

# **CHAPTER 1**

## **OVERVIEW: ACCURACY, PRECISION, AND BIAS IN CONTINUOUS EMISSION MONITORING SYSTEMS**

# **CHAPTER 1**

## **OVERVIEW: ACCURACY, PRECISION, AND BIAS IN CONTINUOUS EMISSION MONITORING SYSTEMS**

### **1.1 BACKGROUND**

Public concern with the environmental impact of acid rain resulted in Title IV of the Clean Air Act Amendments of 1990 which established emission standards for sulfur dioxide (SO<sub>2</sub>) and nitrogen oxides (NO<sub>x</sub>), the primary pollutants causing acid rain. To ensure that the emission standards were met, Title IV required continuous emission monitoring (CEM) systems to be put into operation at all affected utilities to measure SO<sub>2</sub> and NO<sub>x</sub> as well as carbon dioxide (CO<sub>2</sub>), diluent gases (CO<sub>2</sub> or oxygen, O<sub>2</sub>), flue gas velocity, and opacity.

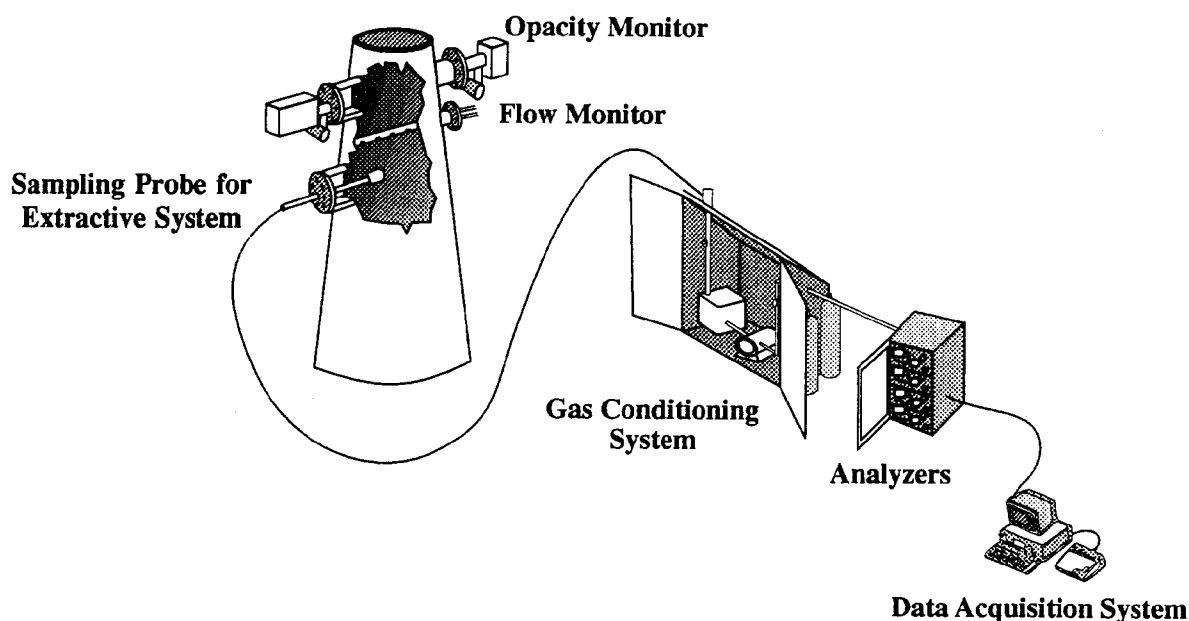
To limit the levels of SO<sub>2</sub> emitted, each source covered under Title IV is allotted a prescribed number of allowances, an allowance being the right to emit one ton of SO<sub>2</sub> per year. Because the total number of allowances issued by the U.S. Environmental Protection Agency (EPA) is strictly limited to the cap established in Title IV, the allowance allocation process provides the means to control SO<sub>2</sub> emissions and, consequently, acid rain.

