be measured to the analyzer within a prescribed tolerance. A number of ways to assess analyzer calibration and sampling system bias exist. Specifications for non-criteria pollutant monitoring systems generally require a performance specification test that includes calibration drift and accuracy tests. These tests are performed soon after installation of the monitor. Daily calibration checks are performed for virtually all CEM systems to assess data accuracy on a day-to-day basis. For some applications, development of a quality assurance (QA) plan and performing periodic accuracy checks may be required. Existing and proposed regulatory mechanisms for performing the initial certification test, conducting subsequent evaluations of monitor performance, and alternate techniques to assess a CEM system's effectiveness are discussed in Sections 7.3.4, 7.3.5, and 7.4.

Because the existing regulatory mechanisms do not address all of the potential issues and applications for non-criteria pollutant monitors, alternate instrumental protocols and acceptance criteria may be necessary for non-criteria pollutant CEM systems data validation. In some cases, approval mechanisms for instrumental test methods may be modified for this application. Techniques such as dynamic analyte spiking may be necessary to properly evaluate CEM system perfor-

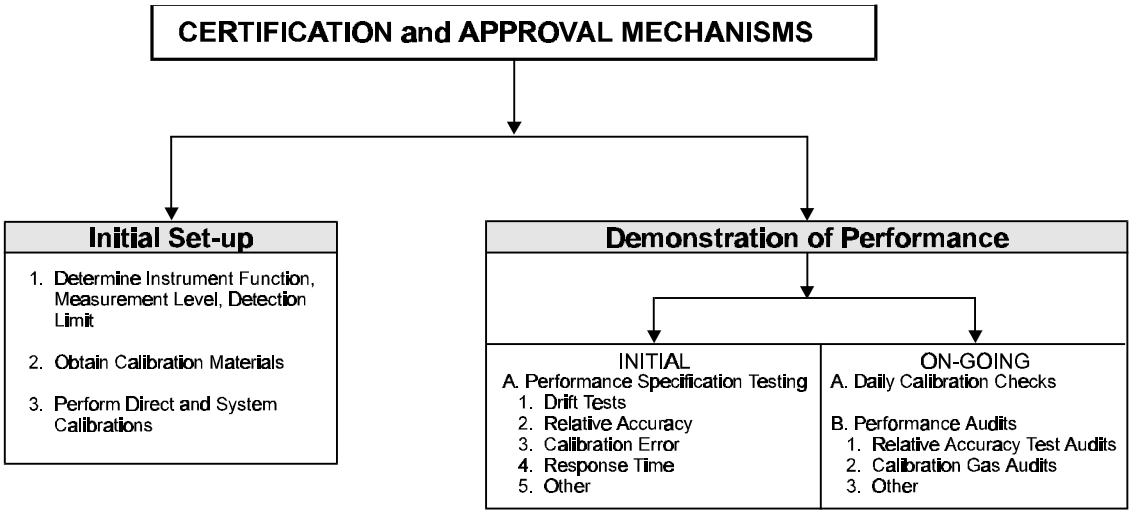

Figure 7-1. Certification and approval mechanisms.

mance necessary to properly evaluate CEM system performance, especially for reactive or condensable compounds, or compounds that have many interferences in the analytical procedures used for quantification. Alternative approaches such as those used in Germany, or those included in standards developed by the International Standards Organization may also be applicable to non-criteria pollutant monitors. These alternative approaches are discussed in Section 7.4.

7.2 Instrument Calibration

The term "calibration" is used to represent many concepts in the CEM field. These concepts include simple daily QA checks, routine adjustments to compensate for analyzer drift, benchtop adjustments made by the manufacturer, comparisons with manual test methods, and fundamental adjustments of various system components. Here, "calibration" is defined as the set of necessary activities that establish the relationship between the analyzer output (instrument response) and a series of standard materials that span the anticipated operating range. In the traditional and most straightforward case, a system measuring gaseous pollutants is calibrated by introducing a series of calibration gases at known concentrations to assess performance. Appropriate adjustments are then made so that the instrument provides a proper response over the measurement range. In other cases, where appropriate calibration materials, performance specifications or quality assurance procedures have not been fully developed, the project manager must evaluate the specific instrument response, the effective measurement range, and the procedures used to establish and verify proper calibration.

In this chapter, the term "analyte" will refer to gaseous compounds that are to be measured. Using this term, calibration can be said to establish the proper measurement system operation for the target analytes.

7.2.1 Instrument Function

"Instrument function" is a term that refers to the instrument's response versus concentration for each analyte. A graphical representation of response versus concentration defines the "shape" of the instrument function. For example, some instruments generate a linear response (i.e., flame ionization detectors, FIDs) through a given calibration range, while other instruments deviate from linearity at low or high concentration ranges (infrared). Some instruments have complex instrument functions such as a logarithmic or quadratic response with respect to concentration.

The instrumental function must be known for each analyte to select the appropriate calibration points over the anticipated operating range. For linear devices such as FIDs, zero and upscale calibration points may be sufficient. However, for non-linear instrument functions, additional calibration points that encompass each analyte concentration are necessary to ensure proper quantification. Figure 7-2 is a simplified representation of two instrument functions.

Multi-component analyzers frequently have different instrumental responses (response factors) for each gas (analyte) that is being measured. In all cases, the instrument function depends upon the analytes' physical properties and interferences from other flue

gas components. For example, an FTIR analyzer may exhibit a linear response for most compounds over a certain measurement range but will exhibit a non-linear response for compounds such as carbon monoxide.

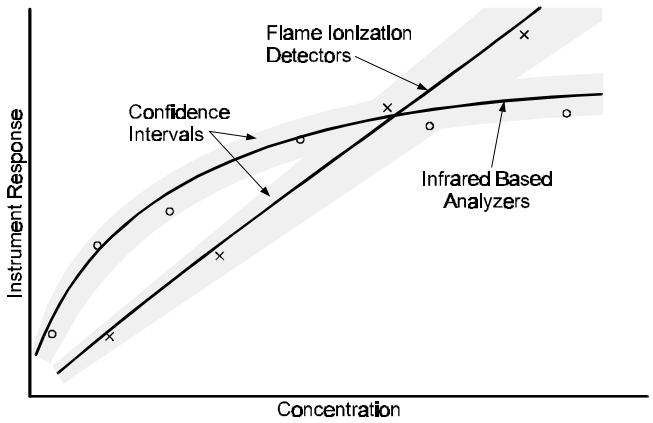

Figure 7-2. Instrument Function of Two Analyzers.

Inherent to the instrument function is the detection level achievable for each analyte. The "detection level," or detection limit, is defined as the lowest concentration level for each analyte that the instrument is able to measure. The detection limit often is determined by the instrument's "signal to noise (S/N) ratio." This is the ratio of the instrument response to the lowest level calibration standard, divided by the instrument response when no analytes are present (noise) which gives a good approximation of the S/N ratio, and thus the instrument's detection limit. An instrument S/N ratio of 10 should be adequate to detect most analytes; however, this should be verified by direct observation.

Another important characteristic of the instrument function is the range of uncertainty associated with the response. Many optical instruments are very repeatable. On the other hand, mass spectrometers exhibit greater imprecision in their quantitative output. Imprecision should not be mistaken for bias. The level of repeatability or precision associated with a specific instrument function must be considered in determining the appropriate number of cali-bration check repetitions to perform and in establishing control limits for initiating corrective action.

7.2.2 Calibration Materials

Calibration standards may be comprised of certified gaseous standards in gas cylinders, permeation or diffusion devices that generate known concentrations of gas, optical filters, gas filled cells, chemical solutions, or reference spectra. The target analytes, sampling system, and analytical techniques determine which type of calibration standard should be used. Regulatory requirements, the instrument function, CEM system design and the expected target analyte concentration range must be considered when selecting calibration standards. In general, a CEM system calibration technique should provide for checks at a zero, a mid-point, and a high-range concentration for each target analyte.

When certifying and calibrating multi-component analyzers, all, or as many of the analytes as possible, should be in the same gas cylinder or set of permeation/diffusion vials. This will enable simul-taneous calibration and calibration checks to be conducted providing for both cost and time savings. The calibration standard manufacturer should be contacted to ensure that calibration standards are available as stable blends at the required concentrations.

7.2.2.1 Gaseous Standards

Virtually all extractive systems can analyze calibration gas directly. Many point in-situ analyzers can analyze calibration gases also. Use of standard reference materials (SRMs) for routine calibrations is cost prohibitive. However, secondary calibration gas standards, called Protocol 1 gases, are referenced to the SRM, and are lower in cost. For some gases, SRMs are available from the National Institute of Standards and Technology. Protocol 1 (USEPA, 1996a) gases are certified by the gas manufacturer under a protocol accepted by EPA. Analyzers that accept calibration gases directly can use Protocol 1 gases (±1-2%), certified reference standards (usually ±2 to 5% accuracy), or permeation and diffusion devices that generate a known concentration of gas. Also, Part 51, Method 205 provides for the dilution of calibration gases for multi-point calibrations to reduce the number of cylinders and amount of gas that is required.

Protocol 1 gases have a period for which they are certified, whereas merely certified reference materials usually do not. Calibration standards for non-criteria pollutants usually are not available as Protocol 1 gases, and therefore, the best available standards are those that have reference values certified by the manufacturer. All calibration standards should be checked periodically to ensure that the compounds have not degraded from their certified values. Suspect calibration gases should be reanalyzed and recertified. Compounds that are particularly problematic in compressed gas form are those that are reactive or have the tendency to polymerize. Examples of these gases are halogenated acids (i.e., HCl and HF) and formaldehyde, respectively.

Some compounds are not available in gaseous form because their vapor pressures are too low, or they are not stable at elevated pressure. Gas standards for these compounds can be generated from the liquid or solid material using either permeation tubes (for liquids) or diffusion vials (for solids). Permeation tubes are sealed devices that contain the analyte and usually are constructed from semipermeable Teflon™ tubing. Diffusion vials are constructed from glass or quartz and have a small orifice through which they are filled with the solid material. Gas generators and permeation/diffusion apparatus are available with emission rates certified within ± 2 to 5 percent.

