CHAPTER 5

CALCULATION OF PRECISION, BIAS, AND
METHOD DETECTION LIMIT FOR CHEMICAL
AND PHYSICAL MEASUREMENTS

Issued by

Quality Assurance Management and Special Studies Staff
Office of Monitoring Systems and Quality Assurance
Office of Research and Development
United States Environmental Protection Agency
Washington, D.C. 20460

March 30, 1984

TABLE OF CONTENTS

TABLE OF CONTENTS
(cont'd)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS

(cont'd)

Chapter 5

CALCULATION OF PRECISION, BIAS, AND METHOD DETECTION LIMIT
FOR CHEMICAL AND PHYSICAL MEASUREMENTS

5.1 INTRODUCTION

One objective of the U.S. Environmental Protection Agency's
(USEPA) quality assurance program is to ensure that environmentally
related measurement data used by the Agency are of known and documented
quality. A most important measure of data quality is the variability
of the measurement system including the acts of sample acquisition
through sample analysis. One goal of the Agency's QA Program is to
establish uniform procedures for calculating and reporting measurement
system data quality indicators (i.e., precision, bias, and method
detection limit).

5.1.1 Purpose

The purpose of this chapter is to present procedures for the
assessment and reporting of precision, bias, and method detection limit
(MDL) for environmentally related chemical and physical measurements.
These procedures should aid in the incorporation of such assessments
into all data acquisition activities and in reporting the assessments
in project reports and environmental data bases.

5.1.2 Scope

This chapter addresses the variability of the measurement system
including the activities of sample acquisition through analysis. It is
concerned with the **error distributions** of the routine measurement
system, not the distributions of the chemical and/or physical measure-
ments from a study project or a monitoring program. This chapter also
does not address the assessment of representativeness of the measure-
ments of the environmental conditions under investigation (i.e., how
the measurement data relate to the target population).

All statistical concepts and analyses involve the use of data
quality assessment results obtained for the specific purposes of

estimating data quality, i.e., bias, precision, and MDL. The proce-
dures discussed in this guideline are based on the assumption that
measurement errors for chemical and physical measurement systems are
normally or near normally distributed. It is stressed that for partic-
ular situations where this does not appear to be a valid assumption or
where problems occur outside the scope of this guideline, e.g., as how
to treat outliers, the advice of a qualified statistician should be
obtained.

5.1.3 Application

Application of these procedures involves assessments of precision,
bias and MDL based on special measurements (hereafter referred to as
data quality assessments) of samples of known composition (e.g., refer-
ence materials, spiked samples, blanks) or of unknown composition
(e.g., replicate study samples or repeated analyses of study samples)
interspersed throughout the periods of routine operation of the
measurement system. Generally, study samples are received for analysis
in batches or lots (use of the term lots applies to any set or subsets
of data obtained from a study project or monitoring program) and data
quality assessments are made during the course of analysis of the study
samples in the lot. Application of the data quality assessments to the
measurements for the complete sample lot or groups of lots is based on
the assumption that all measurements, including the data quality
measurements, are made with the measurement system "in statistical con-
trol". A measurement system is considered to be in statistical control
when its variability is due only to chance causes. Data quality
assessment results used with control charts (Appendix H, ref. 13 of
Section 5.7.5) or statistical techniques such as the construction of
frequency distributions (Chapter 3, ref. 4 of Section 5.7.5) may be
used to assure that the measurement system is in statistical control.
The statistical measures of data quality prescribed in this chap-
ter should be used in conjunction with archived information on the
operational capability of measurement systems "Compilation of Data
Quality Information for Environmental Measurement Systems", (ref. 1,
Section 5.7.1) in the development of Quality Assurance Project Plans
for EPA environmental measurement programs as required in QAMS-005/80
(ref. 2 of Section 5.7.1).

5.1.4 Contents

This chapter includes: a discussion of basic statistical concepts and definitions in Section 5.2; procedures, with examples, for calculating and reporting estimates of precision, bias, and MDL in Sections 5.3. 5.4, and 5.5, respectively; and a case study illustrating the use of the statistical procedures for calculating precision and bias is given in Section 5.6. References, a Glossary of Terms, and recommended formats for reporting data quality indicators are provided, in the above order, in Sections 5.7 through 5.9.

5.2 STATISTICAL CONCEPTS

The principal indicators of data quality are bias and precision. **Bias** is systematic error. **Precision** involves the closeness of data values to each other. **Accuracy** involves closeness of measurements to a reference value and incorporates both bias and precision. To formalize these definitions, some basic statistical concepts are presented and discussed in the ensuing paragraphs.

It is reiterated here that this chapter is concerned with the **error distributions of the measurement system** and not the distributions of the chemical and/or physical measurements from a study project or a monitoring program. The assessments of precision and bias, do, however, apply to the measurements from a study project or monitoring program. All statistical concepts and analyses involve the use of data quality assessments obtained for the specific purposes of estimating data quality, i.e., bias, precision, and MDL.

5.2.1 Sample Statistics and Population Parameters

5.2.1.1 Sample Statistics--

If a quantity X (e.g., a concentration) is measured n times, it is customary to refer to the values X_1, X_2, . . . X_n so obtained as a **sample of size** n of measurements of X. (It should be clear throughout the text from the context whether the word "sample" is used in this statistical sense or refers to, e.g., a chemical sample.) Values obtained in the calculation of quantities which summarize data of this kind are **summary statistics** or **sample statistics**.

More specifically, let X_1, X_2, . . ., X_n be a set of independent data quality assessments taken under fixed and prescribed experimental conditions and regarded as a random sample from some population. For many applications the X_i are considered measurements aimed at estimating a reference or true value T. Some basic statistics for a sample of size n are:

o the **sample average** $\bar{X} = (X_1 + \ldots + X_n)/n$ Eq. 5-1

o the **sample bias** $B = \bar{X} - T$ Eq. 5-2

o the **sample standard deviation** $s = \sqrt{\dfrac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2}$ Eq. 5-3

o   the **sample variance** is s^2

o   the **sample coefficient of variation** (also called
relative standard deviation) $CV = 100\ s/\bar{X}$ Eq. 5-4

o   the **error** of the ith measurement $e_i = X_i - T$,   Eq. 5-5

Sample standard deviation and variance, s and s^2, are measures
of precision with smaller values of s indicating better precision. If
the measurements X_i are widely dispersed, the values of $(X_i-\bar{X})^2$ in Eq.
5-3 will tend to be large, giving a large value of s^2 or s; whereas,
a small spread gives small values. This is illustrated in the follow-
ing diagram.
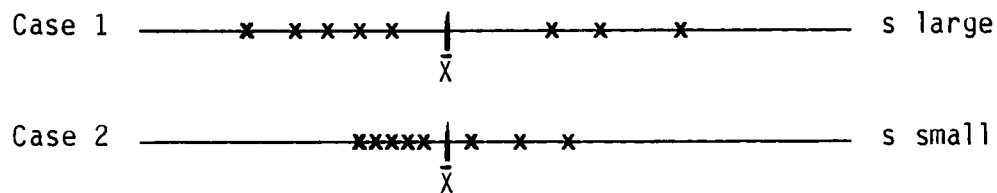
In the above two cases, the measured values X_i are indicated by the
crosses, and the sample averages are the same. For the specific appli-
cation of measurement system error, it is assumed that results of
repeated, independent, measurements of the same sample, under the same
specified conditions, will be normally or near normally distributed
about their average value. There are several techniques for checking
for normality of a set of results including the use of histograms,
normal probability graphs (see Appendices C and D, respectively of
reference 13 in Section 5.7.5), and goodness-of-fit tests. The
assistance of a qualified statistician may be required to develop a
detailed protocol.

The sample size n and values of B and s are fundamental summary
statistics and should be reported in routine sampling situations.
These values may be calculated automatically using a statistical com-
puter (or calculator) program. The following example shows the
"mechanics" of the calculation.

Example 5-1. Calculation of B and s: A reference chemical sample
at an assumed concentration, T = 50 concentration units (CU), was
analyzed four times to assess the bias and precision of the analysis
phase of the measurement system. The values X_1, X_2, X_3, X_4 obtained
were 48, 55, 50, and 45 CU. Values of B and s may be calculated as in
the following table:

Measurements	X_i	$X_i-\bar{X}$	$(X_i-\bar{X})^2$
i = 1	48	-0.5	0.25
2	51	2.5	6.25
3	50	1.5	2.25
4	45	-3.5	12.25

Average $\bar{X}$ = $(X_1+X_2+X_3+X_4)/n$

$\quad$ = $(48+51+50+45)/4$

$\quad$ = $194/4$

$\bar{X}$ = 48.5 CU

Variance $s^2 = \sum_{i=1}^{4} (X_i-\bar{X})^2/n-1$

$\quad$ = $(0.25+6.25+2.25+12.25)/3$

s^2 = 21/3 = 7.0

$\underline{s\ =\ 2.6\ CU}$

Bias $\quad$ B = $\bar{X}-T$

$\quad$ = 48.5-50

$\underline{B\ =\ -1.5\ CU}$

Based on these results the estimated bias of the measurement system, under the above conditions, is -1.5 CU and the estimated precision, expressed as the standard deviation, is 2.6 CU. On a practical basis the bias is not considered to be significant because the assumed value T is contained within the range of the four measured values.

5.2.1.2 Population Parameters--

If a random variable X is measured many times, the calculated sample statistics will approach constant values referred to as **population parameters.** For example, when n is very large, the sample average $\bar{X}$ of n measurements approaches a value called the **population mean**, or simply the **mean**, which is denoted by the symbol μ. In the absence of measurement bias, μ is the true value of the quantity being estimated. Thus, The appropriate estimator for the population mean μ based on n measurements is the sample average $\bar{X}$. The value of $\bar{X}$ generally will not coincide with μ but does give the best estimate based on the measurements available.

Similarly, if many measurements are made, the sample variance s^2 approaches the **population variance** which is customarily denoted by σ^2. Its square root, σ, is the **population standard deviation** (s.d.). Again, σ and σ^2 are generally not known but may be estimated

by the sample statistics s and s^2 (provided at least two measurements are made since n $\geq$ 2 is required to apply Equation 5-3).

The quantities μ, σ, and σ^2 are called **population parameters.** Another commonly used parameter is **coefficient of variation** (CV) or **relative standard deviation** (r.s.d.). This is the ratio 100 σ/μ which expresses the standard deviation as a percentage of the mean and is sometimes useful when the standard deviation changes with levels being measured. It is **estimated** by the ratio 100 s/$\bar{X}$, the **sample** CV.

The following table summarizes the **population parameters** described above and the **sample statistics** used for their estimation.

Population Parameter	Symbol	Sample Statistic	Symbol
mean	μ	average	$\bar{X}$
standard deviation	σ	standard deviation	s
variance	σ^2	variance	s^2
CV (or r.s.d.)	100 σ/μ	CV	100 s/$\bar{X}$

It should be emphasized that there are situations with more structure than that of a simple random sample from a single population. However, as previously noted, for the most part, this overview will be concerned with one or more "simple" samples, rather than more structured situations.

5.2.2 Components of Data Quality

5.2.2.1 Precision--

Precision is a measure of the scatter of a group of measurements, made at the same specified conditions, about their average value. The **sample standard deviation** s and **sample coefficient of variation** CV are used as indices of **precision.** The smaller the standard deviation and coefficient of variation, the better the precision. Precision is stated, in the units of measurement or as a percentage of the measurement average, as a plus and minus spread around the average measured value $\bar{X}$.

In instances where precision estimates are obtained from analyses of **replicate pairs**, the **range** R (maximum value - minimum value) or **relative range** RR (100 R/$\bar{X}$) is sometimes used as an index of **precision**. For replicate pairs the relationship between the range and standard deviation is s = R/ $\sqrt{2}$.

5.2.2.2 Bias--

Bias, as estimated with sample statistics, is the **signed difference** between the average $\bar{X}$ of a set of measurements of a standard and the "true" value of the standard T given by

$$B = \bar{X} - T.$$

Bias can be negative or positive and is expressed in the units of measurement or as a percentage of the value of the standard. Percentage bias is given by

$$\%B = 100 \ (\bar{X} - T)/T.$$

Bias is also estimated by average percent recovery $\bar{P}$ (see Section 5.4.1 for a definition of percent recovery). The relationship between percent bias and average percent recovery is:

$$\%B = \bar{P} - 100.$$

5.2.3 Components of Variance

For any measurement system there are many <u>sources of variation or error</u>, some of which are sample collection, handling, shipping, storage, preparation, and analysis. For each individual or groupings of error sources within a measurement system, there are different <u>classifications of variation or precision</u>. Different classifications are the result of the different conditions and manner in which the data quality assessments are made for estimating precision.

Intralaboratory precision is the variation associated with a single laboratory or organization. Intralaboratory precision must be further subclassified as **short-term** or **long-term** precision depending on the conditions and manner in which the precision data are obtained. Intralaboratory precision is usually referred to as

repeatability. **Interlaboratory** precision or **reproducibility** is
the variation associated with two or more laboratories or organizations
using the same measurement method.

When precision information is used to assess system performance,
care should be taken that the appropriate measure and classification of
precision is used. When reporting precision estimates for environment-
al measurements, the conditions represented by the estimate (e.g.,
short- or long-term, single analyst, etc.) should be documented, and
the component or components of the measurement system included in the
estimate (e.g., the total system, sample preparation and analysis, or
analysis only) should be specified.

5.3 PRECISION

5.3.1 Definitions

The most commonly used estimate of precision is the sample **standard deviation**, as defined in Section 5.2. Other precision measures include the **coefficient of variation** (100 s/$\bar{X}$), **range R** (maximum value - minimum value), and **relative range** RR (100 R/$\bar{X}$).

Collocated samples are independent samples collected in such a manner that they are equally representative of the variable(s) of interest at a given point in space and time. Examples of collocated samples include: samples from two air quality analyzers sampling from a common sample manifold or two water samples collected at essentially the same time and from the same point in a lake.

A **replicated sample** is a sample that has been divided into two or more portions, at some step in the measurement process. Each portion is then carried through the remaining steps in the measurement process.

A **split sample** is a sample divided into two portions, one of which is sent to a different organization or laboratory and subjected to the same environmental conditions and steps in the measurement process as the one retained inhouse.

Collocated samples when collected, processed, and analyzed by the <u>same</u> organization provide **intralaboratory** precision information for the entire measurement system including sample acquisition, handling, shipping, storage, preparation and analysis. Both samples can be carried through the steps in the measurement process together providing an estimate of short-term precision for the entire measurement system. Likewise, the two samples, if separated and processed at different times or by different people, and/or analyzed using different instruments, provide an estimate of long-term precision of the entire measurement system.