Each year, the electric utilities are required to reconcile their total SO<sub>2</sub> emissions against the allowances held. The CEM systems, specified by Title IV, are instrumental accountants for the Acid Rain Program. Not only do they measure emissions, but they also allow utilities to track the consumption of allowances. In so doing, they provide the foundation for this extensive regulatory program.

A CEM system's continuous accounting of emissions allows the utility operator to determine the number of allowances used, the number available for the rest of the year, and the number that need to be acquired to operate for the remainder of the year. Because allowances have monetary value and can be bought, traded, auctioned, and otherwise transferred, it is imperative that CEM system data be accurate. Loss of allowances due to over-representation of emissions or inaccurate CEM systems are a concern to the utility. Under-reporting of emissions due to inaccurate systems are of concern to EPA. This document addresses such concerns by providing guidelines for obtaining accurate, unbiased CEM system data.

### **1.2 CEM SYSTEMS AND CERTIFICATION**

A CEM system is composed of a number of subsystems: a gas monitoring system (which may use either extractive or in-situ sampling techniques and may include either a CO<sub>2</sub> or O<sub>2</sub> diluent correction monitor), a flow monitor, a transmissometer (opacity monitor), and a data acquisition and handling system (DAHS). An extractive system consists of a number of subsystems—the probe and conditioning systems and analyzers. A typical CEM system is shown in Figure 1-1.



**Figure 1-1. A Typical Continuous Emission Monitoring System**

All of these components and subsystems work in concert to provide emissions data. There are, of course, many monitoring options (Jahnke, 1993). For example, in systems that extract gas from the stack, the gas can be cooled and the moisture removed or, alternatively, kept at an elevated temperature above the dewpoint and measured on a wet basis. Instead of measuring the extracted gas directly, it can first be diluted and measured using ambient air analyzers. Another option is to monitor the flue gas in-situ (i.e., directly in the stack or duct), without extraction.

The opacity and flow monitors shown in Figure 1-1 are in-situ monitors—the flue gas is monitored in-place and is not disturbed. The flow monitor is used here, in conjunction with gas concentration measurements, to calculate mass emission rates (i.e., in units of lbs/hr and tons/yr). The transmissometer monitors the flue gas opacity, which indirectly characterizes particulate matter emissions.

Although there are many types of systems, there is no one best system for all applications. CEM systems are application dependent. Regulatory conditions, stack gas composition, environmental and physical conditions, and even management practices can make one system better suited than another for a given application.

### **1.2.1 Performance-Based Standards**

A CEM system is proven through its performance. If the installed system can meet established performance criteria, such as the standards for linearity, calibration drift, and accuracy, it can be approved for use as a regulatory continuous monitoring system. The U.S. EPA, the

International Standards Organization (ISO, 1989), and many European countries have adopted performance-based standards rather than design-based standards. In other words, the standard is not how the system is designed, but whether it works after it has been installed.

The U.S. EPA has established several sets of CEM performance specifications. These can be found in Title 40 of the U.S. Code of Federal Regulations (40 CFR) in Part 60 for New Sources, in Part 266 for facilities that burn hazardous waste, and in Part 75 for sources affected by the Acid Rain Program. Although these specifications are all similar, having evolved from the original Part 60 requirements, the 40 CFR 75 specifications are the most comprehensive and stringent. Because of the central role that Part 75 CEM systems play in the effective functioning of the Acid Rain Program allowance market, the CEM data must be as accurate and precise as possible.

