United States
Environmental Protection
Agency

Office of Air Quality
Planning and Standards
Research Triangle Park NC 27711

EPA 450/4-90-015

April 1991

Air

EPA

Protocol For The Field Validation Of Emission Concentrations From Stationary Sources

EPA 450/4-90-015
April 1991

PROTOCOL FOR THE FIELD VALIDATION OF EMISSION CONCENTRATIONS FROM STATIONARY SOURCES

Technical Support Division
Office of Air Quality Planning and Standards

and

Quality Assurance Division
Atmospheric Research and Exposure Assessment Laboratory

U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

PROTOCOL FOR THE FIELD VALIDATION OF EMISSION CONCENTRATIONS FROM STATIONARY SOURCES

1.0 APPLICABILITY

1.1 This "Protocol for the Field Validation of Emission Concentrations from Stationary Sources" (Protocol) includes procedures for determining and documenting the quality, i.e., systematic error (bias) and random error (precision), of the measured concentrations of the source emissions. This protocol, as specified in the underlying regulations, is to be used whenever a source owner or operator (hereafter referred to as an "analyst") proposes a test method to meet a U.S. Environmental Protection Agency (EPA) requirement in the absence of a validated method. For example, the Protocol may be used to identify and verify post-control emissions for early reduction credit [Section 112(5)(i) of the Clean Air Act Amendments of 1990].

1.2 If EPA currently recognizes an appropriate test method or considers the analyst's test method to be satisfactory for a particular source, the Administrator may waive the use of this protocol or may specify a less rigorous validation procedure. A list of **validated methods** can be obtained by contacting the Emission Measurement Technical Information Center (EMTIC), Mail Drop 19, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, 919/541-2237. Procedures for obtaining a waiver from the protocol are in Appendix A. State agencies may require additional documentation regarding the quality of the emission data.

1.3 This protocol includes optional procedures that may be used to expand the applicability of the proposed method. Appendix B is the "Ruggedness Testing (Laboratory Evaluation)," which demonstrates the sensitivity of the method to various parameters. Appendix C is "Procedure for Including Sample Stability in Bias and Precision," for assessing sample recovery and analysis times; Appendix D is "Procedure for the Determination of Practical Limit of Quantitation" for determining the lower limit of the method.

2.0 PRINCIPLE

The purpose of the validation protocol is to determine bias and precision of a test method at a permissible emission concentration, e.g., emission standard, in the gas stream. The procedures involve (a) introducing known concentrations of an analyte or comparing the test method against a validated test method to determine the method's bias and (b) collecting multiple

or collocated simultaneous samples to determine the method's precision.

2.1 Bias is any systematic positive or negative difference between the measured value and true value of a sample. Three common causes of bias are (a) interfering compounds in the effluent gas, (b) calibration errors, and (c) inefficiencies in the collection of the analyte. Bias is established by comparing the method's results against a reference value and may be eliminated by dividing the measured concentration by an appropriate factor (i.e., average measured concentration/reference value). An offset bias may be handled accordingly. Methods that have bias correction factors outside 0.7 to 1.3 are unacceptable. Method to method comparisons, Section 6.2, requires a more restrictive test of central tendency and a lower correction factor allowance of 0.90 to 1.10.

2.2 Precision is the variability in the data obtained from the entire measurement system (i.e., sampling and analysis) as determined from collocated sampling trains. At least two (i.e., paired) sampling trains shall be used to establish precision. The precision of the method at the level of the emission standard shall not be greater than 50 percent relative standard deviation. For method to method equivalency comparisons the analyst must demonstrate that the precision of the proposed test method is as good as that of the validated method for acceptance.

3.0 REFERENCE MATERIAL

The analyst shall obtain a known concentration of the reference material (i.e., analyte of concern) from an independent source such as a specialty gas manufacturer, specialty chemical company, or commercial laboratory. A list of vendors may be obtained from EMTIC (see Section 1.2). The analyst should obtain the manufacturer's stability data of the analyte concentration and recommendations for recertification. If the reference material is in the gaseous state, the concentration(s) [multiple levels may be required to expand a method's concentration applicability] shall be within 0.20 to 5 times the average concentration in the sample gas stream.