Collocated samples when collected, processed and analyzed by <u>different</u> organizations provide **interlaboratory** precision information for the entire measurement system.

A **replicated or split** sample can be divided into portions (or split) at different points in the sampling and analysis process in order to obtain precision information on the various components of the measurement system. For example, a **field replicated,** or **field split sample,** provides precision information about all steps <u>after</u> sample acquisition including effects of storage, shipment, analysis, and data processing; whereas, information on the intra- and interlaboratory precision of sample preparation and analysis steps of the measurement system is provided by samples subdivided once they are received in the laboratory, i.e., laboratory replicated or laboratory split samples, respectively. A sample divided into two portions just prior to analysis, i.e., an **analysis replicate**, provides information on the precision of the analytical instrumentation.

The **replicated sample** can provide short-term or long-term precision estimates by processing the two portions together or separating them for processing at different times and under different conditions as discussed above for collocated samples.

5.3.2 General Guidelines

The precision assessment should represent the variability of the entire measurement system. Therefore, collocated samples are recommended, when possible, as the preferred method of assessing precision of the entire measurement system.

A sample subdivided in the field and preserved separately is used, where possible, to assess the variability of sample handling, preservation, and storage along with the variability of the analysis process. If the nature of the matrix, sample acquisition procedure, or analytical technique prevent the assessment of the entire measurement system, the replicated samples used to assess precision should be selected to incorporate as much of the measurement system as possible.

Data quality assessments should be made at concentration levels typical of the range observed in routine analyses using the same method in the same laboratory. In situations where the typically observed measurement values are zero but a standard is set at some finite value, repetitive measurements should be made of samples representative of the standard and not zero. The design of the data quality assessment program will depend upon such factors as the data quality needs, the

precision of the measurement system, and the size of the sample lot.
For large sample lots, a fixed frequency for replicate measurements
(such as one sample in ten or twenty) is recommended. For small sample
lots, more frequent repetition may be desirable to ensure that suffi-
cient data are available to assess precision. Alternatively, multiple
sample lots with a common matrix, analyzed by the same measurement sys-
tem, can be combined as discussed under continual precision assessments
(Section 5.6). If the environmental measurements normally produce a
high percentage of results below the MDL (see Section 5.5), samples for
replicate measurement should be selected from those containing measura-
ble levels of analyte. Where this is impractical, such as with complex
multi-analyte methods, sample replicates may be spiked at appropriate
concentration levels to ensure that sufficient data will be available
to assess precision.

Precision information can be dealt with in a variety of ways
depending on the specific situation of interest and the type of data
available.

Interpretation of precision data must always be based on a clear
knowledge of how the data were created. For example, the precision of
the entire measurement system, including sample acquisition, can only
be assessed by analyses of collocated samples. Precision data gener-
ated from multiple analyses of a standard only describe the stability
of the measurement device or instrument and only represent the ultimate
precision which could be achieved for a field sample if the sampling
activity, subsequent sample preparation steps, and the sample matrix
had no impact on final results.

Figure 5-1 graphically outlines these samples in a general sense,
but specific sampling and analysis situations may require additional
precision information and more extensive breakdown of precision evalua-
tion samples. If this is the case, a clear indication of what is being
done and why should be provided in data quality assessment documenta-
tion.

5.3.3 Calculation of the Summary Precision Statistics

Summary statistics provide an assessment of the precision of a
measurement system or component thereof for a project or time period
(i.e., for a sample lot or group of sample lots). They may be used to:

Data Quality Assessment Sample Point of Origination			Data Interpretation
Sample Acquisition	Preparation	Analysis	

Best estimate of intra- or interlaboratory precision of the entire measurement system. (Collocated Samples)

Second best estimate of intralaboratory precision of measurement system from sample acquisition through analysis. (Field Replicate)

Second best estimate of interlaboratory precision of the measurement system from sample acquisition through analysis. (Field Split)

Best estimate of intralaboratory precision of sample preparation and analysis. (Lab Replicate)

Best estimate of interlaboratory precision of sample preparation and analysis. (Lab Split)

Estimate of intralaboratory precision of analysis. (Analysis Replicate)

Estimate of interlaboratory precision of analysis. (Analysis Split)

Figure 5-1 Precision Evaluation Samples.

o estimate precision at discrete concentration levels.

o average estimated precision over applicable concentration ranges.

o provide the basis for a continual assessment of precision of future measurements.

The summary statistics are developed from the basic statistics gathered throughout the project or time period represented. Because the precision of environmental measurement systems is often a function of concentration (e.g., as concentration increases, standard deviation increases), this relationship should be evaluated before selecting the most appropriate form of the summary statistic. An evaluation of the basic precision statistics as a function of concentration will usually lead to one of three conclusions:

o **Case 1: standard deviation (or range) is independent of concentration (i.e., constant);**

o **Case 2: standard deviation (or range) is directly proportional to concentration, and coefficient of variation (or relative range) is constant; or**

o **Case 3: both standard deviation (or range) and coefficient of variation (or relative range) vary with concentration.**

For simplicity of use and interpretation, the relationship most easily described should be selected for use, i.e., for Case 1 the standard deviation (or range) is simplest to work with; whereas, for Case 2, the coefficient of variation (or relative range) is simplest. If the relationship of precision to concentration falls into Case 3, regression analysis can be used to estimate the relationship between standard deviation (or range) and concentration.

The decision as to which case is applicable can be based on plots of precision versus concentration or by regressions of s (or R) or CV (or RR) versus concentration by an approach summarized in Table 5-1 and illustrated in the following examples.

TABLE 5-1. RELATIONS BETWEEN PRECISION AND CONCENTRATION

Case	Plot of s vs $\bar{X}$	Plot of 100 s/$\bar{X}$ vs $\bar{X}$	Slope[a] b_1	Intercept[a] b_0
1. Constant Standard Deviation	s appears constant for all $\bar{X}$	CV decreases as $\bar{X}$ increases	0	>0
2. Constant CV, 100 s/$\bar{X}$	s increases as $\bar{X}$ increases	CV appears constant for all $\bar{X}$	>0	0
3. Other	s increases as $\bar{X}$ increases	CV decreases as $\bar{X}$ increases	>0	>0

[a] Based on the relationships of: $s = b_1 \bar{X} + b_0$, and

$$CV = 100 \ s/\bar{X} = 100 \ (b_1 + b_0/\bar{X})$$

Legend:

The slope b_1 and intercept b_0 are parameter estimates from the linear regression of standard deviation s on average estimated concentration $\bar{X}$. To decide among the cases, use one of the following two methods.

Method 1: Regress s on $\bar{X}$ and base the decision on statistical tests of the hypotheses $b_1 = 0$, $b_0 = 0$. The assistance of a qualified statistician may be required.

Method 2: Review plots of standard deviation vs $\bar{X}$ and 100 s/$\bar{X}$ versus $\bar{X}$. If one plot appears more constant than the other, then elect that case. If nonconstancy is evident from both plots, regress s on $\bar{X}$ to get the precision relationship. The assistance of a qualified statistician may be required.

Example 5-2 (Case 1. s is independent of concentration): Three samples in a 10-sample lot, whose concentrations approximated the range of concentrations of all ten samples, were each analyzed in triplicate to assess the precision of analysis. At estimated concentrations of 20, 32, and 40 CU, the standard deviations were calculated to be 2.2, 2.6, and 2.4 CU, respectively. Based on the plots in Figure 5-2, s appears to be independent of concentration; whereas, CV definitely
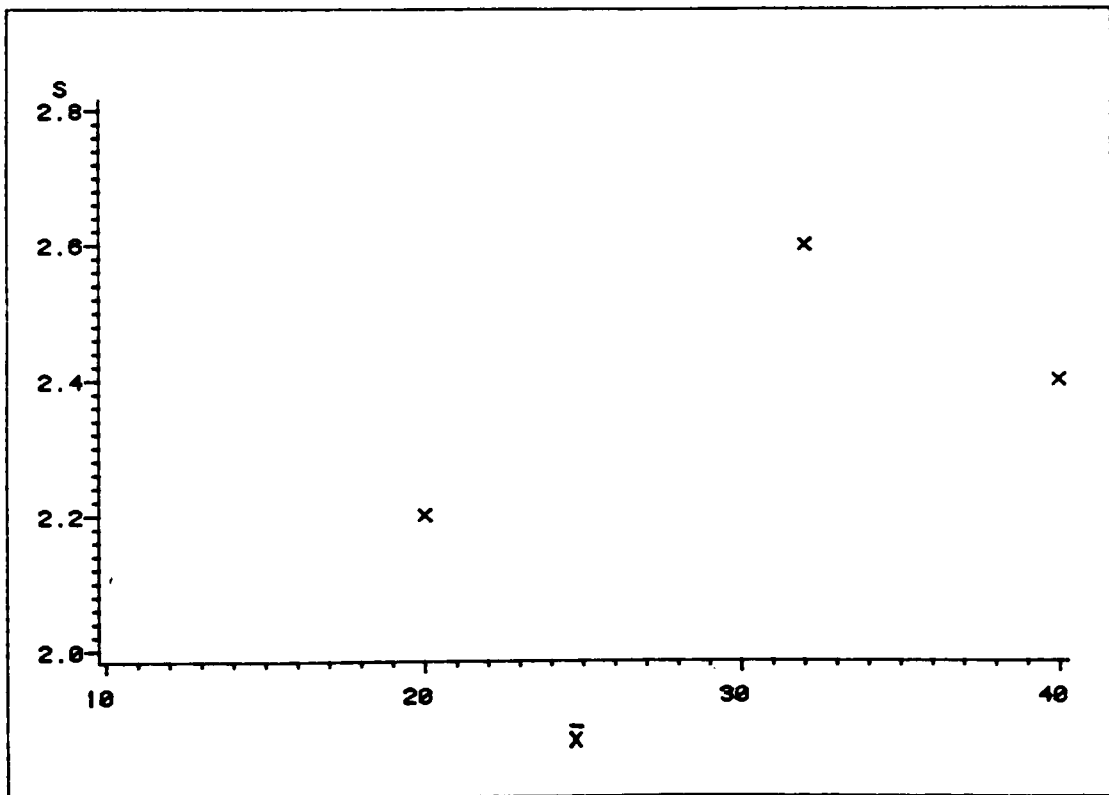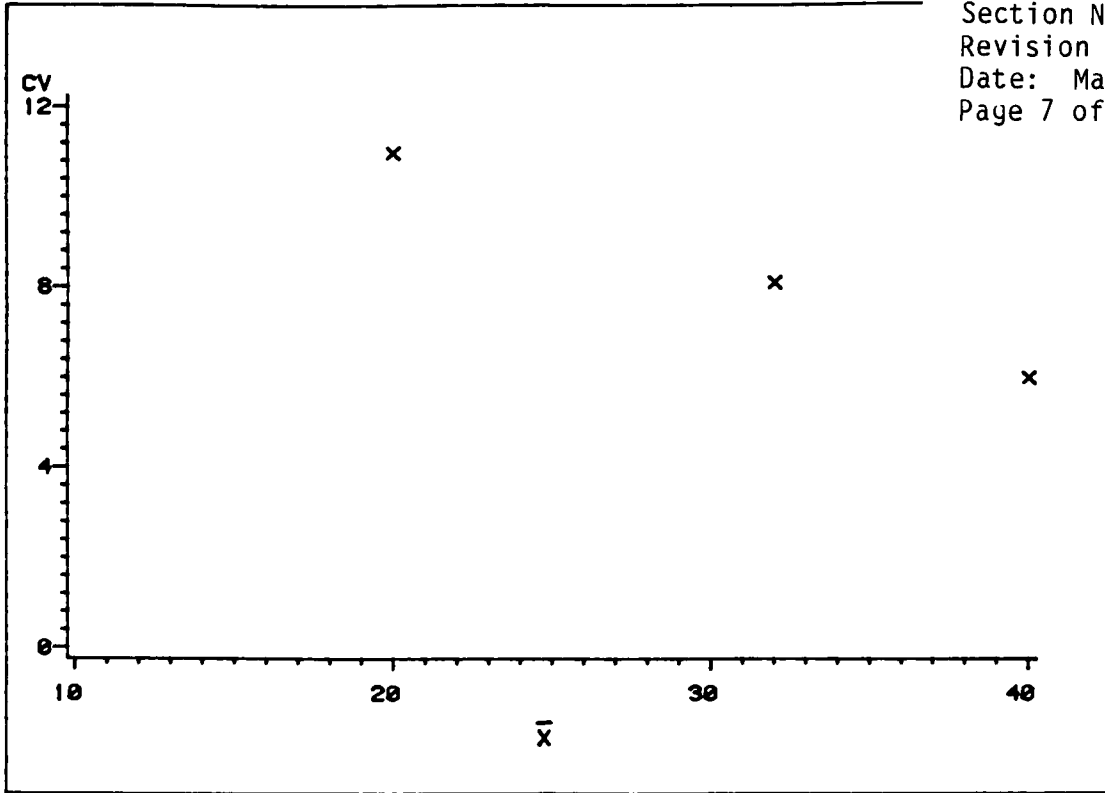
Figure 5-2.   Plots of CV (%) versus $\bar{X}$ (CU) and s (CU) versus $\bar{X}$ (CU)
for Example 5-2.

decreases with increasing concentration so that the precision of analysis for all samples in the lot may be most conveniently summarized by the pooled standard deviation s = 2.4 CU. The pooled standard deviation is calculated as:

$$s = \left[\frac{(n_1-1)\, s_1^2 + (n_2-1)\, s_2^2 + \ldots + (n_k-1)s_k^2}{n_1 + n_2 + \ldots + n_k - k} \right]^{1/2} , \qquad \text{Eq. 5-6}$$

and for this example

$$s = \left\{ [2(2.2)^2 + 2(2.6)^2 + 2(2.4)^2]/(3 + 3 + 3 - 3) \right\}^{1/2}$$

$$s = 2.4 \text{ CU.}$$

The standard deviation of individual measurements in this 10- sample lot then is reported to be 2.4 CU.

Example 5-3 (Case 2, s is directly proportional to concentration and CV is constant): For a study involving a 100-sample lot, 8 pairs of field replicated samples were processed and analyzed to assess the precision of the measurement system from sample acquisition through analysis. The results for each pair of samples (X_1 and X_2) are tabulated in order of increasing X_1, below along with values for replicate average, standard deviation, and coefficient of variation.