### **1.2.2 Relative Accuracy Test Audit**

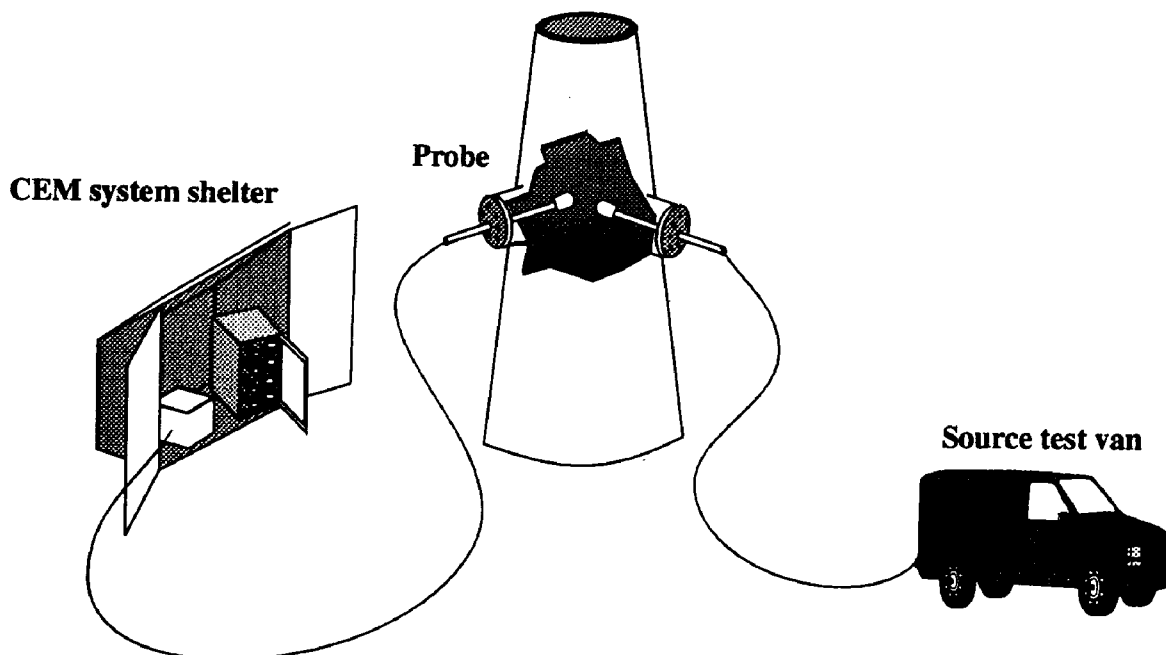
A principal performance testing procedure for Acid Rain CEM systems is the relative accuracy test audit (RATA). The RATA is a comparative evaluation of the CEM system performance against an independent reference method. A reference method can be either (1) a manual wet chemistry method, where, for example, gas is extracted from the stack and bubbled through an absorbing solution which is then analyzed in a chemical laboratory, or (2) an instrumental method, where gas is extracted from the stack and analyzed directly by suitably calibrated analyzers. Under the Acid Rain Program, the applicable reference methods are Method 2 (reference method for determination of stack gas velocity and volumetric flow), Method 6 (manual reference method for SO<sub>2</sub>) or Method 6C (instrumental reference method for SO<sub>2</sub>), and Method 7 (manual method for NO<sub>x</sub>) or Method 7E (instrumental method for NO<sub>x</sub>).

Specifications for both the manual and instrumental reference methods are found in 40 CFR 60 Appendix A. Figure 1-2 illustrates a typical RATA, using a monitoring van with automated test equipment.

In a RATA, a minimum of nine sets of paired monitoring system and reference method test data are obtained. A tester may perform more than nine sets of reference method tests and may reject up to three data sets, as long as the total number of runs used in calculating test results is equal to or greater than nine. Data from the RATA are used to determine both the relative accuracy and bias, if any, of a CEM system.

### **1.3 ACCURACY AND BIAS — A CONCEPTUAL VIEW**

Technically, the accuracy of a measurement refers to the degree of agreement between the measured value and a true value. In source measurements, as in physical science in general, the true value of a physical parameter is rarely known. Instead, an "accepted" true value is generally used for comparison against the CEM system measured values. In source testing, the "true" value is assumed to be that value determined by the EPA Reference Method.



