CHAPTER 7

SOURCES OF BIAS IN THE DATA ACQUISITION AND HANDLING SYSTEM

Chapter 7 Highlights

Data Acquisition and Handling System (DAHS) Problems

Problem: Name	Problem: Description	Corrective Actions	Page Refs
Improper Interfacing			
Distorted Inputs from Analyzer	Input signals to the DAHS from the analyzer, process controller, or sensors are distorted.	To detect problem, compare DAHS readings to strip chart recorder's. Replace or repair faulty components.	7-2
Synchronization Problems	Errors will result if system control and DAHS clocks are out of synchronization.	Prior to certification testing, fix any mismatch between system and DAHS clocks.	7-2, 7-3
Calculation Problems			
Round-Off Problems	Incorrect rounding methods can produce biased results.	Change math to meet accepted professional practices and the conventions in regulations.	7-3
Incorrect Parameters	Entering incorrect values for user-configurable parameters will produce recurring errors.	Re-enter correct values.	7-4
Incorrect Equations	Programming incorrect equations will produce recurring calculation errors.	Require DAHS developers to document and validate all equations and correct code.	7-4
Improper Correction Routines			
Automated Zero/Span "Corrections"	Such adjustments may not be warranted and, at times, can introduce errors.	Do not allow automated corrections, OR Require vendor to precisely define and print out each adjustment. Include definitions in QC plan.	7-5, 7-6
Flow Monitor Correction Factors	If not correlated with actual conditions, these factors can produce systematic error.	Re-test under all prevailing conditions. Then, re-calculate the factors.	7-6, 7-7
Faulty Dilution System P/T Corrections	Pressure and temperature corrections can produce errors if incorrectly derived.	Require vendor to specify factors used and how derived. Correct if wrong.	7-7
Bias Adjustment Factor (BAF)	The BAF is a regulatory remedy, not a technical correction for systematic error.	Avoid having to apply a BAF by eliminating the sources of bias. The lower the BAF, the higher the confidence in the CEM's accuracy.	7-7, 7-8

CHAPTER 7

SOURCES OF BIAS IN THE DATA ACQUISITION AND HANDLING SYSTEM

The CEM data acquisition and handling system (DAHS) must also be addressed when considering bias. Biases can occur in manipulating and presenting data as well as in acquiring the data. Unfortunately, today's sophisticated methods of presenting digital data instill an overconfidence in the validity of the computer print-out. The fact that CEM data are presented by a computer, in digital formats, does not guarantee that the data are true and unbiased.

DAHS biases can occur in two ways: (1) by improperly interfacing the analyzers to the DAHS and (2) by improperly programming the DAHS. Interface problems are usually detected prior to or during system certification. Programming problems can be difficult to detect and may appear either during certification or months later, when inconsistencies begin to appear in the data.

7.1 INTERFACING

7.1.1 Analyzer Inputs

A CEM system data acquisition and handling subsystem must interface with the system analyzers. Data must be entered into the computer before it can be manipulated. However, inputs may vary from analyzer to analyzer. Signals transmitted to the computer may be in analog form, or they may be transmitted digitally in newer systems. In these newer systems, analog to digital (A/D) conversion is performed by microprocessors within the analyzer, simplifying signal transmission to the DAHS.

The analyzer signal may be transmitted as current (milliamps) or as voltage (millivolts). When analyzer signals, plant process parameters, and sensor signals for temperatures, pressures, and alarms, are to be received by the DAHS, care must be taken that these input signals are not modified or distorted. The use of drop-in resistors or simple circuits to convert milliampere current output to a voltage input compatible to the computer may cause shifts or distortions in the signal.

Interface problems can often be detected by connecting a strip chart recorder directly to the analyzer. Time delays, loss of resolution, or shifts in signal magnitude between the strip chart recorder and the DAHS indicate that a problem is present. Such problems need to be resolved during system installation. Although the differences may not appear significant in the mid-range readings of the system, they may become considerably distorted at the high or low values of the range, depending upon the input configuration.