X_1	X_2	Average $(X_1 + X_2)/2$	Standard Deviation	Coefficient of Variation (%)
1.5	1.7	1.60	.14	9
1.7	1.6	1.65	.07	4
2.0	2.1	2.05	.07	3
2.4	2.1	2.25	.21	9
2.7	2.4	2.55	.21	8
3.9	4.3	4.10	.28	7
5.0	4.5	4.75	.35	7
5.2	4.7	4.95	.35	7

Average CV = 7%

An inspection of the data tabulation of the standard deviation and average for each pair suggests that the standard deviation increased as the average increased. Therefore, the coefficient of variation values were tabulated. An inspection of the tabulation of CV versus average reveals no clear relationship, i.e., the CV can be treated as a constant across the concentration range. These findings are confirmed by the plots in Figure 5-3. Therefore, the average CV (7%) is used as the precision summary for this sample lot. The precision, expressed as a standard deviation, of individual measurements in this example, at any concentration, X, is estimated to be, and is reported as, s = 0.07X.

Example 5-4 (Case 3, s increases with increasing concentration and CV decreases with concentration): If a constant relationship does not appear to exist between standard deviation and concentration or between coefficient of variation and concentration, then it is necessary to use a more complex approach, such as a linear regression equation, to describe the relationship. A least-squares linear regression analysis of the precision (i.e., s or CV) versus measured concentration results in two coefficients, a slope and an intercept, which are used to represent the precision of the data set. The range and relative range are sometimes used to estimate precision, particularly for replicate pairs, because of their ease of computation and use.

For a study involving a 100-sample lot, 10 collocated sample pairs were collected for estimating the precision of the entire measurement system. The results for each set of collocated samples are tabulated below along with values for the average, range, standard deviation, relative range, and coefficient of variation of each pair of samples. To simplify visual interpretation, the data have been ordered by increasing values of concentration.

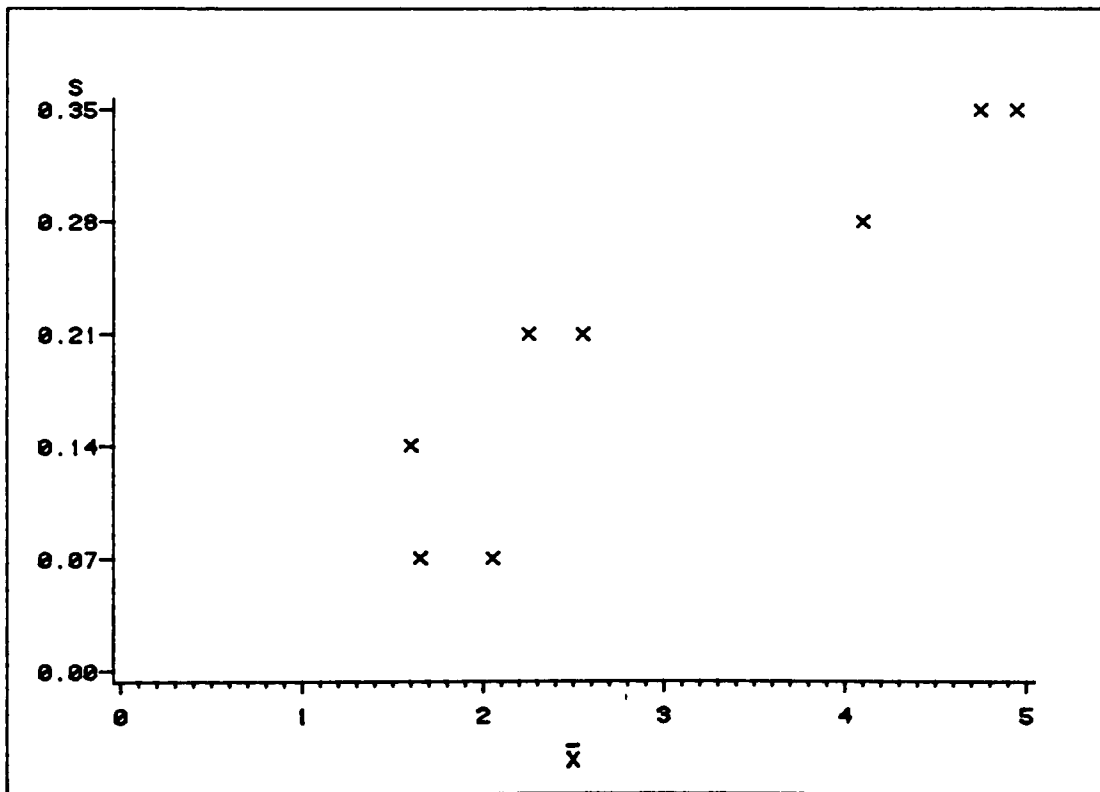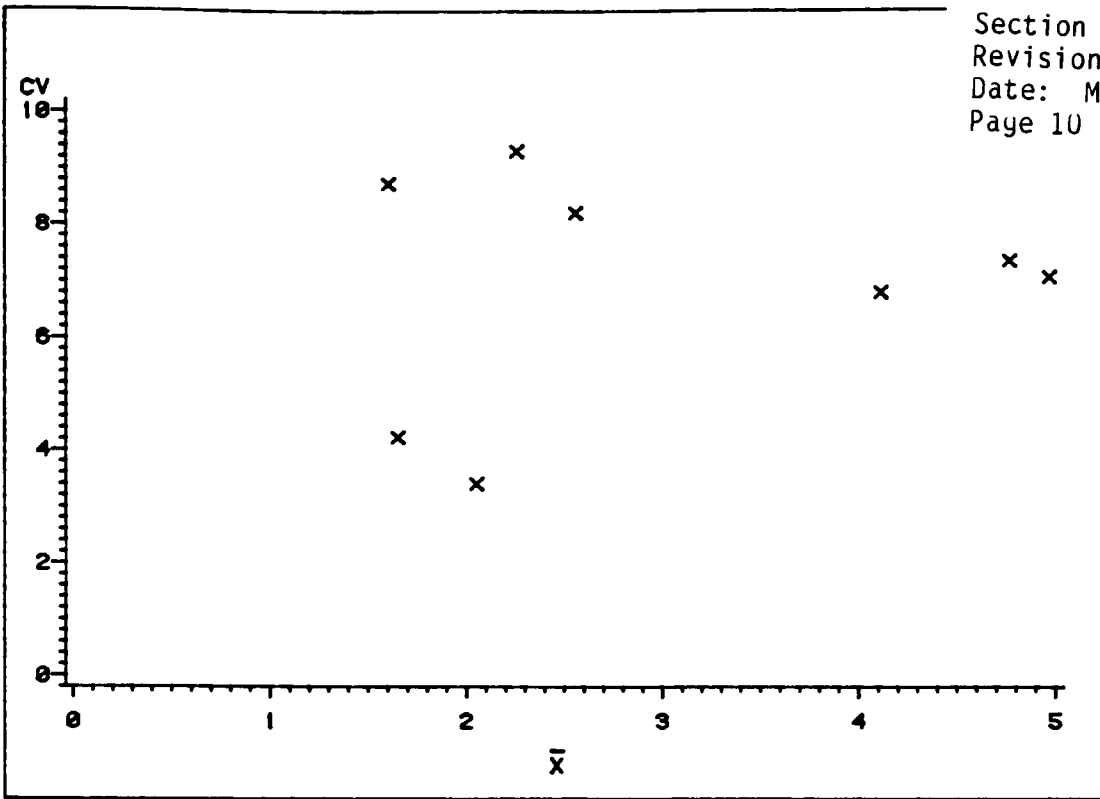| X_1	X_2	Average $(X_1+X_2)/2$	Range	Standard Deviation	Relative Range	Coefficient of Variation (%)
5.33	6.37	5.85	1.04	0.735	17.8	12
10.1	8.65	9.38	1.45	1.03	15.5	11
19.5	17.6	18.55	1.9	1.34	10.2	7
18.6	20.5	19.55	1.9	1.34	9.8	7
32.8	36.1	34.45	3.3	2.33	9.6	7
108.5	102.	105.2	6.5	4.60	6.2	4
132.	124.	128.0	8.0	5.66	6.2	4
186.	197.	191.5	11.	7.78	5.7	4
501.	527.	514.0	26.	18.38	5.1	4
3517	3341	3429	176	124.43	5.1	4

Figure 5-3. Plots of CV (%) versus $\bar{X}$ (CU) and s (CU) versus $\bar{X}$ (CU) for Example 5-3.

For organizations that prefer working with standard deviation and/ or coefficient of variation, plots of standard deviation versus average concentration and coefficient of variation versus average concentration are shown in Figure 5-4. These plots show that s is an increasing, approximately linear function of concentration whereas CV is a decreasing, nonlinear function of concentration. In this situation it is simpler to use a linear regression equation to represent the standard deviation over the concentration range. The calculated regression equation is $s = 0.036 \bar{X} + 0.698$. The individual measurements in this example are estimated to have, at any concentration X, a standard deviation of 0.036 X +0.698 CU.

For organizations that prefer working with range and relative range, the tabulation shows a clear increase in range, and decrease in relative range, with increasing concentration. A least-squares linear regression of range as a function of the average concentration for the data above yields the following regression equation $R = 0.051 \bar{X} + 0.987$. For replicate pairs the standard deviation and range are related by $s = R/\sqrt{2}$.

5.3.4 Reporting Precision

Because each data user must determine the data reliability required for his/her application, the data reporter must provide a standard deviation or alternative measure of precision which applies to measurement.

The data user should be provided with a narrative statement documenting the conditions and manner in which the precision data were obtained and the applicable component or components of the measurement system. Statements for reporting precision estimates for cases 1, 2, and 3 as illustrated in examples 5-2, 5-3 and 5-4, respectively are given as examples.

  o Case 1 (Example 5-2). The estimate of intralaboratory, short-
     term precision (i.e., within lot, single analyst) for each con-
     centration found in the sample lot is s = 2.4 CU, for concen-
     trations in the range of 20 to 40 CU, based on triplicate
     analyses of each of three samples. The standard deviation s
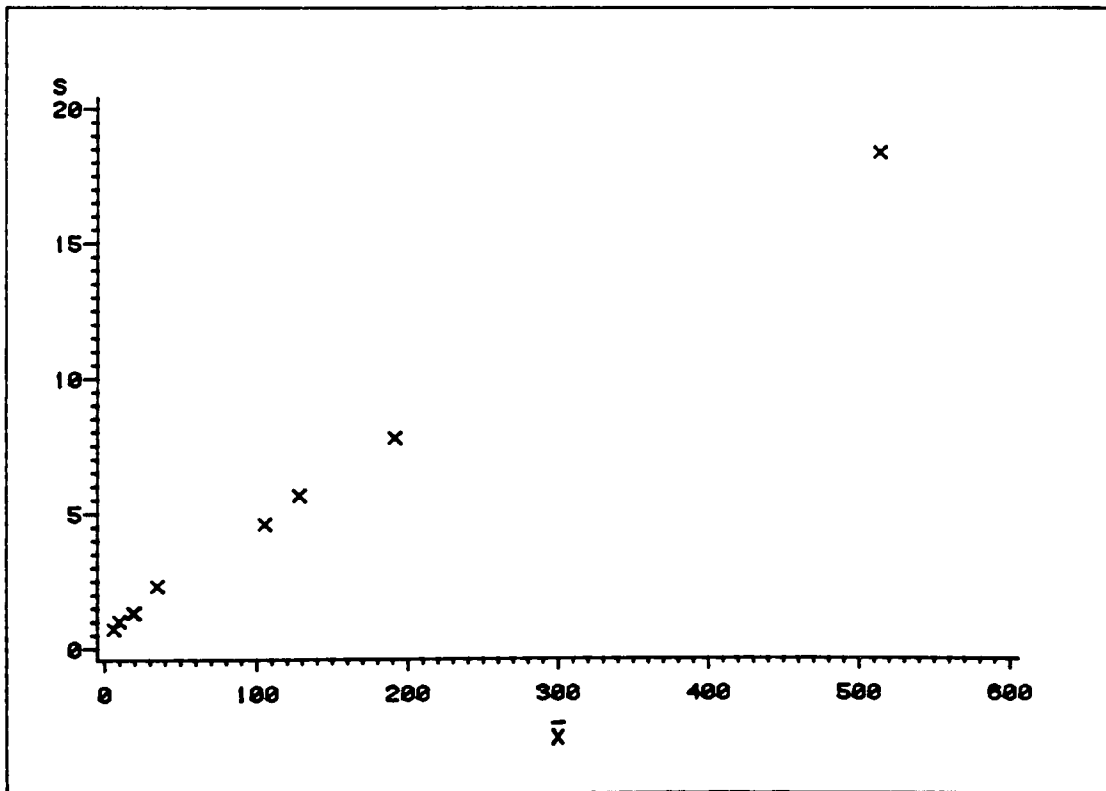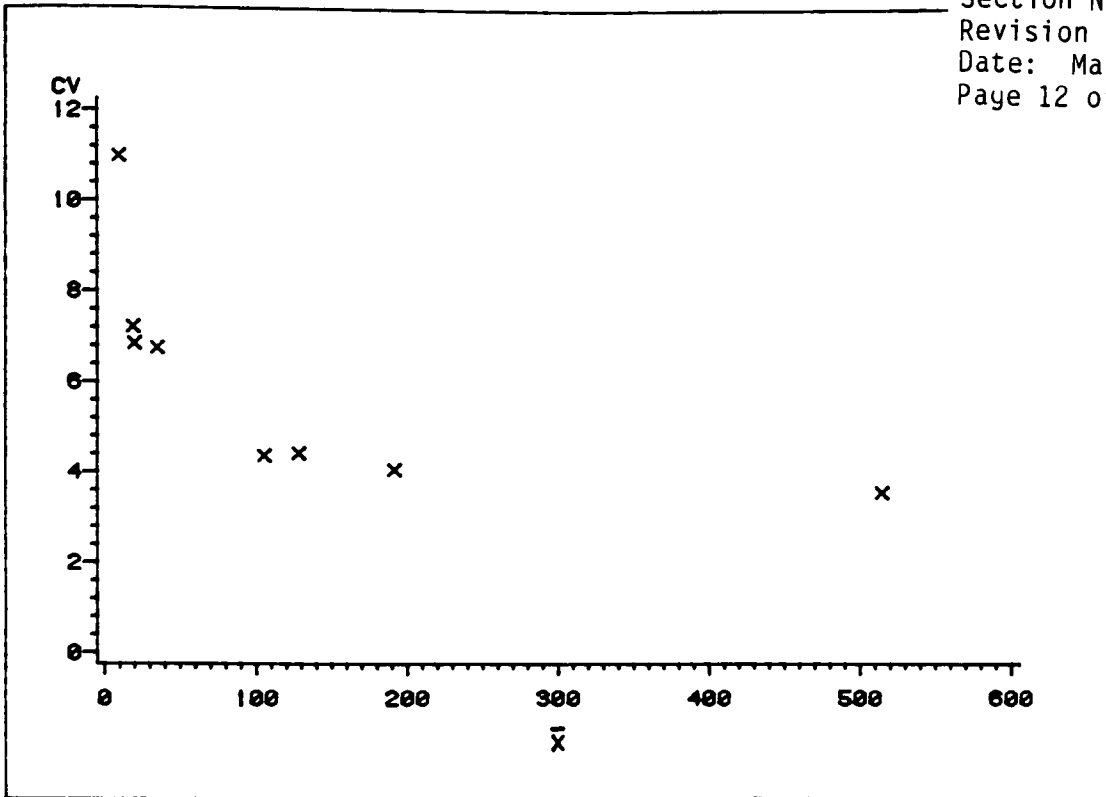   can be used to estimate a probability interval for the random