The generation of gaseous standards from liquids or solids is fairly simple. A sample of the liquid or solid contained in the vial is placed in a temperature controlled oven through which a carrier gas is swept. The concentration of the generated gas is directly proportional to the permeation/diffusion rate of the tube or vial. Permeation and diffusion rates are based upon the vapor pressure of the compound, the temperature and pressure of the oven, and the flow rate of the carrier gas. Figure 7-3 is a genera-lized schematic of a permeation/diffusion based gaseous standards generator.

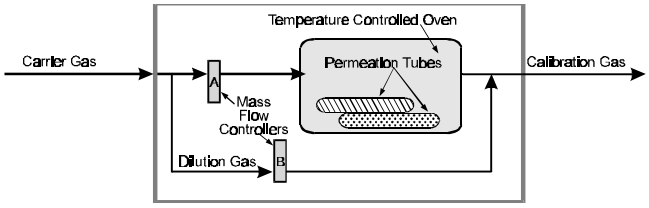

Figure 7-3. Generalized schematic of a permeation/ diffusion-based calibration gas generator.

Another means of generating gaseous standards is termed the "hanging drop" approach. In this procedure, a drop of the standard solution is introduced into a gas stream via a syringe. The syringe assembly is placed into a heated compartment so that the drop is continuously volatilized by carrier gas. Use of the hanging drop technique involves many parameters that introduce errors that could invalidate test results. Factors such as preparing the standard solution and determining the exact volatilization rate affect the resultant concentration of analyte in the calibration gas, and thus the accuracy of the calibration. This procedure should be used only when certified gaseous standards and permeation/diffusion standards are not available commercially.

7.2.2.2 Reference Spectra

Some analyzers employ reference spectra standards for calibration purposes. Reference spectra may be used repeatedly without having to recalibrate the analyzer, and thus provide savings in cost and time. Fourier transform infrared (FTIR) and tunable diode laser analyzers are examples where reference spectra standards are used to quantify data. A reference spectrum is composed of digitally stored data obtained from an instrument's spectral response to a known gas standard, acquired at a known concentration, temperature, pressure, and measurement pathlength. These spectra represent graphically the instrument's response (relative absorbance of electromagnetic radiation) with respect to wavelength. To employ reference spectra acquired at one point in time to quantify data acquired at a later date, the instrument response obtained during reference spectral acquisition is compared with the instrument response obtained during the time of data collection. A comparison of the instrument response at different points in time is performed by analyzing a calibration transfer standard (CTS). Similarly, reference spectra acquired on one instrument may be used to quantify data collected on another instrument provided that the two instrument responses are related through the CTS.

The CTS gas is used as the link between the instrument response at the time of calibration and the instrument response at the time of actual sample acquisition. A CTS spectrum generated during instrument calibration is compared to a CTS spectrum acquired during sampling. A mathematical transformation (correction factor) is then applied to match the reference spectrum CTS to that acquired during sampling. The exact same process is used to adjust the reference spectra so that they represent the instrumental response at the time of sampling, and therefore may be used to quantify data.

Figure 7-4 is a simplified diagram of the use of calibration transfer standards. In this figure, the CTS gas collected during field sampling is 0.85 times the response of the CTS collected with the reference spectra. Therefore, to employ the reference spectra, a mathematical transformation must be applied to the CTS reference spectrum to match the CTS collected with the sample spectrum, and thus adjust for the difference in instrument response. This same transformation is then applied to the analyte reference spectrum to quantify the analyte in the sample spectrum.

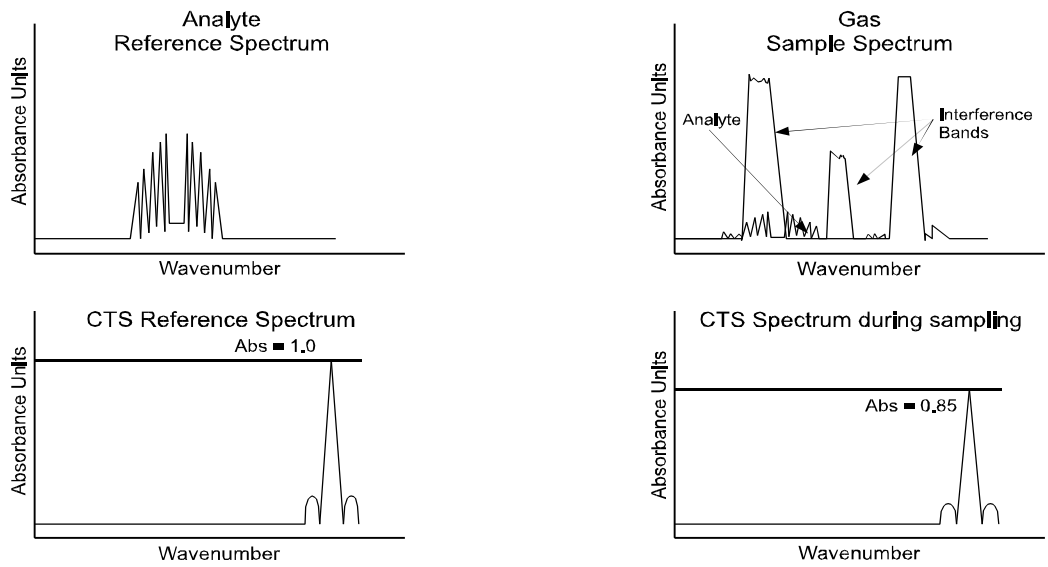

Figure 7-4. Use of calibration transfer standards (CTS).

The concept of calibration transfer is simple in theory; however, in practice it may be quite complex. In theory, any set of reference spectra may be used provided that the calibration transfer is valid. In practice, reference spectra are most often generated on the actual field instrument to minimize errors associated with the necessary CTS mathematical transformations and added noise. Detailed procedures for acquisition of FTIR reference spectra standards and their transference are beyond the scope of this text. These may be found in an EPA FTIR Protocol (USEPA, 1996b).

7.2.2.3 Other Calibration Techniques

Some analyzers are not capable of directly accepting calibration gases. Many cross-stack in-situ analyzers are not designed to accept calibration gases while the design of others allows the use of a flow-through gas cell. The approach used to check the calibration for these monitors also may be constrained because the flue gas cannot be removed from the measurement path and thus the zero concentration is difficult to verify. Some in-situ monitors have been designed or modified to allow the introduction of calibration gases (Jahnke, 1993, 1994). For many of these instruments, the calibration gas cell must be maintained at the same temperature as the flue gas sample to obtain valid calibration measurements.

For in-situ gas analyzers that cannot analyze calibration gases, a zero reflector or other device may be used to simulate the zero concentration condition. Calibration standards may be comprised of a gas-filled cell or an optical filter that is inserted into the measurement path to generate an upscale instrumental response. Where gas cells are mounted within a temperature controlled region of an analyzer, temperature compensation features that are used in the sampling mode may need to be disabled to adjust for optical absorption and/or gas density variations in the flue gas. In this situation, calibration checks with gas cells may not represent fully the performance of the entire measurement system. Problems in the temperature compensation feature cannot be detected if it is disabled during the calibration checks. When a significant component of the monitoring equipment is excluded from the routine calibration check procedure, a separate means of verifying the performance of the excluded component must be used.

The primary calibration for most in-situ analyzers that cannot accept calibration gases is established typically during laboratory bench tests at the manufacturing facility. Based on the design of the instrument, various internal functions are monitored to maintain performance within acceptable limits or to provide an indication to the operator that the initial calibration is no longer valid. The parameters and functions that are monitored depend on the analytical principal and the design of the instrument. Understanding the relationship among these parameters and the acceptable tolerances is usually more complex than simply evaluating a monitor response to an external calibration standard.

Some analyzers employ a correlation factor derived from comparing the instrumental response to another sampling method. The initial instrument function is determined, and the initial calibration is conducted by setting and monitoring internal instrumental parameters. Then the instrument response is correlated to concentration by comparison with reference test methods. Polynuclear aromatic hydrocarbon (PAH) analyzers are an example of such instruments. Calibration is factory set, and correlation to concentration is performed by comparing the instrument response with EPA Office of Solid Waste Method SW846-0010/8270 (SW846 Manual, November 1986). This instrumental correlation to reference method determined concentration must be performed on a source-by-source basis.

Continuous particulate monitors are another example of analyzers calibrated by correlation to manual sampling methods. This can be accomplished in accordance with the provisions of the International Standards Organization 10155 (ISO, 1995, see also 7.3.4 CEM System Approval Mechanisms in Germany, or the proposed Part 60, Performance Specification 11A). This standard establishes performance specifications and procedures for performing the calibration of continuous particulate matter measurement techniques relative to the results of a series of manual tests. Recalibration is required when conditions change at the facility that may affect the calibration.

For particulate monitors, the specific parameters that are monitored on a day-to-day basis (or in some cases on a continuous basis) depend on the measurement technique. For example, transmissometers used as particulate monitors incorporate a simulated zero mechanism (often a reflector to simulate zero opacity) and an internal optical filter (as an upscale check) to verify proper performance. Similarly, in-situ light scattering devices used as particulate monitors include a two point calibration check by temporarily not illuminating the effluent and by using the light source and a filter to check the response of the detector. These and other internal checks are sufficient to determine that the instrument operates within the limits established by the manufacturer.

7.2.3 Direct and System Calibration Procedures

Once the instrument function and necessary calibration materials are determined, the monitor may be calibrated. There are two conventional calibration methods for CEM systems that accept calibration gases. These methods are referred to as “direct” and “system” calibrations, respectively. Figure 7-5 is a generalized schematic of direct and system calibrations.