7.1.2 Control System/DAHS Synchronization

A CEM system computer may be used for both system control and data acquisition and handling. Systems that blend supervisory control and data acquisition are known as SCADA (Supervisory

Control and Data Acquisition) systems. This integrated approach can be useful, since the internal status labeling of data (i.e., for calibration, filter purges, and data errors) can be accomplished with one device.

Alternatively, a separate computer may be used for control or programmable logic controllers may be applied for this purpose. In either of these two cases, the controller must be in synchronization with the data acquisition system. That is, if the two systems are running on different clocks, they must at some point provide a means of manipulating or transferring data on the same time basis. If there is a mis-match between these two systems, signal shifts can result. Again, such problems should be resolved during system installation.

7.2 PROGRAMMING

Programming of the DAHS can also lead to CEM system biases. Calculational problems or data adjustment algorithms can result in nonrepresentative data. Although the CEM system may meet calibration and audit checks and the computer-generated output may "appear" correct, this does not necessarily mean that the data to be reported will "be" correct. Improper manipulation of the analyzer input signals by the DAHS can generate biases just as well as can measurement failures in the extractive or emission analysis systems.

Calculation problems in the programming are relatively simple to detect and easy to resolve. These problems can arise from round-off errors, the use of incorrect parameters, or the use of incorrect equations. The increasing use of computer QA audit programs (either by EPA or commercial vendors) can help identify these problems relatively quickly.

The internal rounding methods used by the computer and the calculation algorithms can have an effect on the end result. For example, it has been shown that the results can depend upon the calculation order. If, however, calculations are performed using double-precision arithmetic, errors on the order of only 10^{-14} would be expected (Xiao et al., 1993).

Alternatively, if data are truncated or rounded to a smaller number of significant decimal places than are actually measured, a bias can result even if double-precision arithmetic is applied. For example, if a CO_2 reading of 10.2% is obtained by the analyzer and the computer uses a rounded value of 10% in the calculation, a significant discrepancy will result from the otherwise true value.

To prevent errors due to cumulative rounding, EPA policy stipulates that intermediate values used to calculate a final test result should be retained to the maximum decimal precision (at least seven decimal places) supported by the computer used. This is in keeping with accepted professional standards and practice. For example, ASTM Standard Practice E29-90, §7.3 (ASTM 1992) states "When calculating a test result from test data, avoid rounding intermediate quantities. As far as practicable with the calculating device or form used, carry out calculations with the test data exactly and round only the final result.

In user-configurable areas of a DAHS system, parameters such as calibration gas values, F-factors, or other constants can be changed in the calculation algorithms. If improper values are entered, biases will certainly result. For example, if an F-factor is used to obtain values of the emission rate in lbs / mmBtu, any error in the F-factor will be reflected directly in the emission rate calculation. This is generally not a problem for sources burning a single fuel, but for oil / gas, coal / gas, and combined or alternating fuel systems, more attention must be paid to the use of these parameters.

Calibration gas values entered into the DAHS can also affect CEM system data. If the DAHS performs automatic daily calibration adjustments, the data will be adjusted using the calibration gas value input as a reference. If the gas value is keyed in incorrectly, if the wrong value was entered, or if the gas manufacturer incorrectly analyzed the cylinder, a constant bias will enter into the reported data. For these reasons, the cylinder gas values should be cross-checked (as discussed in Chapter 6) and the entered values should be verified.

Equations used in the DAHS programs can also cause system biases. It is not uncommon that programmers, unfamiliar with EPA regulations or the technological basis of CEM systems, will develop the CEM system DAHS programs. Simple errors, such as incorrectly ordering the arithmetic steps of a calculation, programming an exponent as positive instead of negative, or using an equation for a dry-basis calculation instead of a wet basis calculation have all occurred in past programs. Unfortunately, the form of the equation is often buried in the code, and it is difficult for the plant environmental engineer or a CEM systems auditor or inspector to uncover the actual calculations used.