Figure 5-4.   Plots of CV (%) versus $\bar{X}$ (CU) and s (CU) versus $\bar{X}$ (CU)
for Example 5-4.

error associated with an individual observation X_i as follows. The approximate 95% probability interval for the difference in an observation X_i and the limiting average $\bar{X}$ (i.e., the value that would be obtained if the sample were analyzed many times) is ± 2 s. The approximate 95% probability interval for errors (excluding bias) for individual measurements in this sample lot is estimated to be ± 2 s = $\pm 2(2.4$ CU$)$ = ± 4.8 CU.

o Case 2 (Example 5-3). The estimate of intralaboratory, short-term precision for individual measurements in this sample lot including steps in the measurement system from sample collection through analysis is s = 0.07X CU for concentrations in the approximate range of 1.5 to 5.0 CU, based on the analyses of 8 field replicate pairs. For example, at X = 5.0 CU, the estimated standard deviation is s = 0.35 CU, the approximate 95% probability interval for errors (excluding bias), for X = 5.0 CU, is ± 2 s or ± 0.7 CU.

o Case 3 (Example 5-4). The estimate of intralaboratory, short-term precision for individual measurements in this sample lot including the total measurement system is s = 0.036X + 0.698, for sample concentrations in the approximate range of 5 to 3500 CU, based on the results from 10 pairs of collocated samples. The approximate 95% probability interval for the errors (excluding bias) in an individual measurement X is ± 2 s = $\pm 2(0.036X + 0.698)$ CU. For example, at X = 100 CU, s = 4.3 CU, the approximate 95% probability interval for errors (excluding bias) for X = 100 CU is ± 8.6 CU.

5.3.5 Continual Precision Assessments

For organizations in which sample lots are routinely analyzed and data are reported on a frequent basis, the basic precision statistics from multiple lots of a given sample matrix may be combined to provide an estimate of long-term precision and an improved estimate of short-term precision as illustrated in Section 5.6.3. This assessment can also be extended to include subsequent lots, unless test results for these new lots indicate that method precision is significantly different. This combining of data quality assessment results permits the laboratory to provide a precision assessment derived from a substantial amount of background data rather than from limited precision data produced in a small study.

This procedure also provides the basis for the use of control charts to monitor the performance of the measurement system. The

procedure is based upon the availability of a precision assessment (normally developed from prior performance of the system), the use of control chart limits, and routine replicate pairs.

Historical data must first be combined as necessary to develop an assessment of precision that defines the expected standard deviation of replicates as a function of concentration. For each replicate pair in the new sample lot, the observed standard deviation s is compared with an upper control limit (UCL) for the expected standard deviation, s^*, calculated for the observed average sample concentration $\bar{X}$. If $s \leq 3.27\ s^*$, the established precision assessment can be applied to the individual members of the new sample lot. If $s > 3.27\ s^*$, either the established precision assessment is not applicable to the new data set, or the measurement system is out of control. (The upper control limit factor, for an s chart is $UCL = B_4\ s^*$ and for replicate pairs $B_4 = 3.27$. Note that B_4 should not be confused with bias B as used in this chapter.) For further information on the use of control charts, including the rationale for the 3.27 constant, see Reference 4 in Section 5.7.5.

At least annually, and preferably after the accumulation of results from 30 to 50 new replicate pairs, new assessments for precision must be calculated to reflect the current precision of the measurement system. This may be done by either expansion or replacement of the historical data base with the most current data.

Example 5-6: A laboratory is analyzing and reporting samples on a continual basis. The regression equation for R versus $\bar{X}$ estimated from historical data is $R = 0.051\ \bar{X} + 0.987$ CU, for replicate pairs. A replicate pair yields measurements of 18.6 and 20.5 CU. The expected range for the pair is calculated for the average concentration of 19.55 CU to be $R = 0.051(19.55) + 0.987 = 1.984$ CU. The control limit for the range of the pair is calculated to be $3.27(1.984)$ CU = 6.49 CU. Since the observed range of 1.9 CU (20.5 - 18.6 = 1.9) is less than the 6.49 CU control chart limit, the result is within expectations and there is no reason to suspect that the historical precision assessment is not applicable. A graphical presentation of R versus $\bar{X}$ may be convenient for use in a laboratory analyzing many samples.

5.4 ASSESSMENT OF BIAS

5.4.1 Definitions

Bias - an estimate of the bias B is the difference between the average value $\bar{X}$ of a set of measurements of a standard and the reference value of the standard T given by:

$$B = \bar{X} - T$$

Alternative estimates of bias are percent bias

$$\%B = 100\ (\bar{X} - T)/T,$$

and average percent recovery $\bar{P}$

$$\bar{P} = \sum_{i=1}^{n} P_i \text{, and}$$

$$P_i = 100\ (A_i - B_i)/T,$$

where A_i = the analytical result from the spiked sample and B_i = the analytical result from separate analysis of the unspiked sample. The relationship between percent bias and percent recovery is:

$$\%B = \bar{P} - 100$$

Reference material - A material of known or established concentration that is used to calibrate or to assess the bias of a measurement system. Depending on requirements, reference materials may be used as prepared or may be diluted with inert matrix and used as blind environmental samples.

Spiking material - A material of known or established concentration added to environmental samples and analyzed to assess the bias of environmental measurements.

Target analyte spiking - Spiking with the analyte that is of basic interest in the environmental sample.

Matrix spike - A sample created by adding known amounts of the target analyte to a portion of the sample.

Field matrix spike - A sample created by spiking target analytes into a portion of a sample in the field at the point of sample acquisition. This data quality assessment sample provides information on the target analyte stability after collection and during transport, and

storage, as well as on losses, etc., during sample preparation and on errors of analysis.

Laboratory matrix spike - A sample created by spiking target analytes into a portion of a sample when it is received in the laboratory. It provides bias information regarding sample preparation and analysis and is the most common type of matrix spike. This type of matrix spike does not necessarily reflect the behavior of the field-collected target analyte, especially if the target analyte is not stable during shipping.

Analysis matrix spike - A sample created by spiking target analytes into a prepared portion of a sample just prior to analysis. It only provides information on matrix effects encountered during analysis, i.e., suppression or enhancement of instrument signal levels. It is most often encountered with elemental analyses involving the various forms of atomic spectroscopy and is often referred to as "standard additions".

Non-target analyte spiking - Spiking of surrogate analytes into the sample. A **surrogate analyte** is one which mimics the behavior of target analytes in terms of stability, preparation losses, measurement artifacts, etc., but does not interfere with target analyte measurement. **This approach** is most frequently used with organic compound determinations and **is a compromise** which is dependent on the target compounds and surrogates involved. **Surrogates**, like matrix spikes, can be added in the laboratory or in the field; results are interpreted in a fashion similar to matrix spikes.

Internal standard spike - An analyte which has the same characteristics as the **surrogate**, but is added to a sample just prior to analysis. It provides a short term indication of instrument performance, but it may also be an integral part of the analytical method in a non-quality control sense, i.e., to normalize data for quantitation purposes.

Figure 5-5 provides a graphic representation of target analyte spiked samples in the preferred order starting at the top, based on point of spiking in the sampling and analysis scheme.

Sample Spiking Point	Data Interpretation

Sample Acquisition	Preparation	Analysis

┌──────────┐
│ Field   │
│ Matrix   │────────────────────────── Provides a best case estimate
│ Spike   │   of bias based on recovery;
└──────────┘   includes matrix effects asso-
                                       ciated with sample preserva-
                                       tion, shipping, preparation
                                       and analysis.

            ┌──────────┐
            │ Lab   │   Provides an estimate of bias
            │ Matrix   │────────────── based on recovery. Incor-
            │ Spike   │   porates matrix effects asso-
            └──────────┘   ciated with sample preparation
                                       and analysis only.

                      ┌──────────┐
                      │ Analysis │   Provides an indication of
                      │ Matrix   │──── matrix effects associated with
                      │ Spike   │   the analysis process only.
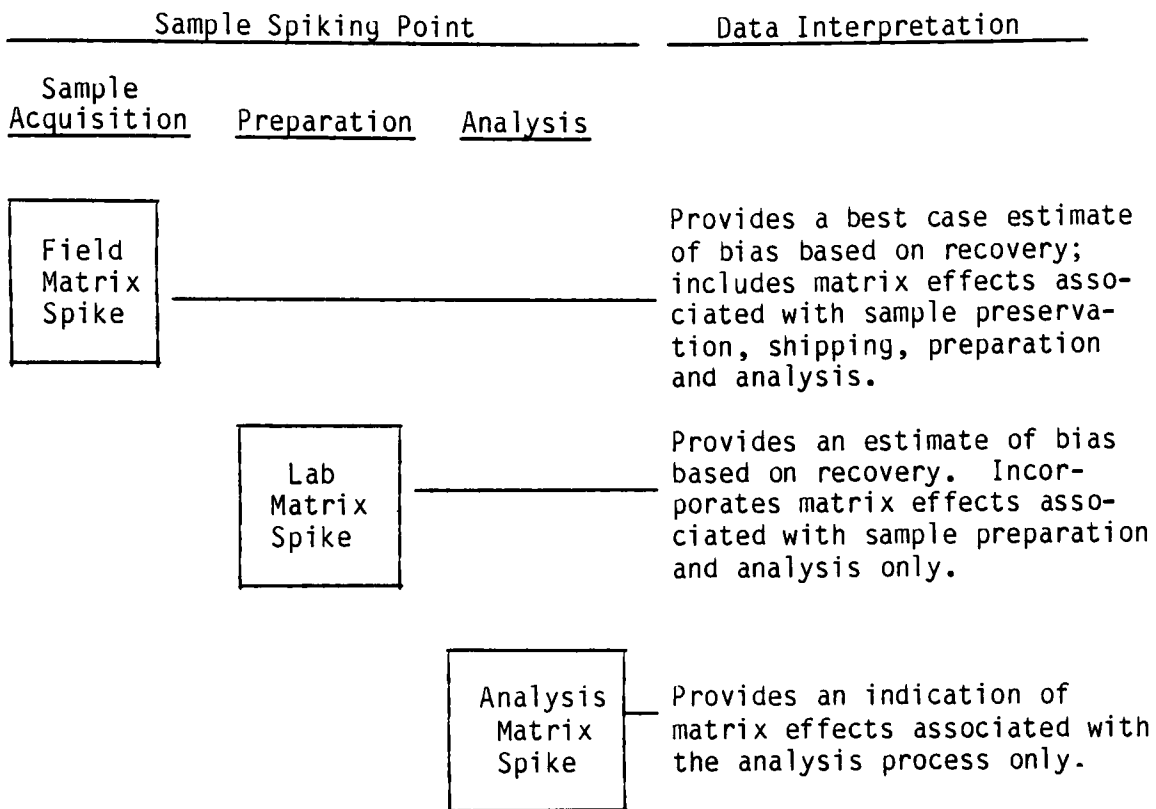   └──────────┘

Figure 5-5. The Use of Target Analyte Spikes for Bias Estimation

5.4.2 Measurement of Bias

Bias assessments for environmental measurements are made using spiking materials or reference materials, personnel, and equipment as independent as possible from materials used in the calibration of the measurement system. Where possible, bias assessments should be based on analysis of spiked samples rather than reference materials so that the effect of the matrix on recovery is incorporated into the assessment. A documented spiking protocol and consistency in following that protocol is an important element in obtaining meaningful data quality estimates. Spikes should be added at different concentration levels to cover the range of expected sample concentrations. For some measurement systems (e.g., continuous analyzers used to measure pollutants in ambient air), the spiking of samples is not practical, and assessments are made using appropriate blind reference materials.

Ideally spiking materials or reference materials should be introduced into samples at the collection site so that the bias assessment includes any losses caused by sample handling, preservations, and storage. If the matrix type or the measurement system prevents such practices, bias assessments should be made for as large a portion of the measurement system as possible.

A representative portion of the sample lot is selected for spiking and the selected samples are analyzed before and after spiking in order to measure recovery. The spiking frequency will depend upon the data quality needs of the program, the bias and precision of the measurement system, the size of the sample lot and other considerations. To properly assess the bias for a small sample lot, it may be necessary to spike a relatively high percentage of the samples. However, where the method performance for multiple lots of samples of similar matrix type is expected to be equivalent, it may be possible to combine information so that fewer spikes are required in each lot.

For certain multianalyte methods, such as EPA Method 608 for organochlorine pesticides and PCBs in water, bias assessments are complicated by mutual interference between certain analytes that prevent all of the analytes being spiked into a single sample. For such methods, lower spiking frequencies can be employed for analytes that are seldom, or never, found. The use of spiked surrogate compounds for multianalyte GC/MS procedures, while not ideal, may be the best

available procedure for assessment of bias. An added attraction is the ability to obtain recovery data on every field sample at relatively low costs. It is used, for example, to evaluate the applicability of methodology and, indirectly, data quality assessments to individual members of a sample lot. Such practices do not preclude the need to assess bias by spiking with the analytes being measured or reported.

5.4.3 <u>Calculation of Bias Statistics</u>

The most widely used summary of bias is by linear regression of bias on T; or, equivalently, regression of data quality assessment results (X_i or $\bar{X}$) on T (as illustrated in Section 5.6.3). For the important special case of spiked samples as described above, the following approach may also be useful.

A portion of the samples in the sample lot is spiked at multiple concentration levels to determine individual measurements of percent recovery. These recoveries are used to calculate summary bias statistics for the entire sample lot. The summary statistics are used to estimate the percent recovery for each individual measurement in the lot. For each sample spike i, calculate the percent recovery P_i by,

$$P_i = 100 \ (A_i - B_i)/T_i,$$

where: A_i = the analytical result from the spiked sample,

B_i = the analytical result from a separate analysis of the unspiked sample,

T_i = the known true value of the spike.