**Figure 1-2. Relative Accuracy Test Audit Using Instrumental Reference Methods**

### 1.3.1 Relative Accuracy Test

Relative Accuracy is a regulatory statistic that expresses CEM accuracy in relative terms, i.e., it quantifies the deviation of the CEM from the reference method relative to the emission levels occurring at the time of the RATA. Derived from the paired data measurements (Natrella, 1963) obtained during the RATA, it is expressed as a percentage of the average of the emission levels encountered during the RATA. This calculation is in contrast to most engineering practice, which expresses accuracy as a percentage of span. As such, relative accuracy is closely associated with the source emission levels occurring at the time of the test, rather than with instrument span.

The relative accuracy is calculated using the following expression:

$$RA = \frac{\bar{d} \times \% cc}{\overline{RM}} \times 100 \quad (\text{Eq. 1-1})$$

To calculate  $\bar{d}$ , the absolute value of the mean difference between data pairs, the arithmetic difference between the reference method and the CEM system measurements for each data pair is first calculated:

$$d_i = RM_i - CEM_i \quad (\text{Eq. 1-2})$$

where  $d_i$  is the difference between a reference method value and the corresponding monitor or CEM system value for the  $i^{\text{th}}$  test run.

The mean difference is then calculated using the expression:

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i \quad \text{(Eq. 1-3)}$$

where

$n$  = the number of data pairs.

The absolute value of  $\bar{d}$  is then used in Eq. 1-1. In calculating the sum of the differences between the data pairs, it is important to note that the signs of the differences are retained (that is, the absolute value is taken of the total summation, not the individual  $d_i$  values).

The confidence coefficient is determined from the following expression:

$$cc = t_{0.025} \frac{S_d}{\sqrt{n}} \quad \text{(Eq. 1-4)}$$

where

$t_{0.025}$  = a statistical parameter used to calculate \*cc\* for a given number of data pairs (Table 1-1).

**Table 1-1. t-Values**

n-1	$t_{\alpha=0.025}$
8	2.306
9	2.262
10	2.228
11	2.201
12	2.179
13	2.160
14	2.145

$S_d$  = the standard deviation of the differences of the data pairs obtained during the relative accuracy test.

$$S_d = \left[ \frac{\sum_{j=1}^n (d_j)^2 + \left[ \frac{\sum_{j=1}^n d_j}{n} \right]^2}{n+1} \right]^{\frac{1}{2}} \tag{Eq. 1-5}$$

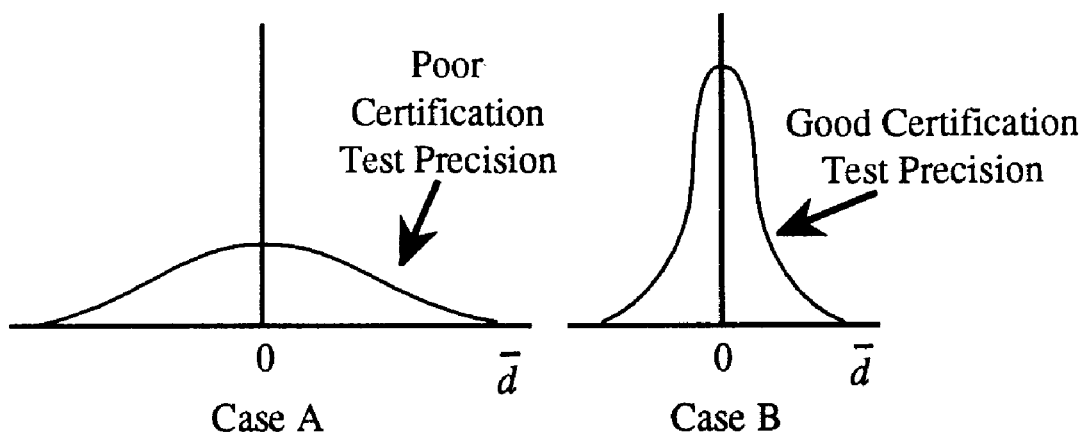
The confidence coefficient is a measurement of the uncertainty in the calculation of  $\bar{d}^*$ : Because the relative accuracy determination is made from a finite set of data, there is a probability that  $\bar{d}$  could be larger or smaller.  $cc^*$  represents the largest variation in  $\bar{d}^*$ , which we would expect to see 97.5% of the time (i.e., it would only be exceeded 2.5% of the time).