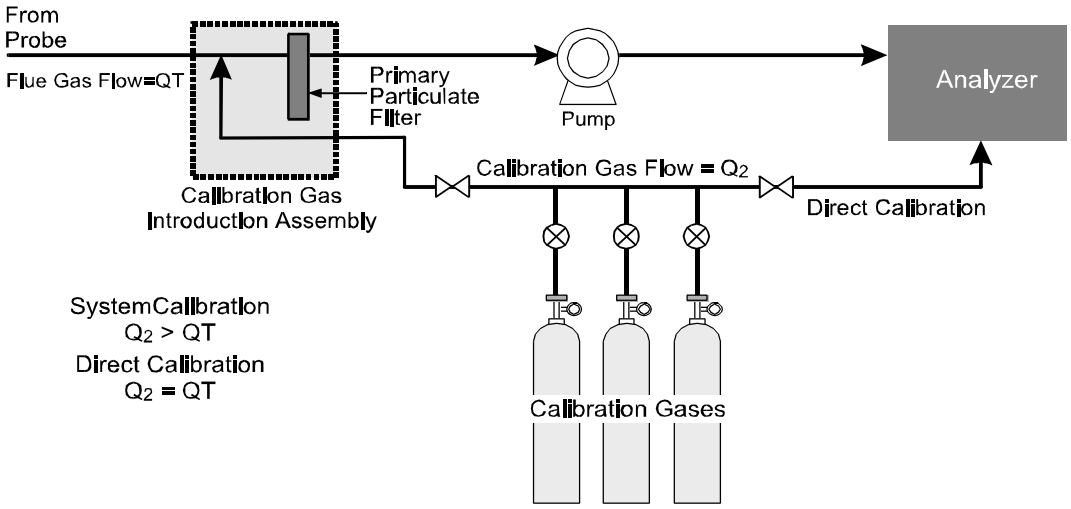

Figure 7-5. Direct and system calibrations.

A direct calibration consists of injecting the standards directly at the analyzer and recording the response. Direct calibration checks assess the stability of the analyzer and the extent to which the instrument has drifted from its zero and span calibra-tion points. Direct instrument calibrations test the instrument function only.

A system calibration injects the calibration stan-dards at a point in the sample delivery system as close to the probe tip as possible. Calibration standards must be introduced at a point upstream of the primary particulate matter filter to account for any filter/particulate cake reactions with the anal-ytes. System calibrations provide an indication of the sample handling system's effectiveness to deliver the analytes to the analyzer, as well as testing the instrument function.

Optimally, direct and system calibrations should agree. Large differences between these two calibration techniques for the same analytes could indicate that a leak is present or that losses due to chemical or physical reactivity are occurring. Losses due to adsorption and condensation are the most typical for non-criteria pollutants. System calibrations are also necessary to determine if compounds may have adsorbed on the filter. Adsorbed or condensed compounds that volatilize later can cause a positive measurement bias.

For those analyzers that measure multiple components, daily calibrations of the instrument may not be economically feasible for all compounds that are measurable. In this instance, surrogate compounds must be chosen for calibration checks. Surrogates are target analytes that can also represent a range of other target analytes through similar instrumental response, sampling system bias response, and other important characteristics. Selection of surrogates depends on their ability to represent more than one of the target analytes, based on similar physical properties. An example of surrogate selection is the use of one water soluble compound to represent many. If a CEM system is designed to measure HCl, formaldehyde, acetaldehyde, ammonia, methanol, and benzene, toluene and xylenes (BTX), the use of HCl and benzene may be possible as surrogates for the entire population of compounds. HCl will represent the water soluble and reactive compounds, while the benzene can represent the aromatic compounds. Protocols using surrogate recovery check compounds have been prepared by the Portland Cement Association based upon using surrogate compounds (Kinner, 1995). In these protocols, three recovery check compounds were used to represent fourteen volatile organic compounds to reduce the cost and complexity of the testing program.

Optical analyzers that employ a reference spectral library for quantification also need to perform calibration status checks on a periodic basis. This ensures that a valid calibration transfer exists be-tween the time of reference spectra acquisition (reference library) and the time of sample collection and analysis. A gaseous calibration transfer standard (CTS) should be analyzed and the response compared to the response at the time the reference spectra were obtained. Direct and system calibrations using the CTS should be performed so that the effect of sampling system bias on instrumental response may be measured. Ethylene and carbon monoxide have been successfully used as calibration transfer standards for extractive FTIR measurement systems (USEPA, 1993; Kinner, 1994).

Software audits also can be used as a tool to evaluate optical CEM system quantitative algorithms and their ability to accurately quantify data. Software audits may consist of synthetic spectra (a number of reference spectra superimposed upon one another), or actual measurement data that challenges the ability of the analysis technique to distinguish between the target analytes and their interferences. Such audits can be useful where a computer is used to analyze broad-band spectral measurements such as those employed by an FTIR. In such cases, the selection of particular spectral regions over which to perform an analysis, or the particular procedure used to remove the affects of water or other interferences, may significantly affect the results.

An audit using artificially generated spectra, or spectral data files from prior measurements obtained at another source, can be employed to gauge the effectiveness of the analytical procedure. However, audits are likely to generate one-sided conclusions. For example, an analysis routine that cannot properly identify target analytes in an audit spectrum most likely will fail to identify these analytes in actual samples. The converse is not necessarily true. Obviously, a software audit cannot evaluate the physical performance of a sampling system or the performance of an analyzer.

7.3 Demonstration of Performance; Certification and Quality Assurance

CEM systems installed to comply with regulatory requirements must demonstrate acceptable performance both initially and on a continuing basis. The specific requirements vary for different regulatory

applications (see Chapter 2) or according to specific conditions or requirements included in the facility's permit. Nevertheless, most CEM systems are subject to a performance specification test, or "certification test," soon after installation. Source operators are required to perform daily checks of CEM systems calibration status to ensure the validity of data on a day-to-day basis. Some sources are required to develop quality assurance plans and conduct periodic audits of CEM systems performance. In other cases, the regulatory agency relies on a general requirement for the proper operation and maintenance of the CEM system and can require performance specification tests to be conducted as necessary.

7.3.1 EPA Performance Specifications for Non-criteria Pollutant Monitoring

Performance specifications are a set of procedures and criteria that are used to evaluate the acceptability of the CEM system at the time of installation and whenever specified in the applicable regulations. Existing EPA performance specifications (PS) for non-criteria pollutant analyzers address monitoring of VOCs and TRS only. Performance specifications for H_2S in refinery fuel gas also have been developed to provide an alternative to measurement of SO_2 emissions at many emission points. Existing EPA performance specifications for non-criteria pollutants are summarized in Table 7-1 (requirements from PS 2 are shown for comparison).

Performance specifications for VOC CEM systems are included in PS 8 and PS 9 of Part 60 for sources subject to New Source Performance Standards. PS 8 is applicable to analyzers that provide a total hydrocarbon response and PS 9 applies to analyzers that speciate hydrocarbons using a gas chromatograph. Neither of these performance specifications require the use of a specific type of detector.

Performance specifications for hydrocarbon monitors also are included in 40 CFR, Part 266, Appendix IX for sources regulated by the boilers and industrial furnaces (BIF) rule, and in 40 CFR, Part 503 (subpart E) for sewage sludge incinerators. These specifications <u>require</u> a heated sampling system and heated flame ionization detector (FID) analyzer.

Performance specifications 8A for hydrocarbons, 10A for multi-metals, 11A for particulate matter (PM), 12A for mercury (Hg), 13A for hydrochloric acid (HCl), and 14A for chlorine (Cl_2) monitors are proposed in the hazardous waste combustor rule. (FR 4/19/96) This proposed rule covers both hazardous waste incinerators and BIF sources. These performance specifications will be included in Part 60 if this proposed rule is adopted. All of the performance-based specifications listed above are derived from PS 2, ISO 10155 (for particulate monitoring), and various instrumental test methods.

7.3.2 Performance Specification Testing

A performance specification test (PST) is performed soon after installation of the CEM system and periodically as specified in the applicable regulations. The test is performed in accordance with procedures detailed in the applicable performance specifications previously described. In some regulatory applications (and in common usage), this initial test is referred to as a "certification test." The PST usually includes a calibration drift test and an accuracy test. (For Part 75 applications, the term "calibration error test" is used to describe the drift test and daily zero and upscale calibration checks. This terminology is avoided here to minimize confusion.) The accuracy test is often a "relative accuracy test" involving comparison with results of an independent reference test method (RM). For some monitors, the accuracy test requires analysis of several calibration standards to determine calibration error only. Depending on the specific regulatory requirements, the PST also may include response time or cycle time tests, short-term drift tests, and demonstrations that the measurement system can operate continuously for prescribed periods without unscheduled maintenance or repairs. (Part 75 performance specifications are described in Chapter 2, Acid Rain Program.)

7.3.2.1 Calibration Drift Test

A typical PST includes a calibration drift test performed over 7 consecutive days. The calibration drift test evaluates the stability of the monitor response to the procedure used for daily checks of the monitoring system (see also Daily Calibration Checks). The calibration drift test is conducted by introducing the zero and span gases, gas cells, optical filters, or electric signals to the system and recording the instrument response. The deviation of the instrument response at the zero and upscale calibration point must meet the applicable criteria. For example, the allowable calibration drift in PS 8A is 2.5% of the span. Therefore, a hazardous waste incinerator required to monitor hydrocarbons subject to PS 8A has an allowable analyzer span value of 100 ppm, and may have a maximum of ± 2.5 ppm drift at the zero and span level for each 24-hour period over the seven consecutive days during the performance specification test. The tester is allowed to adjust the calibration accordingly after each daily check during the drift test.