It is therefore necessary for the purchaser of a CEM DAHS to require the DAHS vendor to present in the system manuals or instructions all equations used in the program algorithms. It is also desirable for sources to require DAHS developers to prepare written specifications that explicitly state the equations to be programmed and to include in the acceptance criteria a requirement for independent verification and validation of the code to ensure (1) that the software developer's equations match those in the regulations and (2) that the code correctly implements the equations specified.

7.3 ADJUSTMENT/CORRECTION ROUTINES

The requirement of 40 CFR 75 monitoring plan (U.S. EPA, 1993) to provide equation formats serves as an excellent check for equation validity. However, some calculations are not required to be provided in the monitoring plan and may be proprietary to the CEM system vendor. Pressure and temperature correction routines for dilution probe systems, linearity corrections, etc., are often programmed into the DAHS, but their existence may not be known to the user. Improper correction equations can introduce as much error, or more, as improper report calculations. The following correction routines are those that are commonly encountered:

1. Daily zero / span corrections,
2. Flow monitor system corrections applied for Reference Method 2 correlation,

3. Dilution system pressure / temperature corrections,
4. Linearity corrections / other corrections not accounted for by the analyzer or analyzer microprocessor, and
5. Bias Adjustment Factor (BAF) as required by 40 CFR 75 Appendix A §7.6.5.

7.3.1 Daily Zero/Span Corrections

Two schools of thought exist among CEM system vendors with regard to computer zero / span corrections. The conservative approach is not to allow the computer to perform any zero span corrections using the daily calibration error check (zero / span) data. Instead, the CEM system operator must manually adjust the analyzers after some designated quality control limit has been exceeded (such as 2.5% of span). The other approach is to automatically correct the analyzer data after each daily calibration error check. Automatic corrections are often performed even if quality control limits have not been exceeded. In either case, Part 75 requires that both zero-level and high-level calibration error be determined and recorded before any adjustments are made, whether manual or automatic.

Each approach has its limitations. When no automatic corrections are performed, bias will be introduced (e.g., up to 2.5% of span) if the analyzer is exhibiting a consistent drift. When the system shows a 1% or 1.5% of span drift consistently, then the system may indeed be biased by that amount and it may be worthwhile to adjust the system, even if the control limits have not been exceeded. If, on the other hand, the zero / calibration values are bouncing back and forth between the control limits, only random noise is being exhibited and, over the long term, no bias will be introduced.

Of course, for manual adjustments, more stringent QC limits can be established. CEM technicians become uncomfortable with 2.5% drift limits and frequently adopt a policy of adjusting the system for drift levels as small as 1%. The danger here, of course, is that the technician may merely be adjusting for noise or other random factors, a procedure that is not particularly productive.

For automatic zero / span adjustments, the system may again adjust merely for random noise. The system may appear that it is doing something, but it may be only correcting for random phenomena, producing a result that would be essentially equivalent to one in which the corrections were not performed. However, if the daily zero and span values exhibit a consistent drift or are consistently high over a period of time (such as when a change of barometric pressure due to a weather front affects the system), the automatic correction may indeed be useful in minimizing the bias associated with the condition.

Several additional problems occur with automatic zero / span correcting systems. The most significant of these is <u>knowing the starting point from which you are correcting</u>. If an automatic correction is performed, the computer should print out the amount of correction (either in terms of ppm, percent of span, or both). The reference point for the correction should also be made clear. That is, is the correction made from the previous day's value, or is the correction cumulative, being made from the original zero and span gas settings? The cumulative correction is preferred—if the drift value refers to drift from the adjusted values of the previous day, automatic adjustment upon automatic adjustment could add up so that one may have drifted less than 2.5% on any one day, but the cumulative drift from the original setting may far exceed a 2.5% drift limit. It is important, therefore, that the operator be able to access the true, raw measurement data versus the "compensated data" in order to double-check if control limits are being exceeded.