3.1 Surrogate Reference Materials. The analyst may use surrogate compounds, e.g., for highly toxic or reactive organic compounds, provided the analyst can demonstrate to the Administrator's satisfaction that the surrogate compound behaves as the analyte. A surrogate may be an isotope or one that contains a unique element (e.g., chlorine) that is not present in the stack gas or a derivative of the toxic or reactive compound, if the derivative formation is part of the method's procedure. Laboratory experiments or literature data may be used to show behavioral acceptability.

3.2 Isotopically Labeled Materials. Isotope mixtures may contain the isotope and the natural analyte. For best results, the isotope labeled analyte concentration should be more than five times the natural concentration of the analyte. Deuterated compounds of interest may be in gaseous or liquid states. The gaseous form should be obtained in compressed-gas cylinders in high-purity nitrogen.

4.0 EPA PERFORMANCE AUDIT MATERIAL

4.1 To assess the method bias independently, the analyst shall use (in addition to the reference material) an EPA performance audit material, if it is available. The analyst may contact EMTIC (see Section 1.2) to receive a list of currently available EPA audit materials. If the analyte is listed, the analyst should request the audit material at least 30 days before the validation test. If an EPA audit material is not available, request documentation from the validation report reviewing authority that the audit material is currently not available from EPA. Include this documentation with the field validation report.

4.2 The analyst shall sample and analyze the performance audit sample three times according to the instructions provided with the audit sample. The analyst shall submit the three results with the field validation report. Although no acceptance criteria are set for these performance audit results, the analyst and reviewing authority may use them to assess the relative error of sample recovery, sample preparation, and analytical procedures and then consider the relative error in evaluating the measured emissions.

5.0 PROCEDURE FOR DETERMINATION OF BIAS AND PRECISION IN THE FIELD

The analyst shall select one of the sampling approaches below to determine the bias and precision of the data. After analyzing the samples, the analyst shall calculate the bias and precision according to the procedure described in Section 6.0.

5.1 Isotopic Spiking. This approach shall be used only for methods that require gas chromatography/mass spectrometry (GC/MS) analysis. Bias and precision are calculated by procedures described in Section 6.1.

5.1.1 Number of Samples and Sampling Runs. Collect a total of 12 samples using either paired (2) or quadruplet (4) collocated sampling trains. For paired trains, conduct six sampling runs. For quadruplet trains, conduct three sampling runs.

5.1.2 Spiking Procedure. Spike all 12 sampling trains with the reference material as follows. The spike shall be introduced as close to the tip of the probe as possible.

5.1.2.1 Gaseous Reference Material with Sorbent or Impinger Trains. Sample the reference material (in the laboratory or in the field) at a concentration equal to the level of the emission standard for the time required by the method, and then sample the gas stream for an equal amount of time. The time for sampling both the reference material and gas stream should be equal; however, the time should be adjusted to avoid sorbent breakthrough.

5.1.2.2 Gaseous Reference Material with Sample Container (Bag or Canister). Spike the containers after completion of the test run with an amount equal to the level of emission standards. The final concentration of the reference material shall approximate the level of the emission standard. The volume amount of reference material shall be less than 10 percent of the sample volume.

5.1.2.3 Liquid and Solid Reference Material with Sorbent or Impinger Trains. Spike the trains with an amount equal to the level of the emission standard before sampling the stack gas. The spiking should be done in the field; however, it may be done in the laboratory.

5.1.2.4 Liquid and Solid Reference Material with Sample Container (Bag or Canister). Spike the containers at the completion of each test run with an amount equal to the level of the emission standard.

5.2 Comparison Against a Validated Test Method. Bias and precision are calculated using the procedures described in Section 6.2. This approach shall be used when a validated method is available and an alternative method is being proposed.

5.2.1 Number of Samples and Sampling Runs. Collect a total of 18 samples using paired trains or 16 samples using quadruplet sampling trains. For paired trains, conduct nine sampling runs. For quadruplet trains, conduct four sampling runs. In each run, the validated test method shall be used to collect and analyze half of the samples.