Table 7-1. Existing Performance Specifications for Selected Pollutant Monitors

PS #	40 CFR	PERFORMANCE SPECIFICATION TEST				DAILY CALIBRATION REQUIREMENTS
		Calibration Drift	Calibration Error	Relative Accuracy	Response Time	
2* SO_2 NO_x	Part 60 Appendix A	7 day test 2.5% of span	N/A	±20% of Reference Method or 10% of standard	N/A	zero and span drift within 2X's PST before required adjustment
5 TRS	Part 60 Appendix A	7 day test 5% of span (6/7 days)	N/A	Same as PS2	N/A	Same as PS2
7 H_2S in fuel gas	Part 60 Appendix A	7 day test 5% of span (6/7 days)	N/A	Same as PS2	N/A	Same as PS2
8 Total VOCs	Part 60 Appendix A	Same as PS2	N/A	Same as PS2	N/A	Same as PS2
9 GC CEMS VOCs	Part 60 Appendix A	N/A	7 day test 3 injections at three levels Precision and Linearity $r^2 = 0.995$ CE $\leq$10% for all levels	Performance Audit during CE testing ±10% criteria	<5 minutes or as specified	Triplicate mid-level calibration checks must be within 10% of certified value or repeat initial 3-point calibration
BIF[1] THC	Part 266 Appendix IX	7 day test 3% of span Annual PST	5% of span for all calibration points	N/A	120 seconds for 95% step change	zero and span drift <3 ppm
SS**[1] THC	Part 503 Subpart E	7 day test zero and span within 6%	zero within 5 ppm mid- and span within 10 ppm	N/A	200 seconds for 90% step change	Same as PS2

* Included for comparison purposes
** Sewage Sludge Incinerators
1 No number assigned

7.3.2.2 Relative Accuracy Test

Relative accuracy (RA) tests are required for almost all criteria pollutant CEM systems except for opacity monitors. A relative accuracy test is also required for certain non-criteria CEM systems. The RA may be conducted concurrently with the 7 day calibration drift test. The relative accuracy test involves comparing the average CEM system response with the integrated average pollutant concentration or emission rate for each corresponding reference method test run. A minimum of nine RA comparative test runs are required; however, EPA specifications allow the rejection of up to three test runs. Therefore, 12 runs usually are performed in practice with the best nine results used in the RA calculations. The RA specification for a CEM system subject to PS 8 is $\leq$20% of the average reference method (RM) result or $\leq$10% of the applicable emission standard, whichever is greater.

Relative accuracy calculations are based upon the results of at least nine comparative runs. The difference between the run averaged CEM system results and the reference method is calculated and averaged. The absolute value of the summed differences for each run plus a confidence coefficient (CC) is divided by the average of the reference method values (RM_{ave}):

$$RA = [|d| + CC]/RM_{ave} * 100 \quad \text{Equation 7-1}$$

The confidence coefficient is expressed in terms of a statistical t-value according to the following equation for a t-value at a = 0.025 for n-1 degrees of freedom:

$$CC = t_{0.975}(S_d/n^{1/2}) \quad \text{Equation 7-2}$$

Where

$t_{0.975}$ = Corresponds to the probability that a measured value will be biased 2.5% at the 95% confidence level.

S_d = The standard deviation of the differences of the 9 data pairs.

n = Number of test runs.

As shown above, the relative accuracy result includes both mean difference and confidence coefficient terms based on the paired CEM system and reference test results. Both the accuracy and precision of the paired measurements are evaluated. A user must recognize that the relative accuracy is affected by the accuracy and precision of both the monitoring system and the reference test method. Failure of a relative accuracy test may be due to problems with the CEM system, problems with the reference test methods, problems with the representativeness of the sampling location, or other factors. A failed test requires careful investigation to determine the cause, after which the test must be repeated.

Relative accuracy is based on comparing concurrent monitoring data and reference method test results. Therefore, the response time of the CEM system must be taken into account to directly compare the monitoring with the reference method results. Where instrumental reference test methods are used, the response time and sampling frequency of the reference method must be considered also. For relative accuracy tests with these systems, sampling runs must be of sufficient duration to allow comparisons of valid averages obtained from both the CEM system and the reference system.

Where CEM systems are compared to reference methods that provide integrated results (e.g., impinger trains and most other manual methods), the exact start time and stop time for manual sampling runs (and time associated with interruptions for port changes or other problems) must be included. When substantial changes in emission concentrations or emission rates occur over the course of the run, or when short duration spikes occur during runs, comparisons between integrated reference methods and CEM systems are often invalid.

The relative accuracy comparison of a CEM system to a reference method assumes that the RM gener-ates accurate and precise data, and that the source testers have adequate experience in correctly executing the reference method. Interferences affecting the test method results can be mistaken for bias in the CEM system measurements. Use of an imprecise test method increases the magnitude of the confidence coefficient in the relative accuracy calculation. Similarly, the imprecision due to sloppy or inexperienced testers can affect the results dramatically. These issues were initially of great concern in criteria pollutant monitoring applications but eventually were minimized through experience and development of more precise instrumental test methods for use in place of wet-chemistry methods for most applications. For non-criteria pollutant monitoring applications, these factors are still of great concern because: 1) interferences affect some reference methods when extended to new applications, 2) sample recovery documentation and other QA criteria are not included in all methods, 3) lower and potentially variable emission concentrations are encountered, 4) the precision of some methods is poor and for some other methods is unknown, 5) some methods are not

demonstrated fully where they applied before promulgation of standards at the specific source category, and 6) source testers have much less experience in using these methods for performing relative accuracy tests. In summary, an imprecise, inaccurate, or improperly conducted reference method test will increase the probability that the CEM system will fail the PST even if it is functioning and calibrated properly.

Further disagreements may result in relative accuracy testing due to actual technical limitations. For example, PS 8 includes a relative accuracy specification, however, the only reference method that generates a total hydrocarbon response is EPA Method 25A (Part 60, Appendix A), which specifies the use of a flame ionization detector (FID). Method 25A is the only reference method directly comparable to a CEM system installed to satisfy PS 8. However, PS 8 does not specify the use of a particular detector; it simply specifies instrumentation that generates a total hydrocarbon response.

Presuming that the CEM system uses an FID, then the relative accuracy test would involve comparing two FID measurement systems. Because the two systems use the same analytical technique, interferences within the effluent stream could go undetected. Additionally, differences in the compound-specific response factors between two FIDs could affect the outcome of the relative accuracy test even though both systems are properly calibrated using the same standard (Cone, 1989).

Also, consider the example where a photoionization detector (PID) is chosen for the CEM system installed to comply with PS 8. A relative accuracy comparison with a Method 25A FID may not be appropriate because PID and FID responses differ substantially for different classes and functional groups of hydrocarbons. Because neither the CEM system nor the reference method provides any information regarding the speciation of individual compounds within the flue gas, one is unable to decide whether apparent agreement or disagreement between the CEM system and the reference method was due to differences in response factors or other calibration/accuracy issues. In this case, the outcome of the relative accuracy test is likely to be indeterminate and may generate more controversy than answers.

The BIF and sewage sludge incinerator performance specifications for hydrocarbon monitors, and the proposed PS 8A, specify using a heated FID and heated sampling system. PS 8A requires a three-point calibration error test rather than a relative accuracy test comparison using Method 25A. This approach avoids the relative accuracy test issues described above for PS 8. However, the three-point calibration error test, using calibration gases, provides only a limited evaluation of the monitoring system. (See also Section 7.3.2.3.)

Proposed performance specifications 10A (multi-metals CEM systems) and 12A (mercury CEM systems) include relative accuracy specifications (USEPA, 1996c). Although multi-metals CEM systems are not commercially available at this time, prototypes have been developed using a variety of approaches. Limited field tests have been conducted. Calibration standards and techniques, measurement sensitivity, and accuracy are critical technical issues for these developing measurement systems. The accuracy and precision of proposed Method 29 has not been demonstrated to allow for relative accuracy testing of multi-metals CEM systems over the expected range of pollutant concentrations (Brown 1996, Laudal 1996). Modifications to proposed Method 29 are being developed.

Mercury may exist as elemental mercury (Hg^0) and as other oxidized Hg (Hg^+, Hg^{+2}) species in effluent streams. Mercury CEM systems generally sense elemental mercury and employ sample conditioning elements that convert all mercury in the sample stream to this form (Richards, 1996, Roberson, 1996). (See also Chapter 3.) Proposed EPA Method 29 attempts to speciate the different oxidation states of mercury. Studies have been conducted on coal-fired facilities to evaluate the performance of Method 29 by spiking elemental mercury and $HgCl_2$ (Laudal, 1996). These investigators have shown that the presence of significant quantities of SO_2 (> 1,500 ppm) and the presence of 10 ppm chlorine (Cl_2) affects substantially the collection of elemental mercury and the determination of mercury species present. Other mercury test methods have been shown to be affected significantly by the presence of NO_x in the sample stream. Until the performance of the manual test methods are understood and the effluent conditions that affect the results are well known, an appropriate and reliable relative accuracy test procedure cannot be devised for mercury CEM systems.

Proposed PS 13A and PS 14A for HCl and Cl_2 specify calibration gas audits instead of comparison to reference methods. The reference method for HCl and Cl_2 is EPA Method 26A, which collects the gaseous components in an impinger train and measures the dissociated chloride ions by ion chromatography. Proposed PS 13A and 14A specify that HCl

and Cl_2 must be measured in the gas phase as opposed to the measurement of the dissociated chloride ion in solution. A cylinder gas audit only is required in these performance specifications, not a relative accuracy test employing Reference Method 26. HCl continuous emission monitoring requirements in Pennsylvania; however, require a relative accuracy test using Method 26. Pennsylvania requirements do not mandate that HCl and Cl_2 be measured in the gas phase.

Even where reference methods have been developed for the pollutants of interest, the comparison of CEM system and RM data may not assess the ability of the CEM systems to measure the compounds of interest accurately. Consider for example measurement of formaldehyde which is a very water soluble, reactive compound and which may polymerize in sample delivery systems. Several reference methods exist for the collection and measurement of formaldehyde. These methods rely on sample collection in high purity water, the Parosanaline Method (EPA TTN EMTIC BBS, 919-541-5742); in dinitro-phenyl hydrazine (DNPH) impregnated silica gel tubes, NIOSH Method 3500, (NIOSH 1994); or in sampling trains containing DNPH (SW846-0011). The results provided by all three of these methods are biased by the presence of other aldehydes and ketones (Serne, 1993). In addition, the reaction of formaldehyde with DNPH is very pH specific. The two DNPH methods are sensitive to the presence of acid gases (e.g., HCl, SO_2, H_2SO_4, etc.) and base gases (NH_3). Thus, the use of these methods at any source having significant concentrations of these compounds in the flue gas would affect adversely relative accuracy test results for a formaldehyde CEM system.