$\overline{RM}$ , the term appearing in the denominator of Eq. 1-1, is the arithmetic mean of the reference method values:

$$\overline{RM} = \frac{1}{n} \sum_{j=1}^n RM_j \tag{Eq. 1-6}$$

In Part 75, the relative accuracy, calculated from concentration units (ppm or percent), for SO<sub>2</sub> and CO<sub>2</sub> monitors must be 10% or less. For an NO<sub>x</sub> monitoring system, the relative accuracy must be 10% or less, calculated from units of lbs/mmBtu (ng/Joule) obtained from both NO<sub>x</sub> and diluent (CO<sub>2</sub> or O<sub>2</sub>) measurements. For flow monitors, the relative accuracy, derived from measurements in units of standard cubic feet per hour, must be 15% or less.

Figure 1-3 offers a graphical representation of the underlying frequency distributions inferred from two hypothetical relative accuracy test audits (denoted Case A and Case B). The graphs show the uncertainty about the estimate of mean differences for the two sample RATAs. Each distribution shows the range and variability in the mean difference that can be inferred from the RATA measurements. The horizontal axis displays the mean difference ( $\bar{d}$ ) found using Eq. 1-3, where CEM system measurements are compared to the "accepted" true values determined by the reference method.



**Figure 1-3. An Illustration of Precision in Two Hypothetical Relative Accuracy Test Audits**

For illustrative purposes in both Case A and Case B, the difference between the CEM and reference method on average is assumed to be zero. (This is represented by each distribution being centered at the zero point on the horizontal axis.) Thus, in both situations bias is not a factor. However, the comparative steepness of two distributions reveals striking differences in the precision of the differences between the CEM system and reference method prevailing during the RATA. In Case A, the curve is squat, indicating that the values of  $d_i$  varied appreciably from run to run, to produce a wide variation in  $\bar{d}$ . It was not possible to reproduce the data well. Such a situation could possibly indicate an erratic CEM system, poor reference method testing, or both. In contrast, the curve in Case B is sharp, indicating that the difference between the reference method values and CEM values were nearly the same for each of the nine test runs used to calculate  $\bar{d}$ . The data were reproducible. The instruments displayed a high degree of precision. The squatness and sharpness of the two curves is captured by  $cc^*$  in the numerator of Eq. 1-1.

### 1.3.2 Bias

The relative accuracy test, used in CEM certification and performance testing, captures the degree of relative imprecision in CEM measurements, but it does not differentiate systematic error from random error. Prior to the promulgation of 40 CFR Part 75, the relative accuracy test alone was used to limit both imprecision (random error) and measurement bias (systematic error).

There is a problem, however, in only using the relative accuracy specification. For example, if a CEM systematically reads 9% low relative to the reference method, it could still pass a 10% relative accuracy standard even though the data subsequently reported to the agency would be consistently 9% low. This situation is particularly serious in the Acid Rain Program, because such a possibility would both jeopardize the achievement of the Program's mandated emission limits and undercut the program-wide uniformity of emission measurements, thereby calling into question the true valuation of  $\text{SO}_2$  allowances.