Average percent bias is calculated from average percent recovery for the sample spikes by the relationship

$$\% \ B = \bar{P} - 100.$$

If reference materials instead of spiked samples are analyzed to assess bias, percent recovery is calculated by the equation above with B_i equal to zero.

Upon completion of the project or time period, the bias assessment for the data set of environmental measurements is calculated from the individual percent recoveries P_i observed through the project period. Unless a relationship between the percent recovery, or its variability, and concentration can be established, all percent recovery measurements

may be combined for the bias assessment. Calculate the average percent recovery $\bar{P}$ and the standard deviation of the percent recovery (s_p). The meaning of the value for s_p may be considerably different from the precision assessment. For spiked samples it includes the variability of the unspiked measurement plus the variability of the spiked final measurement. In addition, the individual recoveries are usually gathered over an extended time period, rather than over short time intervals normally used for replicate measurements and therefore may reflect the presence of many other variables.

Example 5-7: For a 100-sample lot, 10 sample portions were spiked and analyzed along with unspiked portions. In this case the volume of spike to the sample volume was so small that no volume correction was required. The results of the analyses are tabulated below:

Sample Background B_i	Spike T_i	Result A_i	Recovery $A_i - B_i$	Percent Recovery, %
4.0	20.0	22.8	18.8	94.0
7.9	20.0	26.2	18.3	91.5
4.5	20.0	25.4	20.9	104.5
1.3	20.0	21.2	19.9	99.5
17.3	50.0	66.7	49.4	94.8
26.3	100.0	128.0	101.7	101.7
5.7	20.0	24.8	19.1	95.5
5.0	20.0	24.8	19.8	99.0
62.5	200.0	260.5	197.8	98.9
34.5	100.0	135.3	100.8	100.8

$$\text{Average } (\bar{P}) = \sum_{i=1}^{10} P_i/n = 980.2/10 = 98.0$$

$$\text{Standard Deviation } (s_p) = \left[\sum_{i=1}^{6} (P_i-\bar{P})^2/n-1 \right]^{1/2} = 4.0$$

The average percent recovery, $\bar{P}$, and the standard deviation of the percent recovery, s_p, are $\bar{P}$ = 98.0%, s_p = 4.0%.

The approximate 95% probability interval for percent recovery for individual measurements in the sample lot is 98.0% $\pm$ 2(4.0)%, i.e.,

from 90% to 106%. The percent recovery associated with any individual measurement in this 100-sample lot is estimated to be between 90% and 106%, at the 95% probability level.

Example 5-8: In any application of statistics it is important to identify underlying assumptions and check them with the data by graphic display or statistical checks. Figure 5-6 is a plot of recovery versus concentration for analyses of sulfate ion deposited on Teflon® filters. Evidently recovery is not constant, but is a slightly decreasing function of concentration. Therefore, a summary in terms of a single average percent recovery would be of questionable validity. This data is examined in more detail using a linear regression approach in Section 5.6.3.

5.4.4 Reporting Bias

Each environmental measurement must be reported with an assessment of bias. Bias should be expressed as a percent error interval of %B $\pm$ 2 s_p or as a percent recovery interval from $\bar{P}$ - 2 s_p to $\bar{P}$ + 2 s_p. (There are several ways to express bias and accuracy. It should be noted here that expressing bias in this manner is not consistent with the definition given in Subsection 5.4.1. Additional efforts will be made to achieve consistency in the definition and use of bias in future revisions.) Where reference materials are used as a matrix-free check on laboratory performance as a supplement to sample spiking, only the results of the sample spikes should be submitted to an environmental data base.

The data user should be provided with a narrative statement explaining the reported bias estimate along with tabulated percent recovery intervals. The statement might read:

> "Bias is expressed as a 95% probability interval around the average percent recovery. A percent recovery interval of 90 to 106%, for example, means that approximately 95% of the time when a spike of the measured material is recovered, the observed percent recovery can be expected to lie between 90 and 106%."

5.4.5 Continual Bias Assessment

As with precision assessments, laboratories in which small sample lots are routinely analyzed and data are reported on a frequent basis
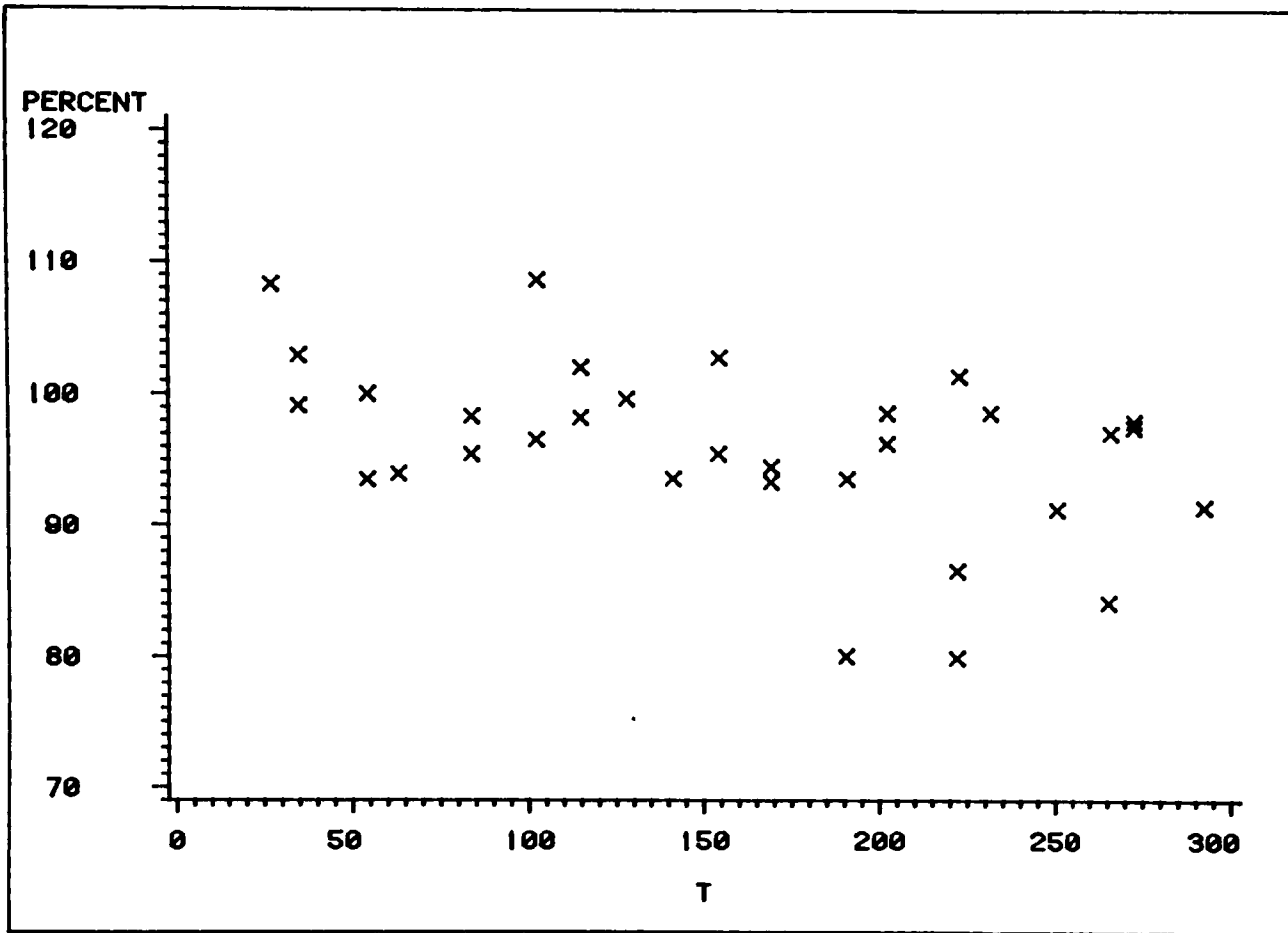
Figure 5-6. Plot of Percent Recovery (P_i) versus Concentration (T_i) for Sulfate Data Quality Assessment Results.

may combine the basic bias statistics of multiple small sample lots of a given matrix into a single bias assessment for the combined sample set. This assessment can also be extended to include subsequent small sample lots, unless test results for these new lots indicate that method bias is significantly different.

Combining bias data in this manner permits the laboratory to provide bias assessments derived from a substantial amount of background data rather than from limited bias data produced in a small study. It also can provide the basis for the use of control charts to monitor measurement system bias over time.

Historical data must first be combined as necessary to develop an assessment of bias which includes the determination of average percent recovery ($\bar{P}$) and the standard deviation of the percent recovery (s_p). These estimates maybe used to develop control chart limits as $\bar{P} \pm 3\ s_p$ for subsequent measurements. Each recovery measurement, P_i, in a new sample lot must be compared with the control chart limits. If each value for P_i falls within the control limits, the historical percent recovery and analysis assessment can be applied to all individual measurements of the new sample lot. If P_i falls outside the control limits, either the historical assessment is not applicable to the new data set or the laboratory operation is out of statistical control.

At least annually, and preferably after no more than 30 to 50 new recovery measurements have been taken, the control chart limits must be recalculated to reflect the current percent recovery capabilities of the measurement system. This may be done by either expansion or replacement of the historical data base to include the most current data.

Example 5-9: A laboratory is analyzing and reporting samples on a continual basis. Historical data for the analysis of spiked samples established that $\bar{P} = 98.0\%$, and $s_p = 4.0\%$; the control chart limits are $\bar{P} \pm 3\ s_p$, or 86 and 110%. A sample with a measured background level (B_i) of 22.0 CU was spiked with the equivalent of 30.0 CU (T) without appreciably changing the sample volume. The result for the analysis of the spiked sample (A_i) was 49.2 CU. Therefore:

$$P_i = 100\ \frac{49.2 - 22.0}{30.0} = 90.7\%.$$

Because P_i falls within the control chart limits of 86 and 110%, the sample is consistent with the historical recovery and analysis assessment. A graphical presentation of $\bar{P} \pm 3 \; s_p$ versus test number may be convenient for use in a high volume laboratory.

5.5 METHOD DETECTION LIMIT

5.5.1 Discussion

There have been many terms used to designate detection limits and they have been defined in various ways. Lower Limits of Detection (LLD), Minimum Detection Amount (MDA), Method Detection Limit (MDL), Detection Sensitivity, and Limit of Detection (LOD) are some of the terms used. Most authorities in the field agree that the smallest detectable quantity, by whatever name, is related to the standard deviation of sample analyses at or near zero analyte concentrations. Since MDL is a basic performance characteristic of an analytical method only its calculation, with example, is discussed.

5.5.2 MDL

Ideally each laboratory should establish and periodically reevaluate its own MDL for each sample matrix type (for one time only matrix types and multianalyte samples in difficult matrices, e.g., soils or fish, this may be impractical) and for each environmental measurement method. The MDL is determined for measurement systems by the analyses of seven or more replicates of spiked matrix samples. As with precision and bias, the assessment of MDL should be based upon the performance of the entire measurement system. The standard deviation of the responses (s_m), in concentration units, is used to calculate the MDL as follows:

$$MDL = s_m \, (t_{.99}) \qquad \text{Eq. 5-7}$$

where:

$t_{.99}$ = "Student's t value" appropriate for a one-tailed test at the 99% confidence level and a standard deviation estimate with n-1 degrees of freedom.

For example, if the MDL is determined using measurements from seven appropriate samples, then use $t_{.99}$ = 3.14 for n-1 = 6 degrees of freedom. If the determination yielded a standard deviation of 0.15 CU, the MDL is calculated (Equation 5-7) to be (3.14)(0.15 CU) = 0.47 CU.

The mechanics of calculating MDL are illustrated in the following example.

Example 5-10. Eight samples, identical in appearance to routine samples, spiked at a level of 0.02 CU were blindly inserted throughout a day of routine analyses to assess the system accuracy at low concentrations on that day.

The observed concentrations were 0.032, 0.016, 0.021, 0.022, 0.024, 0.017, 0.025, 0.019 CU. The standard deviation of these results is s_m = .005 CU. Assuming on the basis of previous experience that the true standard deviation is essentially constant for concentrations this low, a measurement must exceed (ts_m) in order to be significantly greater than zero at the 0.01 level of significance, using a one-tailed test. Here t denotes the upper percentile of Student's t-distribution with n-1 = 7 degrees of freedom, so the MDL estimate using equation 5-7 is (2.998) (.005 CU) = .015 CU.

5.6 A CASE STUDY

To illustrate the application of the statistical concepts previously discussed, data from an actual laboratory measurement program are presented and calculations of appropriate summary statistics are performed. The example includes calculation of within-lot estimates of precision and bias as well as between-lot summary statistics to illustrate a procedure for continual data quality assessment and/or estimating long-term precision of projects of long duration. The within-lot sample sizes are 3 or smaller in all but one case. Within-lot averages and standard deviations computed from such small samples tend to be imprecise estimates of the population mean and standard deviation. The example shows how improved estimates of these quantities can be obtained by averaging over the lots.

The average within-lot standard deviation provides a better estimate of short-term, intralaboratory, precision for the total data set. Likewise, using data quality assessments between- or across-lots can provide an estimate of between-lot variability, i.e., systematic error from lot-to-lot as well as an estimate of total variability or long-term precision over the subject time period.

The estimate of long-term precision is the appropriate measure to use in describing the precision of the total data set. Estimates of short-term (within-lot) precision and between-lot variability are useful as part of a data quality assessment program in monitoring the system's performance and can provide guidance for troubleshooting by indicating which component(s) of the measurement system is experiencing larger than normal variability.

The example goes somewhat beyond the treatment presented in Section 5.4, by employing linear regression to summarize relations between bias and concentration. In this example, a lot is defined as all measurements made within a given day, including routine analysis of study samples and data quality assessment samples at one or more concentrations.

The example data are from a program for ion chromatography analysis of sulfate deposits on Teflon® filters. Data quality assessment samples in this program consisted of sulfate solutions at three different concentrations, known to the analyst. The ion chromatograph was

calibrated at the beginning of each day of operation and data quality assessment samples were interspersed throughout the routine analyses as a check that the system was in control, and to allow assessment of data quality.

An additional data quality assessment was obtained from the regular analysis of samples provided by an external auditor. These samples more closely approximated the actual study samples than did the internal data quality samples, and consisted of known (only to the auditor) concentrations of sulfate deposited on Teflon® membrane filters. Thus, data quality estimates derived from these samples included components of variance due to sample preparation (i.e., extraction) as well as   . sample analysis.