Automatic drift corrections have often been a matter of some confusion to CEM system users. It is therefore imperative that the CEM vendor explain to the user what is actually being done in the adjustment and that that explanation be included in the CEM system QC plan.

Another disadvantage to automatically correcting for zero and span drift is that a strip chart recorder connected directly to the analyzer will not read the same as the computer. If the strip chart record is taken directly from the analyzer output and used to verify system performance, this record should first be compared to both the raw and compensated DAHS data.

7.3.2 Flow Monitor System Corrections

As mentioned in Chapter 5, problems of stratification in flow monitoring systems are frequently accommodated by introducing correction factors into the flow monitoring calculations (Stahlschmidt, 1992; Traina, 1992). This practice is common to all of the flow monitoring techniques: differential pressure, thermal, and ultrasonic. Such corrections are valid under the conditions in which they were originally developed. If the correction factor is established at only one load condition and the flow pattern varies under other conditions, the factor may not be valid. If a correction algorithm is developed under low, mid-range, and high load conditions, the adjustments may be shown to be valid over this range of load conditions.

The introduction of system bias in this type of correction can occur also if the source tester performed the reference method with an uncalibrated pitot tube. If a pitot tube calibration factor, C_p, of 0.84 was assumed (as is allowed by EPA Reference Method 2), a positive bias of 6% could result if the actual calibration factor was found to be 0.79 by wind-tunnel testing. Also, if the Reference Method 2 tests were not performed correctly and carefully or were performed carefully but did not traverse across areas of stratified flow, biases could be again introduced. For example, a common procedural error occurs when the tester fails to determine the proper pitot tube alignment at each point by measurement of the null yaw angle. Incorrectly aligning the tube to achieve the maximum response introduces a 5–7% bias in the reference velocity measurements. The S-type pitot tube used in Reference Method 2 is also sensitive to pitch angle bias (i.e., when it sags). Other types of pitot tubes, such as the 3-D pitot tube described in EPA Method 1 and draft Method 2F, can overcome this problem.

The practice of obtaining correction factors for velocity by conducting a so-called "pre-RATA" prior to the actual certification is widespread. However, if the correlation is not performed with some insight and the correction factor is not constant with changing flue gas conditions or with time, the system may fail semiannual / annual performance testing. Since a pre-RATA is not allowed prior to the semiannual / annual RATA, some element of risk exists in the practice.

7.3.3 Dilution System Pressure/Temperature Corrections

As discussed in Chapter 3, dilution extractive systems apply corrections for changes in absolute stack pressure and, in some cases, stack temperature. Correction algorithms are applied by most dilution / extractive system vendors to improve measured system accuracy. However, many vendors view these algorithms and correction factors as proprietary and do not readily share data on their development with users. Others use only the theoretical expressions, not experimentally derived correction factors. In some cases, the expressions used for these corrections have been wrong. Although the errors here are not great and may have been acceptable for other applications, the importance of Part 75 CEM data accuracy requires attention to this issue.

7.3.4 Linearity/Other Corrections

Other correction algorithms may sometimes be applied in the DAHS. For example in some systems, the gas analyzers consist of merely the sensing elements (i.e., lamp, sample cell, and detector) and the DAHS performs all of the signal manipulation. This manipulation may include signal linearization as well as zero / span adjustments. Here, the distinction between the DAHS and analyzer is blurred. In most systems, this function would be handled internally by the analyzer circuitry or analyzer microprocessor.

This type of system design increases the difficulty of system troubleshooting. In these systems, the DAHS programming becomes more complicated due to the addition of analyzer signal control and manipulation functions as part of the data handling requirements. Biases that may be introduced by the analyzer itself may become difficult to detect.