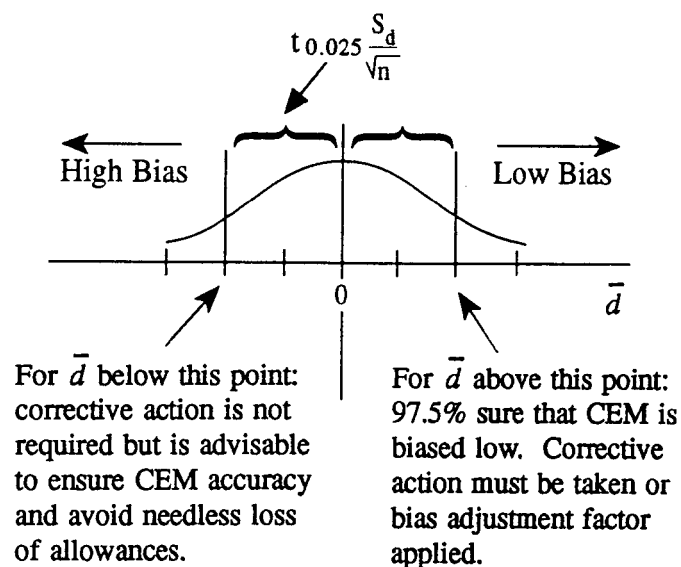
To address this situation, 40 CFR Part 75 tightened the relative accuracy standard to 10% and subjected RATA data to a bias test, specifically designed to detect systematic error.

### 1.3.3 Bias Test

Besides being used to calculate relative accuracy, the paired RATA data are also used to determine if statistically significant systematic error (low bias) is manifested in the CEM measurements. A t-test is applied to the paired differences to test the hypothesis that differences between the CEM and reference method are not statistically different from zero. If the mean difference of the measurements as found in Eq. 1-3 exceeds the confidence coefficient as found in Eq. 1-4, then the hypothesis is rejected. According to well-established principles from classical statistics, if the mean difference exceeds the confidence coefficient then we can be 97.5% confident that the measurement difference was not a random occurrence, i.e., that the difference was due to systematic, not random, error. Thus low bias is considered to be present if

$$\bar{d} > *cc* \quad (\text{Eq. 1-7})$$

This expression merely states that systematic error is considered to be present if on average the CEM measurements are so far below the reference method measurements as to lie outside the confidence limits. That is, they are so low that the  $\bar{d}$  derived from the RATA data falls in a zone where classical statistics predicts with 97.5% confidence that  $\bar{d}_{true}$  will not occur. In other words, the CEM system is reading so low relative to the reference method that we are 97.5% confident that the system is biased low (Figure 1-4).



**Figure 1-4. t Distribution for a RATA, Showing when Bias Occurs**

Equation 1-7 is basically an expression of the one-tailed t-test. By using it, there is at most a 2.5% probability of mistakenly detecting low bias when there really is none. It is important

to note that the bias test is very forgiving. A CEM system is said to be biased only when there is less than a 2.5% probability that the low readings occurred by chance.

The bias test is quite useful in detecting CEM system problems. Although the Part 75 requirements do not allow low-biased systems, high-biased systems are permitted [as long as the relative accuracy specification (10% for SO<sub>2</sub> and NO<sub>x</sub>, 15% for flow) is still met]. Obviously, although a high-biased system is allowed under the Acid Rain Regulations, it would result in the loss of allowances and would not be advantageous to a source owner. Therefore, a CEM system owner should apply the test to check for **both** low and high biases between the CEM system and reference method. Ideally, the cause of the bias should be detected and remedied to give the most accurate data possible.

#### 1.4 ELIMINATING BIAS AND THE BIAS ADJUSTMENT FACTOR

When bias is detected, two options are provided under Part 75. The preferable course of action is to determine the cause of the bias and eliminate the problem. This *Guide* is specifically designed to assist in this process by providing guidance in diagnosing and remedying the sources of measurement bias.

Alternatively, Part 75 provides a regulatory remedy. To compensate for the systematically low CEM measurements detected during the RATA, a bias adjustment factor can be derived from the RATA data and applied to subsequent CEM measurements. A CEM system owner is allowed the option of applying a bias adjustment factor if low bias is detected and the cause of the bias is not corrected. The bias adjustment factor is given in Eq. 1-8:

$$BAF = 1 - \frac{\bar{d}^*}{\overline{CEM}} \quad (\text{Eq. 1-8})$$

where

$BAF$  = bias adjustment factor

$\bar{d}^*$  = absolute value of the arithmetic mean of the difference obtained during the failed bias test using Eq. 1-3

$\overline{CEM}$  = Mean of the data values provided by the monitor during the failed bias test.