The two types of samples will be referred to as solutions and filters, respectively.

5.6.1 Within-lot Summaries

Let $X_1, \ldots, X_n$ be data quality assessments associated with a single lot and made at equal or nearly equal concentration levels. The within-lot summary statistics for n data quality assessments at a fixed reference concentration T in the lot are:

$\bar{X}$ = the average of the data quality assessments made at a fixed concentration in the lot

s_w = the within-lot standard deviation of the n data quality assessments made at a fixed concentration

Obviously, when only a single measurement is obtained at a certain concentration for a given lot, the standard deviation cannot be calculated.

These within-lot summaries of data quality assessments are primarily for internal use. They are useful intermediates in the calculation of summary statistics based on longer time periods, such as six months or a year (Section 5.6.2). Within-lot averages and standard deviations for data quality assessments are also useful as variables to be monitored with control charts. If data quality assessments for a lot are taken at several known concentrations, regression statistics may also provide a useful summary of the data quality of the lot (see Section 5.6.3). To fully utilize these statistics the assistance of a qualified statistician may be required.

Data quality assessments for the example sulfate analysis program and associated within-lot statistics (i.e., n, $\bar{X}$, s_w and T) are presented in Table 5-2 for three values of solution concentration, 0, 120. and 240 CU. As seen in the table, the within-lot sample size (i.e., the number of data quality assessment samples) ranges from n=1 to n=6. The within-lot average and standard deviation can vary widely with such small sample sizes.

5.6.2 Between-Lot Summaries

This procedure allows for the calculation of summary statistics for a project time period that spans several lots. The following statistics are used to report data quality over a given time period (e.g., quarter, year). They are obtained from within-lot statistics described in the previous section, and are calculated for all data quality assessments obtained under a particular set of fixed conditions, such as matrix, concentration, and analysis procedure. For the example sulfate anaysis program, between-lot statistical summaries are calculated for each (fixed) data quality assessment sample solution concentration.

Let n_j = the number of data quality assessments at a given concentration in the jth lot.

$\bar{X}_j$ = the average of the data quality assessments at a given concentration for the jth lot.

s_j = the standard deviation of the data quality assessments at a given concentration for the jth lot.

$n = \sum_{J=1}^{k} n_j$ = total number of data quality assessments for the reporting period, where k = total number of lots.

Then,

The grand weighted average is:

$$\bar{X} = \sum_{j=1}^{k} \frac{n_j}{n} \bar{X}_j \qquad \text{Eq. 5-8}$$

TABLE 5-2. MEASUREMENTS AND WITHIN-LOT SUMMARY STATISTICS FOR
SULFATE DATA QUALITY ASSESSMENT SOLUTIONS

DATE OF OPERATION	T = 0.0 CU[a]				T = 120 CU[a]				T = 240 CU[a]			
	MEASURE-MENT	n	$\bar{X}$	s_w	MEASURE-MENT	n	$\bar{X}$	s_w	MEASURE-MENT	n	$\bar{X}$	s_w
7/16/80	0.00 1.75	2	0.88	1.24	119.8	1	119.8	-	241.0 241.5 238.8	3	240.4	1.44
7/17/80	0.50 1.00	2	0.75	0.35					239.8 239.5 238.0 239.8 238.8 239.5	6	239.2	0.71
7/18/80	2.25	1	2.25	-	118.8	1	118.8	-	241.0 241.5 238.8	3	240.4	1.44
7/31/80	0.00	1	0.00	-					241.2 243.8 236.5	3	240.5	3.70
8/18/80					124.3 119.0 119.3	3	120.9	2.98				
10/6/80	0.00	1	0.00	-					240.2 238.8 239.8	3	239.6	0.72
10/27/80									240.5 236.5 230.5	3	235.8	5.03
10/28/80					123.5 117.8 120.5	3	120.6	2.85				
10/29/80					119.8 117.5	2	118.6	1.63				
11/12/80					127.2 114.2 123.8	3	121.7	6.74				
11/13/80					130.2 113.5	2	121.8	11.81				

[a] CU - concentration unit

The average within-lot standard deviation is:

$$s_{\bar{w}} = \sqrt{\sum_{j=1}^{k} \frac{(n_j-1)s_j^2}{n-k}} \qquad \text{Eq. 5-9}$$

The between-lot (or among lots) standard deviation is:

$$s_a = \sqrt{\frac{\sum_{j=1}^{k} \frac{n_j(\bar{X}_j-\bar{X})^2}{k-1} - s_{\bar{w}}^2}{(n_1 + \ldots + n_k)/k}} \qquad \text{Eq. 5-10}$$

Total variation of single observations at the related concentration level over time is:

$$s_t = \left(s_a^2 + s_{\bar{w}}^2 \right)^{1/2} \qquad \text{Eq. 5-11}$$

Between-lot summary statistics for the sulfate sample data are given in Table 5-3. Example calculations of between-lot summary statistics for Condition 1, where T = 240 CU, k = 6, and n = 21 are given below:

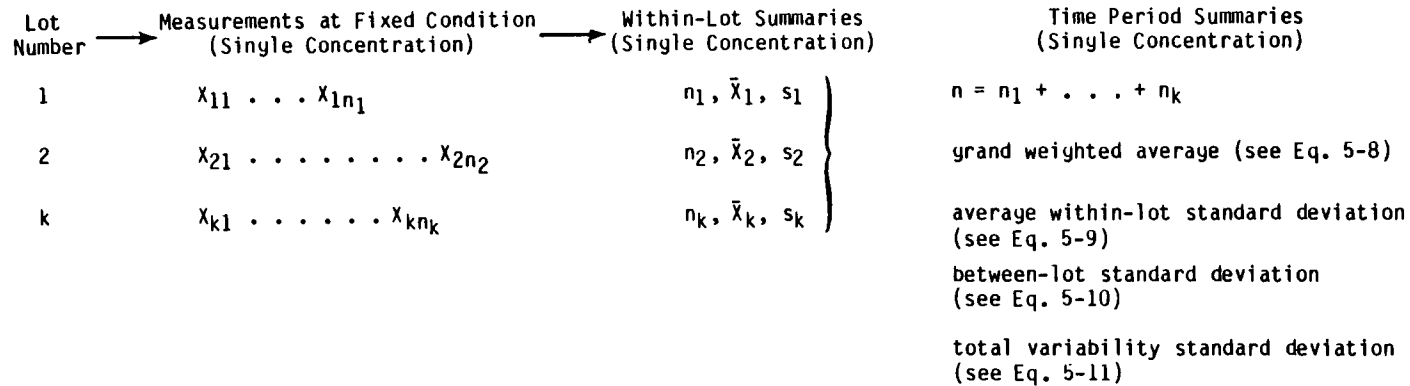$$\bar{X} = \sum_{j=1}^{6} \frac{n_j}{21} (\bar{X}_j)$$

$$= \frac{3}{21} (240.4) + \frac{6}{21} (239.2) + \frac{3}{21} (240.4) +$$

$$\frac{3}{21} (240.5) + \frac{3}{21} (239.6) + \frac{3}{21} (235.8)$$

$$\bar{X} = 239.3 \text{ CU}$$

$$s_{\bar{w}} = \sqrt{\sum_{j=1}^{6} \frac{(n_j-1) s_j^2}{21 - 6}}$$

$$= \left[\frac{2(1.44)^2+5(0.71)^2+2(1.44)^2+2(3.70)^2+2(0.72)^2+2(5.03)^2}{15} \right]^{1/2}$$

TABLE 5-3. SUMMARY STATISTICS FOR SULFATE DATA QUALITY ASSESSMENTS SOLUTIONS AT FIXED CONDITIONS, CONCENTRATION KNOWN

Condition	True Value T	Number of Lots k With Data Quality Assessments	Total Number n of Samples at the Condition Over All Lots	Grand Average $\bar{X}$ Over k Lots of Measurements at the Condition	Bias $\bar{X} - T$	Average Within-Lot Standard Deviation s_W	Between-Lot Std. Dev. s_a	Total Variability s_t
1	240	6	21	239.3	-0.7	2.45	1.03	2.66
2	120	7	15	120.6	.6	5.78	0.00	5.78
3	0.0	5	7	0.79	.79	0.91	0.12	0.92
4								
5								
6								
7								
8								
9								
10								

Lot Number $\longrightarrow$	Measurements at Fixed Condition (Single Concentration) $\longrightarrow$	Within-Lot Summaries (Single Concentration)	Time Period Summaries (Single Concentration)
1	$X_{11} \ldots X_{1n_1}$	$n_1, \bar{X}_1, s_1$	$n = n_1 + \ldots + n_k$
2	$X_{21} \ldots \ldots \ldots X_{2n_2}$	$n_2, \bar{X}_2, s_2$	grand weighted average (see Eq. 5-8)
k	$X_{k1} \ldots \ldots X_{kn_k}$	$n_k, \bar{X}_k, s_k$	average within-lot standard deviation (see Eq. 5-9)

between-lot standard deviation (see Eq. 5-10)

total variability standard deviation (see Eq. 5-11)

$$s_{\bar{w}} = 2.45 \ CU$$

$$s_a = \sqrt{\dfrac{\displaystyle\sum_{j=1}^{6} \dfrac{n_j(\bar{X}_j - \bar{X})^2}{6-1} - s_{\bar{w}}^2}{(n_1 + \ldots + n_6)/6}}$$

$$= \left[\dfrac{\dfrac{\begin{array}{l} 3(240.4-239.3)^2 + 6(239.2-239.3)^2 + 3(240.4-239.3)^2+ \\ 3(240.5-239.3)^2 + 3(239.6-239.3)^2 + 3(235.8-239.3)^2 \end{array}}{5} - 2.45^2}{(3 + 6 + 3 + 3 + 3 + 3)/6} \right]^{1/2}$$

$$s_a = 1.03 \ CU$$

$$s_t = \sqrt{s_a^2 + s_{\bar{w}}^2} = 2.66 \ CU$$

The summary statistics should be reported for each set of sample conditions over the appropriate time period using the format of Table 5-3. (The value $s_a = 0$ at true value 120 resulted from a negative value for s_a^2 indicating that there was not significant variation from lot to lot.)

5.6.3 Regression Summaries and Reporting Requirements for a Project or Time Period

Any measure of bias (e.g., recovery) or precision can depend on the concentration T. Even when T is unknown, dependence of precision on T may be seen as dependence on the apparent (measured) concentration X. Regression summaries of bias and precision over all conditions provide a useful complement to summary statistics for individual sets of conditions. Further, in a case where all the reference values, T, of data quality assessment samples for a given reporting period are different, then between-lot statistical summaries for specific conditions in the Table 5-3 format would be equivalent to lists of the raw assessment data. A much more useful summary involves a regression and plot of measurement X on reference value T.

Table 5-4 contains summary regressions for the sulfate.analysis data of Table 5-2. Associated data plots are shown in Figures 5-7 through 5-9.

Data from analysis of the external filter samples are tabulated in Table 5-5. The summary regression data and plot of $\bar{X}$ versus T are shown in Table 5-6 and Figure 5-10, respectively.

Regression summary statistics should be reported over the appropriate time period using the format of Tables 5-4 and 5-6.

The appropriate rows of each applicable table (i.e. summary statistics and regression summaries) should be completed for each measurement system. If more than ten known concentrations are employed then the ten most frequently used concentrations should be reported in the format of Table 5-3. The summary statistics table can also be used to report on individual lots.

If data quality assessment results are available for only one or two concentrations then the regression summaries are inappropriate and the summary statistics should be limited to individual conditions as in Table 5-3. For three or more concentrations the applicable regression summaries should be reported as well. For bias, either the within-lot averages or grand average for the reporting period can be used as the dependent variable. A similar option is left open for precision in the second row of the summary regression report form (Tables 5-4 and 5-6). The option selected should be indicated by underlining.

In those instances where there is no control over sample concentration T, and thus, calculation of between-lot summary statistics is not possible, a plot and regression of within-lot standard deviation s_j on within-lot average $\bar{X}_j$ is recommended (bottom row in Tables 5-4 and 5-6). In such a case only the last row of the summary regression form is completed.

The second and fourth rows of the summary regression form (Tables 5-4 and 5-6) involve regression of within-lot (average) standard deviation on either the reference value T or the replicate average $\bar{X}_j$. If T is known and only one of rows 2 and 4 is to be completed, then row 2 should be completed.

TABLE 5-4. SUMMARY REGRESSIONS FOR SULFATE ANALYSIS OF DATA FROM PREPARED SOLUTIONS, REPORTING PERIOD JULY, 1980 THROUGH DECEMBER, 1980

Type of Data Quality Indicator	Sample Type	Dependent Variable v[a]	Independent Variable u	Number of (u, v) pairs	slope	intercept
Bias	sulfate ion in solution	lot average, $\bar{X}_j$ or grand average $\bar{X}$	true value T	3	.99	0.98
Precision	'	within-lot std. dev. s_j or average within-lot std. dev. $\bar{s}_w$	true value T	3	.006	2.28
Precision	'	between-lot std. dev. s_a	true value T	3	.004	-0.08
Precision[b]		within-lot std. dev. s_j (primarily for split samples or repeat analysis of routine samples)	$\bar{X}_j$			

[a] Underlined variable is illustrated in this example.
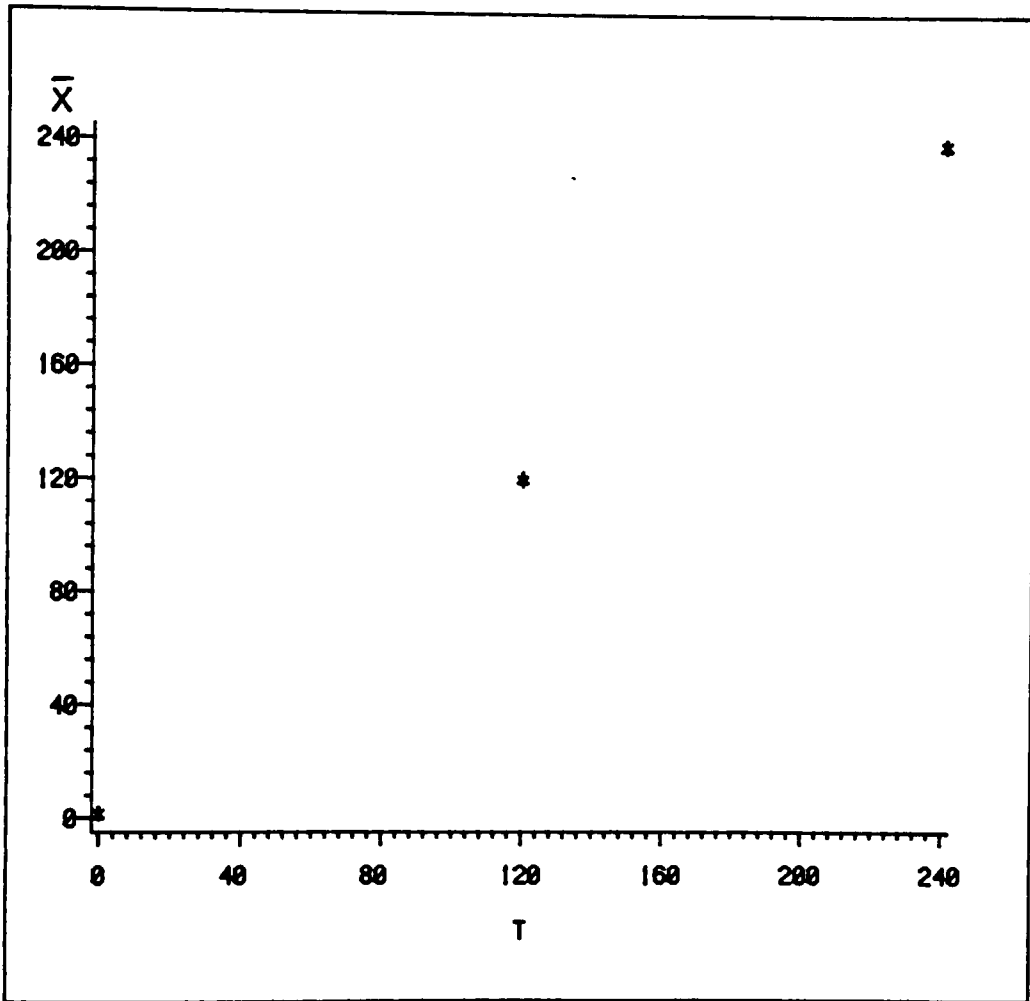
[b] This row not applicable for this data set.