5.2.2 Performance Audit Exception. Conduct the performance audit as required in Section 4.0 for the validated test method. Conducting a performance audit on the test method being evaluated is recommended.

5.2.3 Probe Placement and Arrangement. The probes should be placed in the same horizontal plane. For paired sample probes the arrangement should be that the probe tip is 2.5 cm from the

outside edge of the other with the pitot tube on the outside of each probe. For quad probes, the tips shall be in a 6.0 cm x 6.0 cm square area measured from the inside edge of the probe tip with the pitot tube in the center.

5.3 Analyte Spiking. Bias and precision are calculated using the procedures described in Section 6.3.

5.3.1 Number of Samples and Sampling Runs. Collect a total of 24 samples using quadruplet sampling trains. Conduct six sampling runs.

5.3.2 In each run, spike half of the sampling trains (two out of the four) according to the applicable procedure in Sections 5.1.2.1 through 5.1.2.4.

6.0 CALCULATION OF BIAS AND PRECISION

Data resulting from the procedures specified in Section 5.0 shall be treated as follows to determine bias correction factors, relative standard deviations, and data acceptance. Example calculations are provided in Appendix E.

6.1 Isotopic Spiking. Analyze the data for isotopic spiking tests as outlined in Sections 6.1.1 through 6.1.6.

6.1.1 Calculate the numerical value of the bias using the results from the analysis of the isotopically spiked field samples and the calculated value of the isotopically labeled spike:

$$B = S_m - CS \qquad \text{Eq. 6-1}$$

where:

B = bias at the spike level;
S_m = mean of the measured values of the isotopically spiked samples;
CS = calculated value of the isotopically labeled spike.

6.1.2 Calculate the standard deviation of the S_i values as follows:

$$SD = \sqrt{\frac{\sum (S_i - S_m)^2}{(n-1)}} \qquad \text{Eq. 6-2}$$

where:

S_i = The measured value of the isotopically labeled analyte in the ith field sample;

n = The number of isotopically spiked samples, 12.

6.1.3 Calculate the standard deviation of the mean (SDM) as follows:

$$SDM = \frac{SD}{\sqrt{n}} \qquad \text{Eq. 6-3}$$

6.1.4 Test the bias for statistical significance by calculating the t- statistic,

$$t = \frac{|B|}{SDM} \qquad \text{Eq. 6-4}$$

and compare it with the critical value of the two-sided t-distribution at the 95-percent confidence level and n-1 degrees of freedom. This critical value is 2.201 for the eleven degrees of freedom when the procedure specified in Section 5.1.2 is followed. If the calculated t-value is greater than the critical value the bias is statistically significant and the analyst should procede to evaluate the correction factor.

6.1.5 Calculation of a correction factor. If the t-test does not show that the bias is statistically significant, procede to the precision evaluation. If the method's bias is statistically significant, calculate the correction factor, CF using the following equation:

$$CF = \frac{1}{1 + \frac{B}{CS}} \qquad \text{Eq. 6-5}$$

Multiply all analytical results by CF to obtain the final values.

6.1.6 Calculation of the relative standard deviation (precision). Calculate the relative standard deviation as follows:

$$RSD = \left[\frac{SD}{S_m}\right] \times 100 \qquad \text{Eq. 6-6}$$

where S_m is the measured mean of the isotopically labeled spiked samples.

6.2 Comparison with Validated Method. Analyze the data for comparison with a validated method as outlined in sections 6.2.1 through 6.2.2.5. Conduct the following tests to determine if a proposed method produces results as good as or better than the validated method. Make all necessary bias corrections for the validated method, as appropriate. If the proposed method fails either test, the method results are unacceptable, and conclude that the proposed method is not as good as the validated method. For some highly cyclic emission sources additional precision checks may be necessary. The paired sampling train procedure requires the standard deviation of the validated method to be known. If the standard deviation of the validated method is not available, the paired sampling train procedure shall not be used.