In summary, relative accuracy specifications and test procedures have long served as the primary evaluation standard for the acceptability of criteria pollutant CEM systems. For some non-criteria pollutants and monitoring applications, the relative accuracy test continues to play an important role in demonstrating performance and certifying CEM systems. However, for other applications relative accuracy tests are inappropriate or must await future development of appropriate reference test procedures.

7.3.2.3 Calibration Error Tests

As used here, a "calibration error test" consists of multi-point comparison of monitor response to a series of calibration materials. (The usage of "calibration error test" varies. In the Part 75 Acid Rain Program, this term is used to refer to calibration drift tests and daily checks. In Part 75, the term "linearity test" is used to refer to multi-point accuracy tests relative to calibration materials.) In certain applications, as indicated below, the calibration error test is required in place of the relative accuracy test. However, in other regulatory applications, both a calibration error test and a relative accuracy test are included in performance specification tests.

As previously mentioned, BIF and sewage sludge incinerator performance specifications for total hydrocarbon monitors, and proposed PS 8A, do not include a relative accuracy test but instead require a calibration error test to determine the accuracy of the monitoring system. The test involves analysis of a zero gas and two upscale calibration gases. The calibration gases used for these tests are propane in a balance gas of nitrogen. The instrument response of these measurement systems to different compounds (i.e., "response factors") is known to vary as a function of: 1) the number of the carbon atoms, 2) the degree of carbon bond saturation (carbon double and triple bonds) within the molecule, 3) the presence of chlorine, and 4) other instrument-specific factors. However, the determination of these response factors is not included in the calibration error test. Furthermore, the effect of other effluent stream constituents (e.g., moisture, CO_2, CO, O_2, NO_x, etc.) on the accuracy of the monitoring data can not be determined by the calibration error test. The calibration error test in this case may therefore not adequately validate the monitoring system for actual flue gas hydrocarbon measurements.

PS 9 applies to CEM systems that employ a GC to separate organic compounds. As with PS 8, it also does not specify a particular type of detector. Multiple target analytes are speciated by an instrument-specific calibration curve generated from certified calibration standards. A preliminary calibration error test of the analyzer is accomplished by analyzing low-level (40-60%), mid-level (90-110%), and high-level (140-160%) calibration standards that represent the percent of the measured concentration levels for all analytes. A screening test using the appropriate instrumental test method may be required to identify the effluent analytes and determine their relative concentrations.

The PST required by PS 9 includes a performance audit test and a seven-day calibration error (CE) test period during which the initial instrument performance is evaluated. The CE is performed daily by analyzing calibration standards at each level in triplicate (3 analyses X 3 levels = 9 analyses/day). The calibration error for each analyte is determined by comparing the average instrument response for each level of triplicate analyses to the certified reference

value of the calibration standards. The calibration error may not exceed ± 10% for each analyte.

The performance audit test required by PS 9 consists of analyzing an EPA audit gas during the seven-day CE test. This test is intended to provide a somewhat independent check because the audit gas must be obtained from a different supplier than the gases used for the CE test. The concentration of the analytes in the audit gas generally are not known to the user.

7.3.2.4 Response Time Tests and Cycle Time

The measurement system response time is important for monitoring applications with short averaging or reporting periods or where spikes in the effluent concentration occur. The sampling system response time is determined by monitoring the instrumental response to a calibration standard while performing a system calibration. Upscale and downscale response times are determined by introducing high-range and zero calibration gases respectively. System response time is defined as the time it takes the analyzer to reach 95% of the step-change in concentration.

Determination of the measurement system response time can be important even when it is not specified in the applicable regulations. The system response time must be considered when conducting the relative accuracy test. For measurement systems that accept calibration gases, the system response time must be known so that calibration gases are introduced for a sufficient period to provide a stable measurement response.

Some regulations specify the frequency for obtaining measurements. 40 CFR 60.13 specifies that all measurement systems must complete a cycle of sampling, analysis, and data recording at least once during each 15 minute period. In contrast, BIF regulations specify that a sample must be acquired and analyzed every 15 seconds. The average emission rate is computed and recorded at least every 60 seconds, with an hourly rolling average calculated each minute. Usually, the cycle time is immediately apparent from the design of the CEM system and programming of the data acquisition and handling equipment. In some cases, time-shared extractive systems are used to allow a single set of analyzers to monitor emissions from several locations. Other functions such as particulate filter or sampling line blow-back or purges may also be included for certain CEM systems. These functions must be considered when determining the cycle time for some systems.

7.3.2.5 Other Performance Test Criteria

Some regulations include additional performance specifications and test procedures. For example, transmissometers may be used as particulate mass concentration monitors and as opacity monitors. These instruments, if subject to the proposed revisions to Performance Specification 1, would be subject to a short-term drift test to detect diurnal variations in monitor performance. This drift test includes zero and upscale calibration drift checks at one-hour intervals over a period of 24 hours. Similar short-term drift tests were included in Part 60 performance specifications prior to 1983 and are included in some state monitoring requirements (e.g., Penn sylvania).

Many performance specification test requirements include an “operational test period” during which the CEM system must operate continuously without unscheduled maintenance or repairs. This is typically a one week (168-hour) test. Such a test period may seem insignificant for a measurement system that is intended for continuous monitoring over a period of many years. However, early experience with criteria pollutant monitoring systems revealed that this was a difficult specification for many CEM systems to meet. Although much progress has been made in improving the ruggedness and reliability of conventional systems, the application of new technologies or new measurement systems attempting to utilize laboratory instruments may require that more attention be devoted to achieving reliable operation.

7.3.3 Ongoing Performance Checks and Quality Assurance

Additional performance checks and quality assurance activities are required to be performed after initial installation, calibration, and performance testing (certification) of a CEM system. In general these include whatever is necessary to properly operate and maintain the monitoring system, daily calibration checks to determine the ongoing validity of the data and, in some cases, periodic quality assurance audits.

7.3.3.1 Daily Calibration Checks

Daily calibration checks are the primary basis for evaluating the validity of data on a day-to-day basis. This is of fundamental importance for many sources because 1) reporting of emissions data is based on daily averages, 2) monitor availability may be expressed as a percentage of valid source operating days, and 3) monitor control limits and the corresponding requirement to initiate corrective action are based on the daily calibration check responses. Also, enforcement of regulations under the Clean Air Act is based on the specific days during which a source is considered to be in violation of applicable regulations. Thus, the daily calibration checks for CEM systems provide the critical determination of validity for each day that data are recorded.

As discussed in Chapter 2, the General Provisions of the Part 60 new source performance standards include requirements to conduct daily zero and calibration drift checks for all CEM systems. Specifically, 60.13 states,

> "(d) (1) Owners and operators of all continuous emission monitoring systems installed in accordance with the provisions of this part shall check the zero (or low-level value between 0 and 20 percent of span value) and span (50 to 100 percent of span value) calibration drifts at least once daily in accordance with a written procedure. The zero and span shall be adjusted whenever the 24-hour zero drift or 24-hour span drift exceeds two times the limits of the applicable performance specifications in Appendix B. The system must allow the amount of excess zero and span drift measured at 24-hour interval checks to be recorded and quantified, whenever specified. ..."

Similar requirements are found in the general provisions for Part 61 and Part 63 NESHAP requirements. Most state regulations also contain similar provisions.

For monitors that accept calibration gases, the daily calibration checks usually are performed by conducting system calibrations using zero and mid- or high-range calibration gases. Using an upscale calibration value that approximates either the pollutant concentrations or the level corresponding to the emission standard is preferred. Most, but not all, regulations require that the calibration gases for extractive systems are introduced at or "as near as practical" to the probe outlet so that as much as possible of the sampling system is checked. Some regulations specify that only gases meeting EPA Protocol 1 (USEPA, 1996a) be used for daily calibrations. Other regulations do not specify a traceability protocol or other standard for calibration gases. Such approaches emphasize the use of the daily checks as a stability test and rely on initial and periodic accuracy tests to ensure that monitoring data are accurate.

For in-situ analyzers, daily calibration checks typically include zero checks and upscale checks using the appropriate signals, filters, cells, or other calibration jig assemblies. The approaches that are used are quite diverse depending on the type of instrument to be calibrated. The approaches used for conducting daily calibration checks may rest on a number of important assumptions or conditions which may not be known or well understood by the user. In practice, these assumptions may become invalid because conditions change that result in daily calibration procedures that do not provide results which are representative of monitor performance. Therefore, a thorough understanding of the daily calibration check procedure and its limitations is important to interpret results correctly.

Additionally, many of the approaches that have been used in the past for daily checks of in-situ analyzers do not evaluate important functional parts of the monitoring instrumentation. As previously described, temperature compensation circuitry is often disabled when checking in-situ monitors using cali-bration gas cells within the analyzer. In this case, errors arising within the temperature compensation circuitry (which can significantly bias the CEM system measurement results) are likely to go undetected. A thorough understanding of the daily cali-bration check procedure is needed to interpret results correctly.

An important aspect of the §60.13 (d)(1) regulation is that it requires a written procedure to be developed and followed for performing the daily calibration checks. (For certain sources, this procedure is included in the quality assurance plan.) This written procedure should take into account the important details regarding the calibration of the specific monitor. For extractive systems, it should indicate where the gases are introduced, how the values of check gases are established, provisions to ensure that the gases are injected at the proper flow rate and pressure, etc. For in-situ monitors, the procedure should specify exactly what is done and it should

contain the important assumptions or other conditions that must be maintained for operation of the instrument. Alternatively, the procedure should prescribe the appropriate actions to take when unacceptable performance is observed.

Most importantly, the written procedure provides a standard operating practice to be followed during the initial performance test and each day that emissions data are recorded. In this way, initial and periodic accuracy tests serve to verify the adequacy of the calibration procedure. Implementing the daily calibration procedure verifies the data each day until the next accuracy test is performed. The daily calibration check procedure must conform with applicable regulatory requirements and must also accommodate the technical and practical constraints of the sampling and analysis system as described previously in Section 7.2.