Figure 5-7. Plot of $\bar{X}$ (CU) versus T (CU) for Sulfate Data Quality
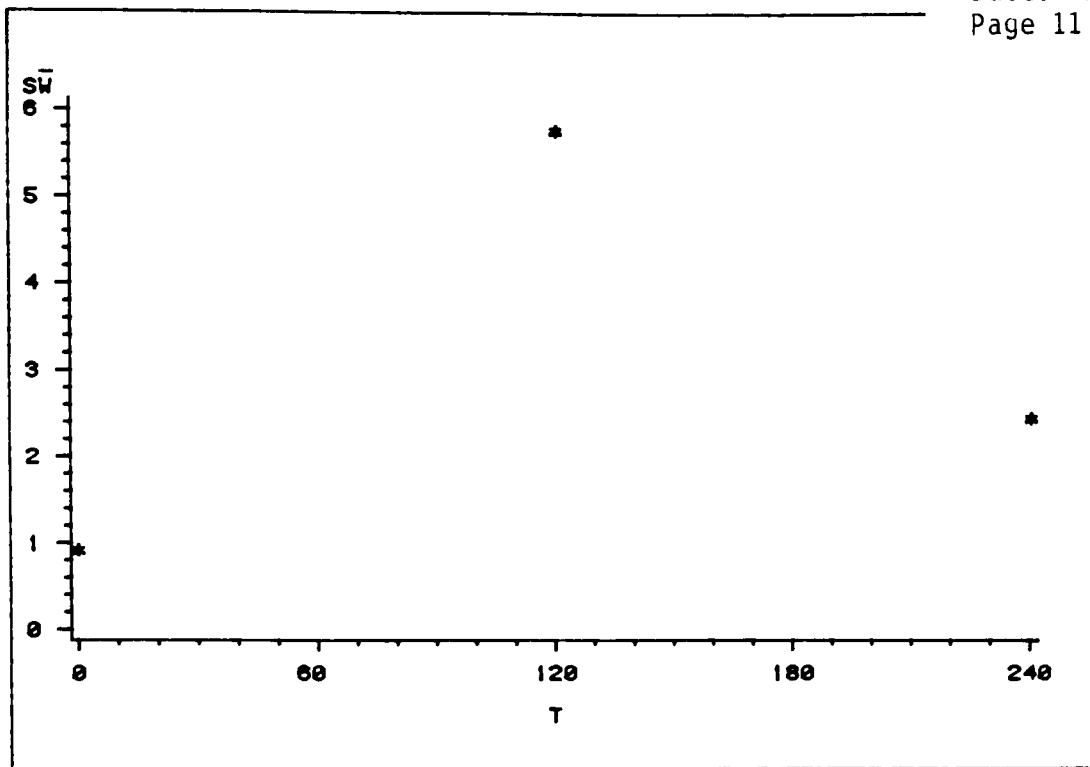Assessment Samples (Table 5-3).

Figure 5-8. Plot of Within-Lot Standard Deviation $\bar{s_w}$ (CU) versus T
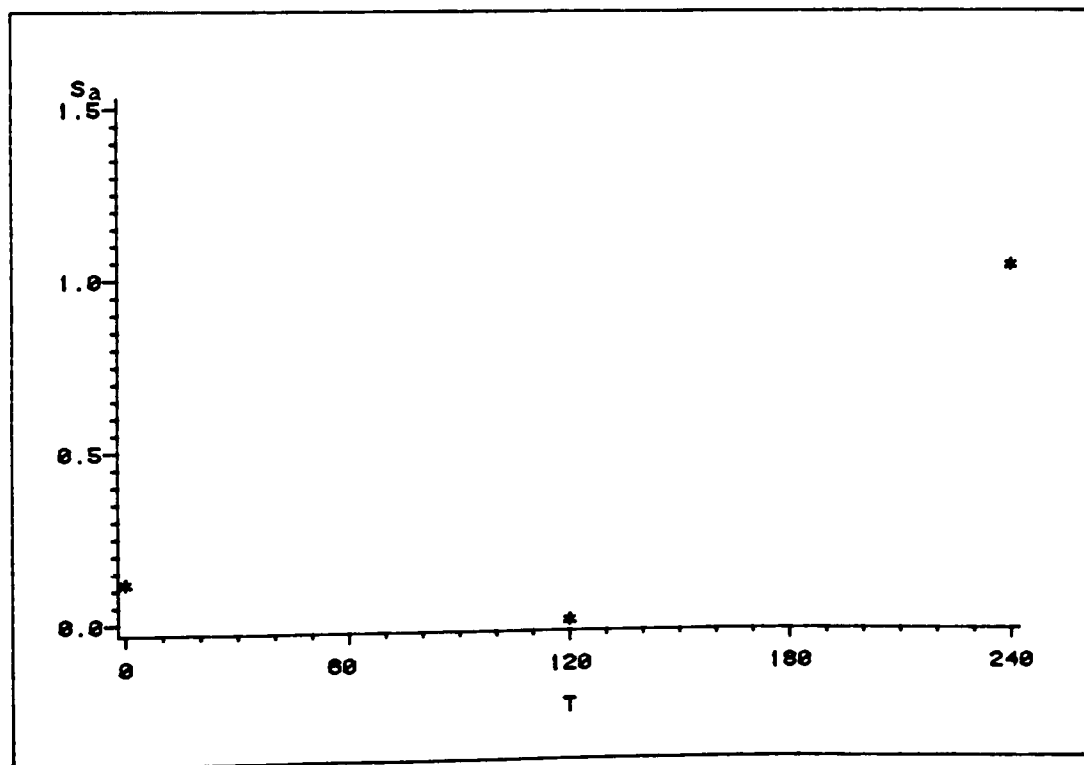(CU) for Sulfate Data Quality Assessment Results (Table 5-3).

Figure 5-9. Plot of Between-Lot Standard Deviation s_a (CU) versus T
(CU) for Sulfate Data Quality Assessment Results (Table 5-3).

TABLE 5-5. MEASUREMENT DATA FOR SULFATE FILTER SAMPLES

X	T	P[a]	X	T	P[a]	X	T	P[a]
152	190	80	225	222	101	98	102	96
265	272	97	79	83	95	116	114	102
50	54	93	157	169	93	199	201	99
132	141	94	59	63	94	159	169	94
35	34	103	127	127	100	267	292	91
223	265	84	147	154	95	158	154	103
34	34	100	227	231	98	29	27	107
177	222	80	82	83	99	112	114	98
110	102	108	192	222	86	228	250	91
178	190	94	54	54	100	194	201	97
257	265	97	266	272	98			

[a] P = Percent Recovery = 100 X/T (See Section 5.4 for definition)

TABLE 5-6. SUMMARY REGRESSIONS FOR SULFATE ANALYSIS OF FILTER DATA,
REPORTING PERIOD AUGUST, 1980 TO DECEMBER, 1980

Type of Data Quality Indicator	Sample Type	Dependent Variable v^a	Independent Variable u	Number of (u, v) pairs	slope	intercept
$\bar{X}$	sulfate ion on teflon filter	lot average, $\bar{X}_j$ or grand average $\bar{X}$	true value T	32	0.91	5.80
Bias					-0.09	5.80
Precision[b]		within-lot std. dev. s_j or average within-lot std. dev. $\bar{s}_w$	true value T			
Precision[b]		between-lot std. dev. s_B	true value T			
Precision[b]		within-lot std. dev. s_j (primarily for split samples or repeat analysis of routine samples)	$\bar{X}_j$			

[a] Underlined variable is illustrated in this example.

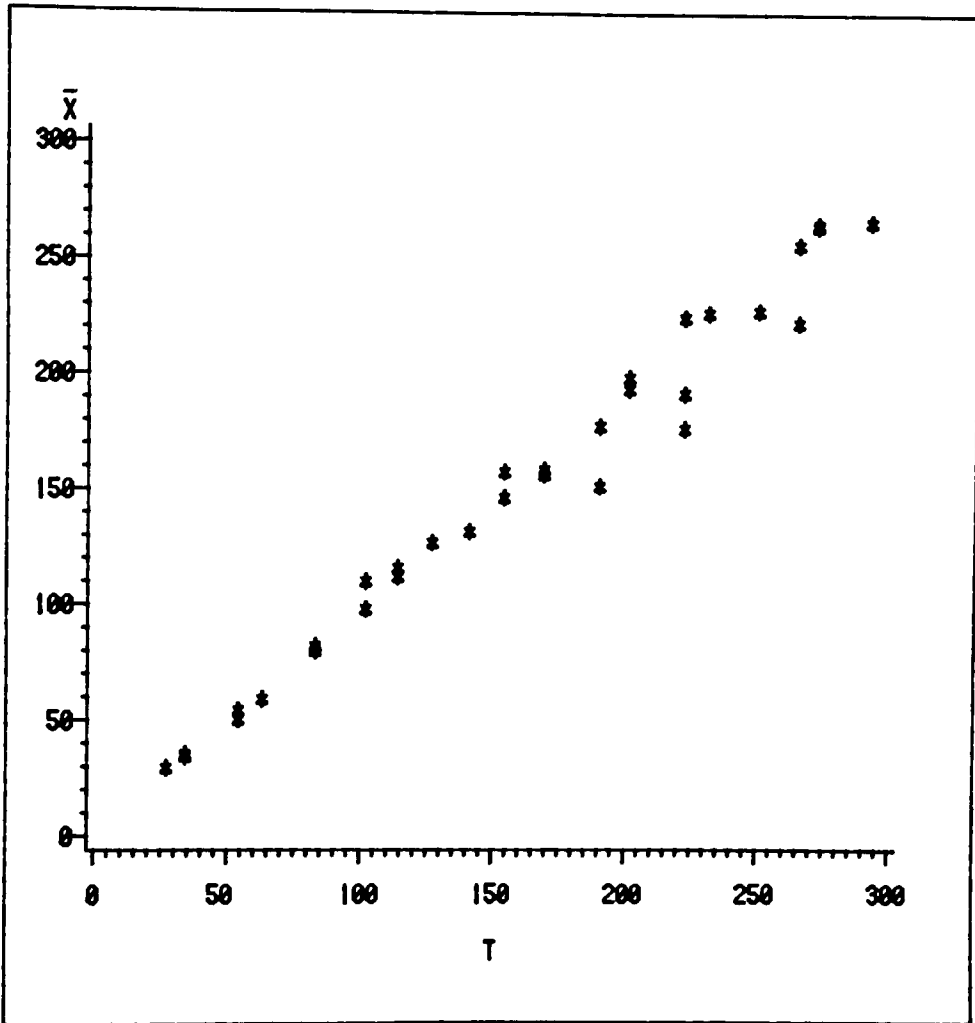[b] This row not applicable for this data set.

Figure 5-10. Plot of $\bar{X}$ (CU) versus T (CU) for Sulfate Data
Quality Assessment Results Reported from August,
1980, to December, 1980.

5.7   SOURCES OF ADDITIONAL INFORMATION

5.7.1   Study Planning

1.   "Compilation of Data Quality Information for Environmental Measurement Systems," QAMS _____, U.S. EPA, Office of Research and Development, Washington, DC, 1983 (Draft).

2.   "Interim Guidelines and Specifications for Preparing Quality Assurance Project Plans, "QAMS-005/80. U.S. EPA, Office of Research and Development, Washington, D.C. 20560, December, 1980.

3.   Natrella, M.G., Experimental Statistics, NBS Handbook 91, U.S. Department of Commerce, National Bureau of Standards, 1966.

4.   Davies, O. L., The Design and Analysis of Industrial Experiments, 2nd edition, Hafner Publishing Co., New York, 1956.

5.   Cox, D.R., Planning of Experiments, Wiley, New York, 1958.

6.   Box, G.E.P., W.G. Hunter and J.S. Hunter, Statistics for Experimenters, Wiley, New York, 1978.

7.   Youden, W.J., "Statistical Aspects of Analytical Determinations, "Journal of Quality Technology, 4(1), 1972, pp. 45-49.

8.   Elder, R.S., "Choosing Cost-Effective QA/QC Programs for Chemical Analysis," EPA Contract No. 68-03-2995, Radian Corporation, Austin, Texas, 1981 (draft).

5.7.2   Sampling

1.   Environmental Monitoring and Support Laboratory, Handbook for Sampling and Sample Preservation of Water and Wastewater, EPA-600/4/82-029, U.S. EPA, Office of Research and Development, Cincinnati, 1982.

2.   Brumbaugh, M.A., "Principles of Sampling in the Chemical Field," Industrial Quality Control, January 1954, pp. 6-14.