6.2.1 Paired sampling trains.

6.2.1.1 Acceptable precision for equivalency. Determine the acceptance of the proposed method's variance with respect to the variability of the validated method results. If a significant difference is determined, the proposed method and the results are rejected. Proposed methods demonstrating F-values equal to or less than the critical value have acceptable precision.

6.2.1.2 Calculate the variance of the proposed method, S_p^2, and the validated method, S_v^2, using the following equation:

$$S^2 = \sqrt{SD} \qquad \text{Eq. 6-7}$$

Where: SD_v = The standard deviation provided with the validated method;

SD_p = The standard deviation of the proposed method calculated using Equation 6-9a.

6.2.1.3 The F-test. Determine if the variance of the proposed method is significantly different from that of the validated method by calculating the F-value using the following equation:

$$F = \frac{S_p^2}{S_v^2} \qquad \text{Eq. 6-8}$$

Compare the experimental F-value with the critical value of F. The critical value is 1.0 when the procedure specified in Section 5.2.1 for paired trains is followed.

If the calculated F is greater than the critical value, the difference in precision is significant and the data and proposed method are unacceptable.

6.2.1.4 Bias analysis. Test the bias for statistical significance by calculating the t-statistic and determine if the mean of the differences between the proposed method and the validated method is significant at the 80-percent confidence level. This procedure requires the standard deviation of the validated method, SD_v, to be known. Employ the value furnished with the method. If the standard deviation of the validated method is not available, the paired sampling train procedure shall not be used. Determine the mean of the differences, d_m, and the standard deviation, SD_d, of the paired differences, d_i's, using Equation 6-2. Calculate the standard deviation of the proposed method, SD_p, as follows:

$$SD_p = \sqrt{SD_d - SD_v} \qquad \text{Eq. 6-9a}$$

(If $SD_y > SD_d$, let $SD = SD_d/1.414$.) Calculate the value of the t-statistic using the following equation:

$$t = \frac{d_m}{\left(\frac{SD_p}{\sqrt{n}}\right)} \qquad \text{Eq.6-9}$$

where n is the total number of paired samples. For the procedure in Section 5.2.1, n equals nine.

Compare the calculated t-statistic with the corresponding value from the table of the t-statistic. When nine runs are conducted, as specified in Section 5.2.1, the critical value of the t-statistic is 1.397 for eight degrees of freedom. If the calculated t-value is greater than the critical value the bias is statistically significant and the analyst should procede to evaluate the correction factor.

6.2.1.5 Calculation of a correction factor. If the statistical test cited above does not show a significant bias with respect to the reference method, assume that the proposed method is unbiased and use all analytical results without correction. If the method's bias is statistically significant, calculate the correction factor, CF, as follows:

$$CF = \frac{1}{1 + \frac{d_m}{V_m}} \qquad \text{Eq. 6-10}$$

where V_m is the mean of the validated method's values. Multiply all analytical results by CF to obtain the final values.

The method results, and the method, are unacceptable if the correction factor is outside the range of 0.9 to 1.10.

6.2.2 Quadruplet sampling trains.

6.2.2.1 Acceptable precision for equivalency. Determine the acceptance of the proposed method's variance with respect to the variability of the validated method results. If a significant difference is determined the proposed method and the results are rejected.

6.2.2.2 Calculate the variance of the proposed method, S_p^2, and the validated method, S_v^2, using the following equation:

$$S^2 = \frac{\sum d_i^2}{2n} \qquad \text{Eq. 6-11}$$

where the d_i's are the differences between the validated method values and the proposed method values.

6.2.2.3 The F-test. Determine if the variance of the proposed method is significantly different from that of the validated method by calculating the F-value using Equation 6-8. Compare the experimental F-value with the critical value of F. The critical value is 1.0 when the procedure specified in Section 5.2.2 for quadruplet trains is followed.

If the calculated F is greater than the critical value, the difference in precision is significant the results and the proposed method are unacceptable.