For both extractive and in-situ CEM systems, daily calibration checks are performed and adjustments to the CEM system are required if the indicated drift exceeds specified control limits. When such adjustments are made, or when other corrective action is undertaken, a daily calibration check must be performed after these activities are completed to demonstrate that the monitor has been returned to service properly. Records of adjustments, corrective actions, and the results of daily drift tests are required by virtually all regulatory monitoring requirements.

7.3.3.2 Quality Assurance/Quality Control

40 CFR 60, Appendix F, Procedure 1 contains QA procedures for CEM systems used to demonstrate compliance with emission standards. Although non-criteria pollutant monitoring is not specified as the compliance demonstration method in Part 60 regulations, many permits and state regulations applicable to such monitors adopt Appendix F, Procedure 1 or very similar provisions. The quality assurance procedures for CEM systems in the acid rain program are contained in Part 75, Appendix B and are similar in principal to the Part 60 requirements. Part 75 requirements require that all monitors be able to analyze calibration gases.

Appendix F, Procedure 1 requires that sources develop and implement a QC program with written procedures that describe in detail the complete, step-by-step procedures for calibration of CEM systems, calibration drift determination and adjustment procedures, preventative maintenance, data recording calculations and reporting, accuracy audit procedures, and corrective action procedures for malfunctioning CEM systems. The results of the daily drift determinations are used within Procedure 1 to determine if the CEM system is in control. Procedure 1 specifies that a monitor system is out-of-control if the zero or high-level response exceeds either 1) twice the Appendix B calibration drift performance specification limit for five consecutive days, or 2) four times the Appendix B calibration drift performance specification limit on any one day.

Appendix F, Procedure 1 requires that sources conduct accuracy audits once each calendar quarter. For systems that can use calibration gases, the accuracy audits may include cylinder gas audits conducted during three of four calendar quarters and a relative accuracy test conducted during the other quarter. Alternatively, relative accuracy audits may be performed instead of cylinder gas audits for three quarters. A relative accuracy audit is simply a three-run relative accuracy test evaluating only the mean difference between the CEM system and the RM. The relative accuracy test required by Appendix F, Procedure 1 is identical to the test required during the initial performance specification (i.e., “certification test”).

One can see from the above, which is confirmed by experience, that the selection of a CEM system is constrained by the applicable QA requirements. The cost and inconvenience of conducting three cylinder gas audits per year is much less than the cost and inconvenience of conducting three relative accuracy audits per year. The majority of criteria pollutant CEM systems incapable of analyzing calibration gases were excluded from compliance monitoring applications based on this requirement.

7.3.4 CEM System Approval Mechanisms in Germany

Other approval mechanisms have been used internationally that may serve as models or may contain elements that are useful for certification of non-criteria CEM systems. For example, Germany has implemented a rigorous and comprehensive approval process for many years for CEM systems installed for regulatory purposes. Frequently when monitoring equipment used in Europe is first offered for sale in the United States, potential suppliers will claim that the measurement system is “TÜV certified.” Various control agencies have indicated interest in these certification procedures and they may be useful in the evaluation of non-criteria pollutant monitors in the absence of EPA approval procedures. The German approach to monitor approval and TÜV certification often are misunderstood and sometimes misrepresented in the United States. For these reasons,

information regarding CEM system QA requirements in Germany are described below.

Monitoring equipment must first successfully complete certification testing by the TÜV before it can be sold for use in regulatory programs. This is a rigorous technical evaluation of each equipment model to ensure that it is suitable for particular applications. In addition, testing and evaluation is performed for each monitor at the time that it is installed and this testing is repeated every three to five years depending on the type and size of the facility. This initial calibration program for monitors in Germany is similar to performance specification tests for monitors in the United States. Additional tests are performed by an independent authority on an annual basis to verify performance of the monitoring equipment. Finally, maintenance procedures are performed by the facility's personnel on an ongoing basis in accordance with the manufacturer's instructions and results from the suitability test.

7.3.4.1 Initial TÜV Certification (Suitability) Tests

Each instrument model must first pass a Technischer Überwachungs-Verein (TÜV) certification before it may be offered for sale in regulatory monitoring applications. The TÜV certification is a rigorous evaluation requiring at least three to six months to complete and includes both laboratory evaluations and field suitability testing. The fee for a TÜV certification test is substantial (typically $50,000 to $100,000) and is paid for by the monitor manufacturer. Repeating portions or all of the certification test increases the cost to the manufacturer. Basic TÜV monitor performance specifications are listed in the Table 7-2. The reproducibility specification demands that two randomly selected instruments provide readings with a mean difference of less than 3.3% over a three-month period. The reproducibility specification also ensures that instruments of the same model number are interchangeable. Zero and calibration drift are limited to ± 2% over the maintenance interval. The instrument must provide an alarm indicating the need for immediate maintenance when zero or calibration drift limits are reached. The time needed to reach these limits defines the maintenance interval. The maintenance interval is recommended by the manufacturer but is determined by the TÜV as part of the suitability tests. The maintenance interval should be at least one week; however, four weeks is preferable. Determination of the maintenance interval based on suitability tests and drift limits (relative to the maintenance interval) provides incentives for manufacturers to build stable instruments.

Tests to determine the effects of line voltage variations, ambient temperature variations, and other factors that may influence monitor performance are also performed as part of the TÜV certification. The TÜV may investigate other areas or issues that are considered to be relevant to the performance of a particular measurement system. Field suitability tests are performed for each type of application (e.g., gas-, oil-, coal-, refuse-fired boilers) but in some cases, success at more difficult monitoring applications is taken as a sufficient demonstration for less demanding applications.

7.3.4.2 Initial and Periodic Monitor Calibration Tests

These tests are in many ways similar to CEM system performance specification tests performed in the

Table 7-2. Principal Performance Specifications for TÜV Suitability Tests of Emission Monitoring Instruments

System Parameter	Criteria
Analytical Function (ma vs. concentration)	By Reference Method (regression analysis)
Detection Limit	2% of most sensitive range
Reproducibility	R = measurement range ≥ 30 mean difference
Zero and Calibration Drift	± 2% in maintenance interval
Availability	90% in three months minimum; 95% is the goal
Interference from other species	± 4% of full scale
Maintenance interval	Determined by test program (limited to three months maximum)

United States. However, a number of important differences exist in both the philosophy, nomenclature, and procedures used for these tests. The term "calibration" as used in VDI 3950, Part 1, "Calibration of Automatic Emission Measuring Instru-ments" refers to the entire process of establishing and verifying the correct performance of a monitoring system (VDI 3950). This includes: 1) procedures for the selection and preparation of measurement sites, 2) procedures to check for proper installation of equipment, 3) leak checks of extractive monitoring systems, 4) a five-point verification of the instrument characteristic (similar to calibration error in the United States using zero and upscale calibration gases), 5) monitor-specific procedures for evaluating measurement interferences, 6) zero and reference (upscale calibration) drift stability tests, 7) response time tests, 8) procedures for measuring and processing data, 9) procedures for checking the representativeness of sampling points, 10) procedures for determining the analytical function for the complete measuring system relative to independent flue gas measurements (reference methods), and 11) procedures for analysis and reporting of the calibration results.

In the United States, many of the calibration issues that are specifically addressed by VDI 3950, Part 1 are either lumped together and evaluated as a group during the relative accuracy test or are simply ignored. For example, the German procedures evaluate separately: 1) the representativeness of the sampling points, 2) the influence of interferences, and 3) the system performance relative to independent emission measurements through a series of tests. In the United States, we rely on a single relative accuracy test to determine the acceptability of all of these factors taken together. The German "calibration" procedure also relates test procedures and results back to the initial TÜV certification test. For example, the calibration interference test procedures specifically address the measurement interferences identified during the initial certification because this is obviously dependent on the specific monitor design and analytical principal. Also, the zero and reference point drift determinations are determined relative to the maintenance interval established during the initial certification test rather than an arbitrary time interval.

The German comparison with independent reference methods relies on a minimum of fifteen paired measurements (CEM and reference method) conducted at different emission levels to facilitate a regression analysis that is used to determine subsequent emission levels. Linear or quadratic regression analysis may be used and calculations of the confidence coefficient and tolerance intervals are included. These statistical quantities also are employed in the interpretation of the monitoring data.

The field calibration test is performed by an independent expert agency (the TÜV) authorized by the German government to perform these evaluations. However, the TÜV office that performs the calibration test at a particular industrial facility is not necessarily the TÜV office that performed the initial certification. Again, the field calibration procedure is performed initially and then is repeated every three to five years depending on the type and size of the facility.

7.3.4.3 Annual Calibration Tests and Maintenance

Section 8, "Periodical Functional Test" of VDI 3950, Part 1 describes the annual evaluation requirements. These include: 1) checks of the operational status (leak checks, optical contamination, etc.), 2) records review of zero and reference point checks, 3) zero and reference drift tests, 4) checks for interferences, 5) multi-point verification of the instrument characteristic (calibration error test) using zero and upscale calibration gases or other calibration materials, 6) certain monitor functional tests, 7) a minimum 4-run comparison with independent reference methods, and 8) inspection of the data transmission to the chart recorder, integration device, or data logging system. These are more extensive evaluation procedures than are required for monitors subject to EPA regulations. These procedures also are performed by an independent licensed agency (TÜV) rather than by source personnel. Maintenance is performed by the industrial facility's trained personnel. Minimum procedures are specified for in-situ devices and for extractive systems.

7.3.5 International Standards Organization

International Standards Organization (ISO) has developed standards for certain monitoring applications. ISO Standard 7935 is for SO_2 CEM systems and the standard for NO_x CEM systems is currently in draft form (ISO, 1991). A discussion of these standards for criteria pollutant monitoring and a comparison with United States regulations is presented by Jahnke (Jahnke, 1993). Other ISO standards for emissions test methods are under development.