3.   Kratochvil, B. and J.K. Taylor, "Sampling for Chemical Analysis," Analytical Chemistry, 53(8), 1981, pp. 928A-938A.

4.   Currie, L.A. and J.R. DeVoe, "Systematic Error in Chemical Analysis", In: Validation of the Measurement Process, ACS Symposium Series 63, American Chemical Society, Washington, D.C., 1977, pp. 114-139.

5.7.3 Assessment of Precision

1. Bennett, C.A. and N.L. Franklin, Statistical Analysis in Chemistry and the Chemical Industry, Wiley. New York, 1954.

2. Rhodes, R.C., "Components of Variation in Chemical Analysis." In: Validation of the Measurement Process, ACS Symposium Series No. 63, American Chemical Society, Washington, D.C. 1977, pp. 176-198.

3. Wilson, A.L., "The Performance Characteristics of Analytical Methods-II," Talanta, 17, 1970, pp. 31-44.

4. Bicking, C.A., "Precision in the Routine Performance of Standard Tests," ASTM Standardization News, January 1979, pp. 12-14.

5. Merten, D., L.A. Currie, J. Mandel, O. Suschny and G. Wernimont, "Intercomparison, Quality Control and Statistics." In: Standard Reference Materials and Meaningful Measurements, NBS Special Publication 408, U.S. Department of Commerce, National Bureau of Standards, 1975, p. 805.

6. Janardan, K.G. and D.J. Schaeffer, "Propagation of Random Error in Estimating the Levels of Trace Organics in Environmental Sources, "Analytical Chemistry, 51(7), 1979, pp. 1024-1026.

7. Bicking, C.A., "Inter-Laboratory Round Robins for Determination of Routine Precision of Methods." In: Testing Laboratory Performance, NBS Special Publication 591, U.S. Department of Commerce, National Bureau of Standards, 1980, pp. 31-34.

8. Wernimont, G., "Use of Control Charts in the Analytical Laboratory, "Industrial and Engineering Chemistry, 18(10), 1946, pp. 587-592.

9. Frazier, R.P., et al., "Establishing a Quality Control Program for a State Environmental Laboratory," Water and Sewage Works, 121(5), 1974, pp. 54-57.

10. Dorsey, N.E. and C. Eisenhart, "On Absolute Measurement." In: Precision Measurement and Calibration, NBS Special Publication 300, U.S. Department of Commerce, National Bureau of Standards, 1969, pp. 49-55.

11. Suschny, O. and D.M. Richman, "The Analytical Quality Control Programme of the International Atomic Energy Agency." In: Standard Reference Materials and Meaningful Measurements, NBS Special Publication 408, U.S. Department of Commerce, National Bureau of Standards, 1975, pp. 75-102.

12. Taylor, J.R., An Introduction to Error Analysis -- The Study of Uncertainties in Physical Measurements, University Sciences Books, Mills Valley, California, 1982 (This is a general text).

5.7.4 Assessment of Bias

1. Uriano, G.A. and C.C. Gravatt, "The Role of Reference Materials and Reference Methods in Chemical Analysis," CRC Critical Reviews in Analytical Chemistry, 6(4), 1977, pp. 361-411.

2. Uriano, G.A. and J.P. Cali, "Role of Reference Materials and Reference Methods in the Measurement Process." In: Validation of the Measurement Process, ACS Symposium Series No. 63, American Chemical Society, Washington, D.C., 1977, pp. 140-161.

3. Skogerboe, R.K. and S.R. Koirtyohann, "Accuracy Assurance in the Analysis of Environmental Samples." In: Accuracy in Trace Analysis, Vol. 1, NBS Special Publication 422, U.S. Department of Commerce, National Bureau of Standards, 1976, 1976, pp. 199-210.

4. Watts, R.R., "Proficiency Testing and Other Aspects of a Comprehensive Quality Assurance Program." In: Optimizing Chemical Laboratory Performance through the Application of Quality Assurance Principles, Association of Official Analytical Chemists, Arlington, VA, 1980, pp. 87-115.

5. Horwitz, W.L., R. Kamps and K.W. Boyer, "Quality Assurance in the Analysis of Foods for Trace Constituents," Journal of the Association of Official Analytical Chemists, 63(6), 1980, pp. 1344-1354.

6. Colby, B.N., "Development of Acceptance Criteria for the Determination of Organic Pollutants at Medium Concentrations in Soil, Sediments, and Water Samples," EPA Contract No. 68-02-3656, Systems Science and Software, LaJolla, CA, 1981.

7. Bicking, C., S. Olin and P. King, Procedures for the Evaluation of Environmental Monitoring Laboratories, Tracor Jitco, Inc., EPA-600/4-78-017, U.S. EPA, Office of Research and Development, Environmental Monitoring and Support Laboratory, Cincinnati, 1978.

8. U.S. Department of the Army, "Quality Assurance Program for U.S. Army Toxic and Hazardous Materials Agency," Aberdeen Proving Ground, MD., August 1980 (draft).

9. Freeberg, F.E., "Meaningful Quality Assurance Program for the Chemical Laboratory." In: Optimizing Chemical Laboratory Performance Through the Application of Quality Assurance Principles, Association of Official Analytical Chemists, Arlington, VA, 1980, pp. 13-23.

10. American Society for Testing and Materials, "Standard Practice for Determination of Precision and Bias of Methods of Committee D-19 on Water," ASTM Designation: D2777-77. In: 1977 Annual Book of ASTM Standards, Part 31, pp. 7-19.

11. Frazier, R.P., et al., "Establishing a Quality Control Program for a State Environmental Laboratory." Water and Sewage Works, 121(5). 1974, pp. 54-57.

5.7.5 Use of Control Charts

1. Shewhart, W.A., Economic Control of Manufacture Products, Van Nostrand, New York, 1931.

2. McCully, K.A. and J.G. Lee, "Quality Assurance of Sample Analysis in the Chemical Laboratory." In: Optimizing Chemical Laboratory Performance through the Application of Quality Assurance Principles, Association of Official Analytical Chemists, Arlington, VA, 1980, pp. 57-86.

3. Duncan, A.J., Quality Control and Industrial Statistics, 3rd edition, Richard D. Irwin, Inc., Homewood, IL, 1968.

4. Grant, E.L. and R.S. Leavenworth, Statistical Quality Control, 4th edition, McGraw-Hill, New York, 1972.

5. Environmental Monitoring and Support Laboratory, Handbook for Analytical Quality Control in Water and Wastewater Laboratories, EPA-600/4-79-019, U.S. EPA, Office of Research and Development, Cincinnati, 1979.

6. Wernimont, G., "Use of Control charts in the Analytical Laboratory, Industrial and Engineering Chemistry, 18(10), 1946, pp. 587-592.

7. Bennett, C.A. and N.L. Franklin, Statistical Analysis in Chemistry and the Chemical Industry, Wiley, New York, 1954.

8. Eisenhart, C., "Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems." In: Precision Measurement and Calibration, NBS Special Publication 300, U.S. Department of Commerce, National Bureau of Standards, 1969, pp. 21-47.

9. Wernimont, G., "Statistical Control of the Measurement Process." In: Validation of the Measurement Process, ACS Washington, D.C., 1977, pp. 1-29.

10. Moore, P.G., "Normality in Quality Control Charts," Applied Statistics, 6(3), 1957, pp. 171-179.

11. Morrison, J., "The Lognormal Distribution in Quality Control," Applied Statistics, 7(3), 1958, pp. 160-172.

12.   Iglewicz, B. and R.H. Myers, "Comparison of Approximations to the Percentage Points of the Sample Coefficient of Variation," _Technometrics_, 12(1), 1970, pp. 166-170.

13.   Environmental Monitoring and Support Laboratory, _Quality Assurance Handbook for Air Pollution Measurement Systems, Volume I - Principles_, EPA-600/9-76-005, U.S. EPA, Office of Research and Development, Research Triangle Park, NC, 1976.

14.   Grubbs, F.E. "The Difference Control Chart with an Example of Its Use," _Industrial Quality Control_, July, 1946, pp. 22-25.

15.   Page, E.S., "Cumulative Sum Charts," _Technometrics_, 3(1), 1961, pp. 1-9.

16.   Jackson, J.E., "Quality Control Methods for Several Related Variables, "_Technometrics_, 1(4), 1959, pp. 359-377.

17.   Jackson, J.E. and R.H. Morris, "An Application of Multivariate Quality Control to Photographic Processing," _Journal of the American Statistical Association_, 52, 1957, pp. 186-199.

18.   Montgomery, D.C. and H.M. Wadsworth, "Some Techniques for Multivariate Quality Control Applications," _ASQC Technical Conference Transactions_, 1972.

19.   Frazier, R.P., J.A. Miller, J.F. Murray, M.P. Mauzy, D.J. Schaeffer and A.F. Westerhold, "Establishing a Quality Control Program for a State Environmental Laboratory," _Water and Sewage Works_, 121(5), 1974, pp. 54-57.

20.   Hillier, F.S., "X and R-Chart Control Limits Based on a Small Number of Subgroups, "_Journal of Quality Technology_, 1(1), 1969, pp. 17-26.

5.7.6 Method Detection Limit

1.   Glaser, J.A., D.L. Foerst, G.D. McKee, S.A. Quave, W.L. Budde, "Trace Analysis for Wastewaters," _Environmental Science and Technology_, 15, 1981, pp. 1425-1435.

2.   Hubaux, A. and G. Vos, "Decision and Detection Limits for Linear Calibration Curves," _Analytical Chemistry_, 42, 1970, pp. 849-855.

3.   "Guidelines for Data Acquisition and Data Quality Evaluation in Environmental Chemistry," _Analytical Chemistry_, 52, 1980, pp. 2242-2249.

4. Currie, L.A., "Limits for Qualitative Detection and Quantitative Determination - Application to Radiochemistry," Analytical Chemistry, 40, 1968, pp. 586-594.

5. Ramirez-Munoz, J., "Qualitative and Quantitative Sensitivity in Flame Photometry," Talanta, 13, 1966, pp. 87-101.

6. Parsons, M.L., "The Definition of Detection Limits," Journal of Chemical Education, 46, 1969, pp. 290-292.

7. Ingle, J.D., Jr., "Sensitivity and Limit of Detection in Quantitative Spectrometric Methods," Journal of Chemical Education, 51, 1974, pp. 100-105.

8. Wilson, A.L., "The Performance Characteristics of Analytical Methods - III," Talanta, 20, 1973, pp. 725-732.

9. Kaiser, H., "Guiding Concepts Relating to Trace Analysis," Pure and Applied Chemistry, 34, 1973, pp. 35-61.

10. Liteanu, C. and I. Rica, "Statistical Theory and Methodology of Trace Analysis," John Wiley and Sons, 1980.

5.8 A GLOSSARY OF TERMS

Arithmetic mean
  (μ for populations,
  $\bar{X}$ for samples)

In a sample of n units, X_1, X_2 . . . . . X_n, the sum of the observed values in the sample divided by the number of units in the sample.

Bias

The difference between the population mean and the true or reference value, or as estimated from sample statistics, the difference between the sample average and the reference value.

Coefficient of variation

A measure of relative dispersion. It is equal to the standard deviation divided by the mean and multiplied by 100 to give a percentage value.

Correlation Coefficient

A number between -1 and 1 that indicates the degree of linear relationship between two sets of numbers.

Error

The difference between an observed value and its true value or the probability interval that contains the systematic and and random error with $1-\sigma$ confidence.

Lot, batch

A definite quantity of samples collected under conditions that are considered uniform.

Lot size (N)

The number of units in a particular lot.

Matrix

The material in which the analyte(s) of primary interest is embeded.

Method Detection Limit (MDL)

The lowest concentration of an analyte that a measurement system can "consistently detect" and/or measure in replicated field samples.

Observed value, observation, or variate (X)

The particular value of a characteristic and designated X_1, X_2, X_3, and so on.

Parameter

A constant or coefficient that describes some characteristic of a population (e.g., standard deviation, mean, regression coefficients).

Precision

Degree of mutual agreement among individual measurements made under prescribed conditions.

Quality

The totality of features and characteristics of a product or service that bear on its ability to satisfy given needs.

Quality Assurance

A system of activities whose purpose is to provide adequate confidence that a product or service will satisfy given needs.

Quality Control

The operational techniques and the activities which are aimed at maintaining a product or a service at a level of quality that will satisfy given needs.

Range (R)

The difference between the largest and smallest numbers in a set of $\underline{n}$ numbers.

Range, Relative (RR)

The range divided by the mean of a particular set of numbers.

Regression coefficients

The quantities describing the slope and intercept of a regression line.

Intercept (β_0 for populations, b_0 for samples),
Slope (β_1 for populations, b_1 for samples).

Regression line or equation

The function that indicates the regression relationship. For example, $X = \beta_0 + \beta_1 T$ for populations, and $X = b_0 + b_1 T$ for samples.

Sample (statistical)

A group of samples (chemical) taken from a lot or batch of samples (chemical).

Sample average

Same as arithmetic mean.

Sample size (n)

The number of units in a sample.

Standard deviation (σ)	A measure of the dispersion about the mean of the elements in a **population.**
Standard deviation(s)	A measure of the dispersion about the average of the elements in a **sample.** An estimate of the standard deviation of a population.
Statistic	A constant or coefficient that describes some characteristic of a sample. Statistics are used to estimate parameters of populations.
Universe or population	The totality, finite or infinite, of a set of items, units, elements, measurements, and the like, real or conceptual, that is under consideration.

REGRESSION SUMMARY STATISTICS FORM

Type of Data Quality Indicator	QA/QC Sample Type	Dependent Variable v	Independent Variable u	Number of (u, v) pairs	slope	intercept
Bias		lot average, $\bar{X}_j$ or grand average $\bar{X}$ (underline one)	true value T			
Precision		within-lot std. dev. s_j or average within-lot std. dev. $\bar{s}_w$ (underline-one)	true value T			
Precision		between-lot std. dev. s_a	true value T			
Precision		within-lot std. dev. s_j (primarily for split samples or repeat analysis of routine samples)	$\bar{X}_j$			

BETWEEN-LOT SUMMARY STATISTICS FORM

Condition	True Value T	Number of Lots	Number of Lots k With QC Measurements	Total Number n of Samples at the Condition Over All Lots	Grand Average $\bar{X}$ Over Lots of Measurements at the Condition	Bias $\bar{X} - T$	$s_{\bar{w}}$ [a]	s_a [b]	s_t [c]
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									

[a] Average within-lot standard deviation.

[b] Between-lot (among-lot) standard deviation.

[c] Total variability standard deviation.