6.2.2.4 Bias Analysis. Test the bias for statistical significance at the 80 percent confidence level by calculating the t-statistic. Determine the bias (mean of the differences between the proposed method and the validated method, d_m) and the standard deviation, SD_d, of the differences. Calculate the standard deviation of the mean of the differences, SD_d, using Equation 6-2 where:

$$d_i = \frac{(V_{1_i} + V_{2_i})}{2} - \frac{(P_{1_i} + P_{2_i})}{2} \qquad \text{Eq. 6-12}$$

and: V_{1i} = The first measured value of the validated method in the ith test sample;

P_{1i} = The first measured value of the proposed method in the ith test sample.

Calculate the t-statistic using Equation 6-9 where n is the total number of test sample differences (d_i). For the procedure in Section 5.2.2, n equals four.

Compare the calculated t-statistic with the corresponding value from the table of the t-statistic and determine if the mean is significant at the 80-percent confidence level. When four runs are conducted, as specified in Section 5.2.2, the critical value of the t-statistic is 1.638 for three degrees of freedom.

If the calculated t-value is greater than the critical value the bias is statistically significant and the analyst should procede to evaluate the correction factor.

6.2.2.4 Correction factor calculation. If the method's bias is statistically significant, calculate the correction factor, CF, using Equation 6-10. Multiply all analytical results by CF to obtain the final values. The method results, and the method, are unacceptable if the correction factor is outside the range of 0.9 to 1.10.

6.3 Analyte Spiking. Conduct sampling as described in Section 5.3, and analyze the data for analyte spike testing as outlined in Sections 6.3.1 through 6.3.6.

6.3.1 Calculate the numerical value of the bias using the results from the analysis of the spiked field samples, the unspiked field samples, and the calculated value of the spike:

$$B = S_m - M_m - CS \qquad \text{Eq. 6-13}$$

where

B = absolute bias at the spike level
S_m = mean of the spiked samples
M_m = mean of the unspiked samples
CS = calculated value of the spiked level.

6.3.2 Determine the precision of the spiked samples. Calculate the difference, d_i, between the pairs of the spiked proposed method measurements for each sampling run. Determine the standard deviation (SD_s) of the spiked values using the following equation:

$$SD_S = \sqrt{\frac{\sum d_i^2}{2n}} \qquad \text{Eq. 6-14}$$

where: n = the number of samples.

6.3.3 Calculate the standard deviation of the mean using Equation 6-3.

6.3.4 Test the bias for statistical significance by calculating the t- statistic using Equation 6-4 and comparing it with the critical value of the two- sided t-distribution at the 95-percent

confidence level and n-1 degrees of freedom. This critical value is 2.201 for the eleven degrees of freedom.

6.3.5 Calculation of a correction factor. If the t-test does not show that the bias is statistically significant use all analytical results without correction. If the method's bias is statistically significant, calculate the correction factor using Equation 6-5. Multiply all analytical results by CF to obtain the final values.

6.3.6 Determination of precision of the unspiked samples. Calculate the standard deviation of the unspiked values using Equation 6-14 and the relative standard deviation of the proposed unspiked method using Equation 6-6.

7.0 FIELD VALIDATION REPORT FORMAT

The field validation report shall include a discussion of the regulatory objectives for the testing which describe the reasons for the test, applicable emission limits, and a description of the source. In addition, validation results shall include:

7.1 Summary of the results and calculations shown in Section 6.0.

7.2 Reference material certification and value(s).

7.3 Performance audit results or letter from the reviewing authority stating the audit material is currently not available.

7.4 Laboratory demonstration of the quality of the spiking system.

7.5 Discussion of laboratory evaluations.

7.6 Discussion of field sampling.

7.7 Discussion of sample preparations and analysis.

7.8 Storage times of samples (and extracts, if applicable).

7.9 Reasons for eliminating any results.

8.0 FOLLOWUP TESTING

The correction factor calculated in Section 6.0 shall be used to adjust the sample concentrations in all followup tests conducted at the same source. These tests shall consist of at least three sample collections, and the average shall be used to determine the emission rate. The number of samples per sample collection

period (run) of the method shall be as follows, depending on the validated method precision level:

8.1 Validated relative standard deviation (RSD) ≤ ±15 Percent. One sample per run or three total samples.

8.2 Validated RSD ≤ ±30 Percent. Two samples per run or six total samples.

8.3 Validated RSD ≤ ±50 Percent. Three samples per run or nine total samples.

8.4 Equivalent method. One sample per run or three total samples.

APPENDIX A

PROCEDURE FOR OBTAINING A WAIVER FROM THE VALIDATION PROTOCOL

A.1 INTRODUCTION

The validation protocol may be waived or a less rigorous protocol may be granted for site-specific applications. The following are three example situations for which a waiver may be considered.

A.1.1 "Similar" sources. If the test method has been validated previously at a "similar" source, the validation protocol may be waived provided the requester can demonstrate to the satisfaction of the Administrator that the emission characteristics are "similar." The methods's applicability to the "similar" source may be demonstrated by conducting a ruggedness test as described in Appendix B.

A.1.2 "Documented " methods. In some cases, bias and precision may have been documented through laboratory tests or protocols different from the protocol in this document. If the analyst can demonstrate to the satisfaction of the Administrator that the bias and precision apply to a particular application, the Administrator may waive the entire validation protocol or parts of the validation protocol.

A.1.3 "Conditional" test methods. When the method has been demonstrated to be valid at several sources, the analyst may seek a "conditional" method designation from the Administrator. "Conditional" method status provides an automatic waiver from the protocol provided the method is used within the stated applicability.

A.2 APPLICATION FOR WAIVER

In general, the requester shall provide a thorough description of the test method, the intended application, and results of any validation or other supporting documents. Because of the many potential situations in which the Administrator may grant a waiver, it is neither possible nor desirable to prescribe the exact criteria for a waiver. At a minimum, the requester is responsible for providing the following.

A.2.1 A clearly written test method, preferably in the format of 40 CFR 60, Appendix A Reference Methods. The method must include an applicability statement, concentration range, precision, bias (accuracy), and time in which samples must be analyzed.

A.2.2.2 Summaries (see Section 7.0) of previous validation tests or other supporting documents. If a different protocol from that

described in this document was used, the requester shall provide appropriate documents substantiating (to the satisfaction of the Administrator) the bias and precision values.

A.2.3 Discussion of the applicability statement and arguments for approval of the waiver. This discussion should address as applicable the following: Applicable regulation, emission standards, effluent characteristics, and process operations.

A.3 REQUESTS FOR WAIVER

Each request shall be in writing and signed by the analyst.

A.3.1 "Conditional" Method. Submit requests to Director, OAQPS, Technical Support Division, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711.

A.3.2 "Similar" Source or "Documented" Methods. Submit requests to the appropriate reviewing authority. The appropriate reviewing authority may be identified by contacting the EPA Regional Offices listed in Table A.1.

Table A.1 EPA Regional Offices

Region	Address	States included
I	Air Management Division John F. Kennedy Bldg. Rm 2203 Boston, MA 02203 (617) 565-3245	Maine, Vermont, New Hampshire, Connecticut, Rhode Island, Massachusetts
II	Air and Waste Management Division 26 Federal Plaza Room 900 New York, NY 10278 (212) 264-2517	New York, New Jersey, Puerto Rico, Virgin Islands
III	Air Management Division 841 Chestnut Bldg. Philadelphia, PA 19107 (215) 597-3989	Pennsylvania, West Virginia, Virginia, Delaware, Maryland
IV	Air Management Division 345 Courtland St., N.E. Atlanta, GA 30365 (404) 347-2904	Alabama, Georgia, Kentucky, Tennessee, Mississippi, Florida, North Carolina, South Carolina
V	Air and Radiation Division 230 S. Dearborn St. Chicago, IL 60604 (312) 353-2081	Minnesota, Wisconsin, Michigan, Illinois, Indiana, Ohio
VI	Air, Pesticides and Toxics Division 1445 Ross Ave. Dallas, TX 75202 (214) 655-7220	Arkansas, Louisiana, Oklahoma, Texas, New Mexico