The ISO committee TC146/SC1/WG1 has prepared ISO standard 10155 "Stationary Source Emissions, Automated Monitoring of Mass Concentration of Particles - Performance Characteristics, Test Procedures, and Specification". The standard, prepared over the last 10 years, was published in its final form on April 1, 1995 (ISO, 1995). It does not prescribe a particular method or analytical technique but instead reflects a general approach and provides performance specifica-

tions to evaluate any specific method that is offered. Fundamental to the ISO standard is the requirement to perform a site-specific calibration or correlation to ISO manual method 9096 (which is similar to EPA isokinetic particulate sampling methods) (ISO, 1991). As should be expected, ISO 10155 specifies that the calibration procedure must be repeated when changes in emission controls, fuel type, or other factors occur that may be expected to influence the calibration. Although this is a general requirement, it reflects the fact that the need for recalibration is best based on practical judgment.

ISO 10155 prescribes several performance specifications including the following:

- Response time of less than 1/10 of the manual sampling time
- Zero and span drift less than ± 2% of working range per month
- Accumulated automatic zero and span adjustments less than 6% of working range
- Calibration (correlation) line specification
  1. Correlation coefficient ≥0.95
  2. 95% confidence interval shall be ≤10% of emission standard
  3. Tolerance interval; 95% confidence that all possible values are within ± 25% of emission standard

The ISO 10155 standard requires that sample runs be performed at three different emission levels to establish the calibration. A minimum of 9 sampling runs must be performed but 12 or more runs typically are expected. The ISO standard prescribes that the <u>process operating conditions</u> should be varied if possible to create the different emission levels. Where this is not possible, variation of the control equipment operating parameters to create a range of emission levels is accepted.

The Central European Normalization Committee (CEN committee TC264/WG5) is developing requirements applicable to continuous particulate monitoring (Peeler, 1996, b). Unlike ISO, which is a voluntary organization of participating countries, CEN mandates requirements for the 14 countries comprising the European Community. The CEN committee will establish: 1) emission standards and limits, 2) manual test methods, and 3) automated monitoring methods. Each participating country must adhere to the CEN requirements or adopt more restrictive/more rigorous requirements. CEN typically adopts ISO standards/ methods where available. The CEN committee has adopted the ISO 10155 continuous particulate monitoring standard and applied it as a requirement for hazardous waste incinerators.

The calibration line specifications are illustrated in Figure 7-6.

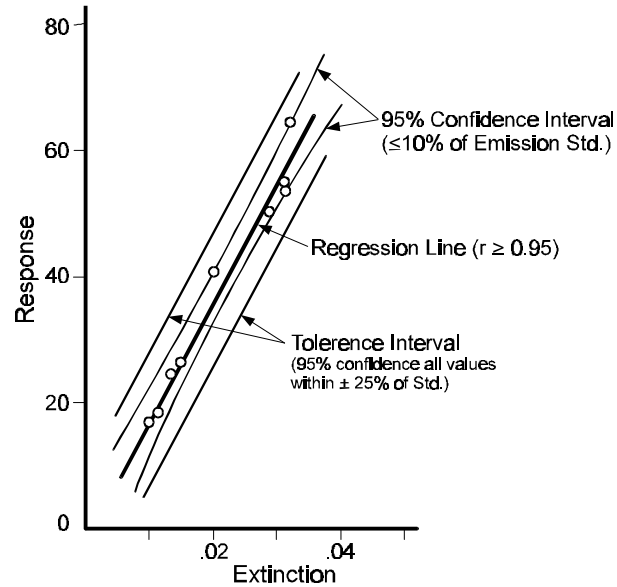

Figure 7-6. Continuous particulate monitoring calibration line specifications.

7.4 Suggested Approval Mechanisms and Approaches for Non-criteria Pollutant and Application Testing

The requirements and procedures discussed in Sections 7.3.1-7.3.3 for conducting initial performance specification tests and audits are consistent with existing regulatory programs and can be used for many non-criteria pollutant monitoring applications. However, some of these procedures do not address problems fully that may be encountered in evaluating non-criteria CEM systems. New evaluation procedures likely will be needed as technology continues to evolve and as monitoring applications for additional specific non-criteria pollutants expand through regulations, permit requirements, or market based trading programs. Several alternative approval mechanisms that may be applicable in these situations are described in this section.

7.4.1 EPA Method 301

Method 301 of Part 63 (USEPA, 1996a) is an approval process designed for industrial sources to use in the development of alternate test methods in the absence of other validated methods, or in cases where an alternate method is desirable. Method 301 initially

was developed to allow a specific source to demonstrate the validity of a hazardous air pollutant test method for that source to qualify for an "early reduction" option. Method 301 was intended to encourage development of methods where methods did not already exist. The practical application of Method 301 has been expanded greatly beyond its original purpose; it now serves as a model for evaluating many measurement methods for both criteria and non-criteria air pollutants.

Method 301 incorporates procedures to determine if the precision and bias of a candidate method are acceptable based on comparisons with a validated method, use of isotopic spiking, or analyte spiking. Method 301 is directed primarily at traditional source testing approaches that include discrete sampling and analysis phases. The implementation of the procedure requires the use of quad-trains or paired sampling trains to evaluate precision and bias. Alternative procedures have been developed and accepted by the EPA to allow the application of Method 301 to direct interface methods (methods where the effluent is directly injected into the analyzer) such as instrumental methods using FTIR and GCMS analytical techniques (USEPA, 1994). These alternative Method 301 procedures include collecting a series of spiked and unspiked effluent data for calculation of the precision and bias of a method. These analyte spiking procedures have been used to demonstrate the acceptability of various test methods used by industry in MACT standard development programs (Kinner, 1996, LaCoss, 1995).

Note: Method 301 is only a source specific test method determination. Method 301, Section 12 includes a general discussion regarding “conditional approval” of a method which might allow the transfer of results to additional sources. Conditional approval waiver requirements include: method documentation, ruggedness tests, sample stability, and practical limits of quantification. EPA has granted test method approval to groups of sources in several circumstances. A voluntary administrative procedure documenting the applicant’s and EPA responsibilities for instrumental field test methods has been proposed by an instrument vendor, and is being evaluated for a direct interface gas chromatograph mass spectrometer method (Peeler, 1996).

As noted above, Method 301 can be adapted to evaluate direct interface or other instrumental test methods. Such methods are similar to CEM systems in that successive samples are acquired and analyzed. However, instrumental test methods generally involve a much more detailed protocol and many more specific procedures than are associated with a CEM system. Also, an instrumental test method usually requires the full time attendance of an operator as compared to a CEM system, which is designed to function automatically with very little human intervention. Because of these differences, many modifications to Method 301 are necessary to apply it to the evaluation of CEM systems.

7.4.2 Dynamic Analyte Spiking

Dynamic analyte spiking involves the quantitative introduction of a calibration gas of known concentration containing one or more analytes, to an effluent sample stream. The spike gas represents only a small fraction of the combined stream. Therefore, the spiked sample stream contains essentially the same constituents at the same concentrations as the unspiked samples. A comparison of spiked and unspiked samples provides an evaluation of both the effectiveness of the sampling system and the performance of an analytical system.

Because of the similarity of the spiked and unspiked samples, “matrix effects” (interference effects due to the presence of other sample constituents such as moisture, oxygen, carbon dioxide, etc.) can be assessed by this procedure. The analyte spiking approach is particularly useful in the evaluation of non-criteria pollutants that are reactive, condensable, water soluble, or that have the tendency to polymerize in the sampling system. These types of compounds represent the most difficult measurement challenges and can invalidate traditional approaches for assessing monitor performance.

Selection of the spike analytes depends upon two items: the number of target analytes, and the ability to obtain these compounds in gaseous form. For multi-component CEM systems, analyte spiking with all of the target analytes may not be economically feasible. In this case, some surrogates that represent the analytes must be chosen. Selection of surrogates depends on their ability to represent more than one of the target analytes based on similar physical properties.

Dynamic analyte spiking is a more rigorous evaluation procedure than the traditional system calibration. Simply introducing dry calibration standards in a system calibration is useful to check for leaks, adsorption/desorption in the sampling system and to check the analyzer calibration. However, even the straightforward absorption of analytes by condensate formed from sample moisture cannot be detected in the system calibration because dry calibration gases are used. Similarly, the system calibration cannot detect

analytical interferences due to analyte interaction with other components present in the effluent samples because the calibration gas does not usually contain these components.

By performing dynamic analyte spiking, the effectiveness of the sample delivery system and the analytical components are assessed more thoroughly for the compounds of interest in the presence of actual source matrix interferences and moisture. This procedure should be performed in conjunction with both direct and system calibrations. Together these procedures provide a check of the CEM system both on a dry basis and under actual sampling conditions.

7.4.2.1 Analyte Spike Procedure

Any CEM system that has the capability to accept calibration gases at a point upstream of the particulate filter is a candidate for analyte spiking. The spike gas is delivered into the sampling system, typically in a ratio of 1 part spike to 9 parts sample gas, resulting in a ten-fold dilution of the spike gas. The spike should not exceed this 1:10 ratio to avoid excessive dilution of the analytes of interest or substantially changing the sample matrix. Figure 7-7 is a generalized schematic of the analyte spiking technique.

The spike gas should be preheated to prevent localized condensation at the point of injection. The spike must be delivered to the sampling system at a point upstream of the particulate matter filter to detect possible gas reactions with the accumulated particulate material on the filter.

Calibrated mass flow meters or controllers are necessary to deliver the spike at a precise, measured flow rate. A calibrated rotometer for measurement of total flow is also necessary. The rotometer should be installed in the sample delivery system at a point that provides an accurate measurement of total sample flow. Experimental errors associated with calculating the dilution factor of the analyte spike arise from the inaccuracy of the spike and total sample flow measurement devices (see Appendix G, Equation G-2). In addition, calibration of the total flow measurement device with wet stack gas can be difficult or impractical. In addition, the effects of errors associated with the analyte spike calibration standard values are also magnified.

As an alternative to the above, the measurement of the spike dilution factor may be determined directly through measurements of a tracer compound contained in the analyte spike gas. A tracer must be chemically inert and not present in the source effluent. Sulfur hexafluoride (SF_6) has been used successfully as a tracer in extractive FTIR test methods, and is included in draft Methods 318 and 320. (Method 318 may be proposed as an FTIR test method for inclusion in 40 CFR Part 63.) The amount of spike gas introduced during a spiking experiment can be determined from the concentrations of the analyte and tracer components of the calibration standard. (See Appendix G). Errors associated with use of the tracer technique are associated with the analyzer's ability to measure the tracer gas concentration accurately and errors associated with the analyte and tracer concentrations in the spike gas standard.