7.3.5 Bias Adjustment Factor (BAF)

The bias adjustment factor has been discussed in Chapter 1 (see Eqs. 1–8 and 1–9). Here, in considering sources of bias in the data acquisition and handling systems, it is important to clearly reiterate the purpose of the BAF. The BAF is a regulatory remedy, not a technical adjustment factor. It was adopted by EPA in direct response to an industry proposal to provide a compliance alternative to elimination of the sources of systematic error in situations where corrective actions were unusually difficult or expensive. As such, the BAF serves a twofold purpose: It provides flexibility in compliance options and, at the same time, serves as a safeguard against reporting artificially low emission measurements that nevertheless meet regulatory requirements for relative accuracy. For example, if a relative accuracy of 10% were permitted without a corresponding bias test and bias adjustment requirement, data that were systematically low relative to the standard but meeting the relative accuracy specification, would be acceptable (as in 40 CFR 60). However, this would give an acid rain trading allowance advantage to a source with a low-biased CEM system; the source would be reporting emissions lower than true and could possibly trade allowances that should not have been.

In any case, the BAF should not be viewed as a multiplier that the DAHS employs to correct CEM system bias. From a measurement standpoint, it is always preferable to eliminate all sources of bias in the CEM system and, thereby, completely avoid having to apply a BAF at all. The next best alternative is to minimize the sources of systematic error and, in so doing, minimize the value of the BAF. This may require care in system design, installation, and certification. But in general, the lower the BAF, the higher the confidence in the accuracy of the CEM system data.

7.4 Summary

Problems that can occur in the generation of CEM system data by the DAHS are summarized in the table on page 7–1. Many of the errors can be readily corrected once the problem is uncovered.

A number of these problems can be detected through the application of computer data validation programs. These audit programs, or routines, are being developed by EPA for validation of Part 75 data submitted to the agency (Moritz et al., 1993). Routines are also being developed by commercial programmers. In either case, a common technique used is the development of a test data set that contains traps and errors designed to challenge the CEM DAHS. If the DAHS correctly produces summary data from the test data set, there is increased assurance that those algorithms checked by the audit program are satisfactory.

Computerized audit programs that check all CEM system algorithms are unquestionably difficult to design. Manually checking the DAHS for accuracy and system logic against known parameters should still be conducted using actual sampled data.

7.5 References

American Society for Testing Methods (ASTM). 1992. Standard Practice for Using Significant Digits in Test Data to Determine Conformance with Specifications, E29-90. Annual Book of ASTM Standards. Philadelphia: 14.02:73-76.

Moritz, L., Price, D., and Salpeter, A. 1993. Electronic Data Reporting and the Acid Rain Program. Continuous Emission Monitoring - A Technology for the 90s. Air & Waste Mgmt. Assoc., Pittsburgh, pp. 177-185.

Stahlschmidt, D. 1992. Ultrasonic Flow Measurement of Gases for EPA Monitoring. Air & Waste Mgmt. Assoc. Meeting Paper. Kansas City: 92:121-07.

Traina, J.E. 1992. Feasibility of Installing Volumetric Ultrasonic Flow Monitors on Non-Optimal Ductwork. Air & Waste Mgmt. Assoc. Meeting Paper. Kansas City: 92-66.14.

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.

Xiao, Y., Butcher, J., Warren-Hicks, W. 1993. Memorandum - Equivalency Criteria Study. Cadmus Group, Inc.

7.6 ADDITIONAL READING

Mohnen, V.A., and Durham, J. 1992. Quality Assurance for Air Quality Monitoring in the United States of America. Staub. Reinhaltung der Luft 52, pp. 13-17.

U.S. Environmental Protection Agency. 1993. DAHS Verification. Description Pamphlet. U.S. EPA Acid Rain Program Office. Washington, DC.

U.S. Environmental Protection Agency. 1993. Electronic Data Reporting. Instruction Pamphlet. U.S. EPA Acid Rain Program Office. Washington, DC.

U.S. Environmental Protection Agency. 1993. Monitoring Plan & Instructions. EPA Form 7610-12 (1-93).

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.