The magnitude of the bias adjustment factor is such that if the original CEM data were multiplied by the BAF, the average of the resulting values would exactly equal the average of the reference method readings and, consequently,  $\bar{d}^*$  would equal zero. Using Eq. 1-9, this factor is applied to all subsequent CEM system data for the measured parameter until the next relative accuracy test has been performed.

$$CEM_i^{Adjusted} = CEM_i^{Monitor} \times BAF \quad \text{(Eq. 1-9)}$$

where

$CEM_i^{Adjusted}$  = Data value, adjusted for bias, at time  $i$

$CEM_i^{Monitor}$  = Data (measurements) provided by the monitor at time  $i$ .

If the CEM system passes the bias test at the time of the next relative accuracy test, no adjustment would then be required. If the system fails, a new bias adjustment factor must then be calculated and applied unless the cause of the bias is determined and corrected.

When bias is detected but not corrected, CEM system bias adjustment factors are typically on the order of 3 to 4% of the CEM system measurement values. Before purchasing a CEM system, it should be decided by the user whether this level of adjustment would be acceptable. If not, the CEM system contract should specify to the CEM system vendor that bias-free or less biased Part 75 systems are to be provided.

It must be noted that it is always preferable from a measurement standpoint to eliminate the sources of bias in a CEM system rather than resort to the regulatory remedy provided by the bias adjustment factor.

## 1.5 SOURCES OF ERROR IN CEM SYSTEMS

Systematic and random errors can occur in all of the subsystems and components of a CEM system. It is left to the skill and experience of the CEM system manufacturer, integrator, and operator to minimize biases and obtain the best possible accuracy and precision. It is then the responsibility of the CEM system owner and operator to maintain the system to specified levels of accuracy and precision.

This guideline document will discuss sources of CEM system bias and possible methods of detecting and correcting bias problems. Specifically, bias problems associated with the following, will be discussed:

1. Sampling location and stratification
2. Dilution-extractive system biases
3. Source-level extractive biases
4. In-situ gas and flow monitor biases
5. Pollutant and diluent analyzer biases
6. Data acquisition and handling system problems

This document cannot identify all CEM system problems and sources of bias as many are system specific. However, it can point out some of the primary sources of systematic error that can be addressed when evaluating CEM system performance.

## 1.6 REFERENCES

International Standards Organization (ISO). 1989. Stationary Source Emissions - Determination of the Mass Concentration of Sulfur Dioxide - Performance Characteristics of Automated Measuring Methods. ISO Standard 7935. Central Secretariat, Geneva, Switzerland.

Jahnke, J.A. 1993. Continuous Emission Monitoring. Van Nostrand Reinhold. New York.

Natrella, M.G. 1963. Experimental Statistics. National Bureau of Standards. Handbook 91. Washington, DC.

## 1.7 ADDITIONAL READING

U.S. Environmental Protection Agency. 1977. Calculation and Interpretation of Accuracy for Continuous Emission Monitoring Systems (CEMS). Section 3.0.7 of the Quality Assurance Handbook for Air Pollution Measurement Systems, Volume III, Stationary Source Specific Methods. EPA-600/4-77-027b. August 1977. U.S. Environmental Protection Agency. Cincinnati, OH.

U.S. Environmental Protection Agency. 1993. Performance Specifications. U.S. Code of Federal Regulations - Protection of the Environment. 40 CFR 60 Appendix B. U.S. Government Printing Office.

U.S. Environmental Protection Agency. 1993. Test Methods. U.S. Code of Federal Regulations - Protection of the Environment. 40 CFR 60 Appendix A. U.S. Government Printing Office.

U.S. Environmental Protection Agency. 1993. Acid Rain Program: Continuous Emission Monitoring. U.S. Code of Federal Regulations - Protection of the Environment. 40 CFR 75. U.S. Government Printing Office.