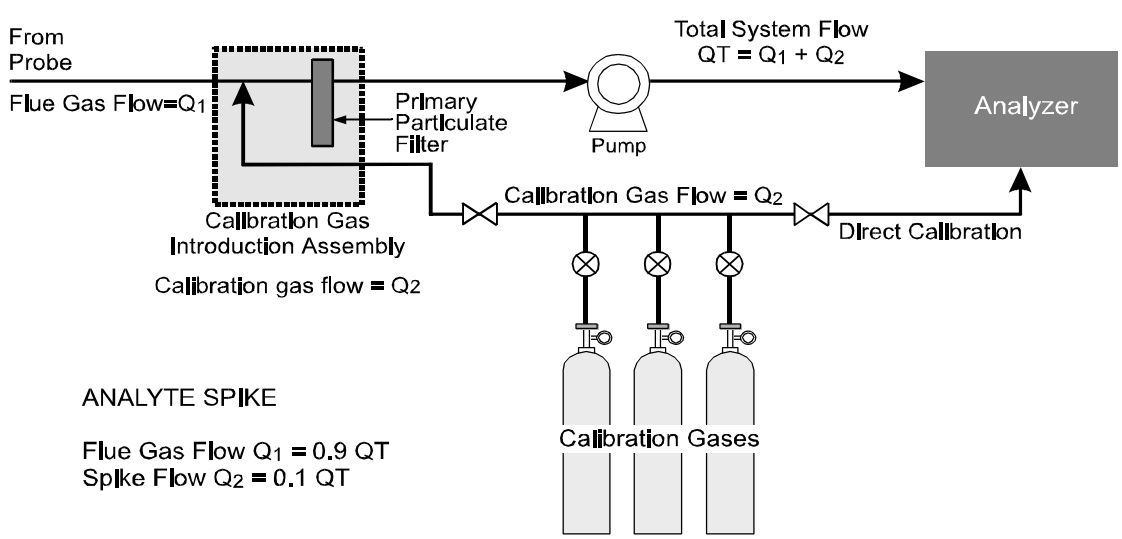

Figure 7-7. Analyte spiking.

Acceptable measurement system performance should be based upon the intended use of the data. However, analyte spike recovery values of ±30% of the expected value are typical. These limits are consistent with the Method 301 bias limits and have been applied in determining the acceptability of direct interface instrumental test methods (Powell, 1996). Analyte spike recoveries and their calculations are discussed in further detail in Appendix G.

REFERENCES

Brown, Thomas D. 1996. "ND Energy and Environmental Research Center for Air Toxic Metals," Newsletter, Volume 2, Issue 2.

Cone, L.A. 1989. "Carbon Monoxide and Total Hydrocarbon Continuous Monitoring at Hazardous Waste Incineration Facilities," *Proceedings Air & Waste Management Association CEM Specialty Conference,* Chicago, IL, pp. 338-349.

International Standards Organization. 1991. *Stationary Source Emissions Determination of Mass Concentration of Sulfur Oxides - Performance Characteristics of Automated Measurement Systems*. ISO 7935. (Available from American National Standards Institute, New York, NY.)

International Standards Organization. 1995. *Automated Monitoring of Mass Concentration of Particles in Stationary Source Emissions: Performance Characteristics, Test Procedures and Specifications.* ISO 10155. (Available from American National Standards Institute, New York, NY.)

Jahnke J.A. 1994. *An Operators Guide to Eliminating Bias in CEM Systems*, EPA/400/R94/016, Washington, D.C., U.S. Environmental Protection Agency.

Jahnke J.A. 1993. "In-situ Monitoring Systems for the Measurement of Gas Concentrations and Flue Gas Velocities," in *Continuous Emission Monitoring*, Van Nostrand Reinhold, NY, pp 156-199.

Kinner, L.L., Peeler, J.W. 1996. "Development of Fourier Transform Infrared Spectrometry and Gas Filter Correlation Infrared Spectrometry Test Method Protocols for the Portland Cement Industry," *Proceedings Air & Waste Management Association*, Nashville, TN, 96-TA37.03.

Kinner, L.L., Peeler, J.W. 1995. "Protocols for Measurement of Gaseous Volatile Organics and HCl from Cement Kilns by FTIR and GFCIR," Available from the Portland Cement Association, Skokie, IL.

Kinner, L.L., Geyer, T.J., Plummer, G.M., Dunder, T.A. 1994. "Application of FTIR as a Continuous Emission Monitoring System," *Proceedings Air & Waste Management Association Specialty Conference on International Incineration,* Houston, TX, pp. 178-196.

LaCoss, J., McCarthy, J. 1995. "Measurement of Air Toxics Using Extractive FTIR Spectroscopy," *Proceedings Emerging Technologies for Hazardous Waste Management, Special Symposium of Industrial and Engineering Chemistry, Division of American Chemical Society*, Atlanta, GA, paper no. 194.3.

Laudal et. al. 1996. "Mercury Speciation: A Comparison between EPA Method 29 and Other Sampling Methods," *Proceedings Air & Waste Management Association,* Nashville, TN, paper no. 96-WA64A.04.

NIOSH. 1994. *NIOSH Manual of Analytical Methods, 4th Edition.* U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Division of Physical Sciences and Engineering, Cincinnati, Ohio, August, 1994.

Peeler, J.W., Kinner, L.L., DeLuca, S. 1996. "General Field Test Method Approval Process and Specific Application for a Direct Interface GCMS Source Test Method," *Proceedings Air & Waste Management Association*, Nashville, TN, paper no. 96-RP132.01.

Peeler, J.W., Jahnke, J.A., and Wisker, S.M. 1996. Continuous Particulate Monitoring in Germany and Europe Using Optical Techniques. *Continuous Compliance Monitoring Under the Clean Air Act Amendments.* Air & Waste Management Association. Pittsburgh, PA, pp. 208-220.

Powell, J.H., Dithrich, E.C. 1996. "Hot-Wet Instrumental Hydrogen Chloride Emissions Quantification Using GFCIR Method Validation and Comparison," *Proceedings Air & Waste Management Association Specialty Conference on Boilers and Industrial Furnaces*, Kansas City, MO.

Richards, Marta K., et. al. 1996. "Performance Tests of Mercury Continuous Emissions Monitors at the U.S. EPA Incineration Research Facility," *Proceedings Air & Waste Management Association*, Nashville, TN, paper no. 96-WA64A.02.

Roberson, R. 1996. "Status of CEM Systems for HAP Emissions," *Proceedings Electric Power Research Institute CEM Users Group*, Kansas City, MO.

Serne, J.C., White, M.O., Burdette, J.W. 1993. "Recent Experience Measuring Emissions of Aldehydes and Ketones Using EPA Method 001," *Proceedings Air & Waste Management Association Waste Combustion in Boilers and Industrial Furnaces Specialty Conference*, Clearwater, FL.

U.S. Environmental Protection Agency. 1977. *Traceability Protocol for Establishing True Concentrations of Gases Used for Calibration and Audits of Continuous Source Emissions Monitors (Protocol Number 1),* June 1978. Section 3.0.4 of the Quality Assurance Handbook for Air Pollution Measurement Systems. Volume III. Stationary Source Specific Methods. EPA 600/4-77-027b. U.S. EPA, Office of Research and Development Publications, 26 M.L. King Dr., Cincinnati, OH 45268.

U.S. Environmental Protection Agency. 1981. *A Procedure for Establishing Traceability of Gas Mixtures to Certain National Bureau of Standards Reference Materials*. Joint publication NBS/EPA-600/7/81-010. Available from U.S. EPA, Quality Assurance Division (MD-77), Research Triangle Park, NC.

U.S. Environmental Protection Agency. 1986. *Test Methods for Evaluating Solid Waste, Physical/Chemical Methods, SW-846 Manual, 3rd, ed.* Document no. 955-001-0000001. Available from Superintendent of Documents, U.S. Government Printing Office, Washington, D.C.

U.S. Environmental Protection Agency. 1993. *Field Validation Testing at a Coal Fired Boiler,* U.S. EPA Report, EPA Contract No. 68D20163, WA No. 2., Research Triangle Park, NC.

U.S. Environmental Protection Agency. 1994. *Field Validation Test Using FTIR Spectrometry to Measure Formaldehyde, Phenol, and Methanol at a Wool Fiberglass Production Facility,* United States Environmental Protection Agency Report, EPA Contract No. 68D20163, WA No. 32.

U.S. Environmental Protection Agency. 1995. *Protocol for the Use of Extractive Fourier Transform Infrared (FTIR) Spectrometry in Analysis of Gaseous Emissions from Stationary Industrial Sources*, United States Environmental Protection Agency, Research Triangle Park, NC.

U.S. Environmental Protection Agency. 1996a. Code of Federal Regulations. *Continuous Emission Monitoring, Appendix H to Part 75 - Revised Traceability Protocol No. 1.* 40 CFR 75 Appendix H. Washington, D.C.

U.S. Environmental Protection Agency. 1996b. Code of Federal Regulations. Method 318, *Extractive FTIR Method for the Measurement of Emissions from the Mineral Wool and Wool Fiberglass Industries*, 40 CFR 63 Appendix A., Washington, D.C.

U.S. Environmental Protection Agency. 1996c. *Revised Standards for Hazardous Waste Combustors*, Proposed Rule, 61 Federal Register 17358.

U.S. Environmental Protection Agency. 1996d. Code of Federal Regulations. Method 301, *Field Validation of Pollutant Measurement Methods from Various Waste Media,* 40 CFR, Part 63 Appendix A., Washington, D.C.

VDI Guideline 3950. 1994. "Calibration of Automatic Emission Measuring Instruments," Part 1. Available from Beuth Verlag Gmbh, 10772, Berlin.