Table A.1 EPA Regional Offices (continued)

Region	Address	States included
VII	Air and Toxics Division 726 Minnesota Ave. Kansas City, KS 66101 (913) 551-7020	Iowa, Kansas, Missouri, Nebraska
VIII	Air and Toxics Division 999 18th Street Suite 500 One Denver Place Denver, CO 80202-2405 (303) 293-1750	Montana, Colorado, Utah, Wyoming, North Dakota, South Dakota
IX	Air and Toxics Division 1235 Mission Street San Francisco, CA 94103 (415) 556-5568	California, Nevada, Guam, Hawaii
X	Air and Toxics Division 1200 6th Ave. Mail Stop 50121 Seattle, WA 98101 (206)442-4166	Alaska, Washington, Oregon, Idaho

APPENDIX B
RUGGEDNESS TESTING (LABORATORY EVALUATION)

B.1 INTRODUCTION

B.1.1 Ruggedness testing is a useful and cost-effective laboratory study to determine the sensitivity of a method to certain parameters such as sample collection rate, interferant concentration, collecting medium temperature, or sample recovery temperature. This appendix discusses generally the principle of the ruggedness test.

B.1.2 In a ruggedness test, several variables are changed simultaneously rather than one variable at a time. This reduces the number of experiments required to evaluate the effect of a variable. For example, the effect of seven variables can be determined in eight experiments rather than 128 (W.J. Youden, **Statistical Manual of the Association of Official Analytical Chemists**, Association of Official Analytical Chemists, Washington, DC, 1975, pp. 33-36).

B.1.3 Data from ruggedness tests are helpful in extending the applicability of a test method to different source concentrations or source categories.

B.2 RUGGEDNESS TEST DESIGN

B.2.1 If an evaluation of seven factors of a method is desired, then eight experiments can be conducted in the combinations shown in Table B.1. The uppercase letters A, B, C, D, E, F, and G represent the nominal values for the seven different factors, and the lowercase letters a, b, c, d, e, f, and g represent alternative values for the same factors. The results from these combinations are denoted by the letters s, t, u, v, w, x, y, and z.

B.2.2 To evaluate the effect of A-a, the average of the results from combinations 1, 2, 3, and 4 (level A) are compared to combinations 5, 6, 7, and 8 (level a). That is, the average $(s + t + u + v)/4$ is compared to $(w + x + y + z)/4$. As can be seen from Table B.1, each of the two groups contains the other seven factors, which occur twice at the uppercase level and twice at the lowercase level. The effects of these other factors cancel out, which leaves only the effect of changing A to a. In the same way, the effects of the other factors can be evaluated.

B.3 EXAMPLE RUGGEDNESS TEST DESIGN

B.3.1 The example given below was taken from Appendix B of the **Statistical Manual of the Association of Official Analytical Chemists** (Youden, 1975).

B.3.2. A ruggedness test was performed on a distillation method for determining total water in phosphoric acid. Table B.2 lists the seven variables and the levels assigned to the uppercase letters and the lowercase letters.

Table B.1. Ruggedness Test Matrix

	Combination Number							
Factor Value	1	2	3	4	5	6	7	8
A or a	A	A	A	A	a	a	a	a
B or b	B	B	b	b	B	B	b	b
C or c	C	c	C	c	C	c	C	c
D or d	D	D	d	d	d	d	D	D
E or e	E	e	E	e	e	E	e	E
F or f	F	f	f	F	F	f	f	F
G or g	G	g	g	G	g	G	G	g
Result	s	t	u	v	w	x	y	z

Table B.2. Test Conditions

Condition	No.	Letter	Value for capital letter	Value for lower-case letter
Amount of H_2O	1	A, a	ca. 2 Ml	ca. 5 Ml
Reaction time .	2	B, b	0 min	15 min
Distillation rate	3	C, c	2 drops/s	6 drops/s
Distillation time	4	D, d	90 min	45 min
n-Heptane used	5	E, e	210 Ml	190 Ml
Aniline used .	6	F, f	8 Ml	12 Ml
Reagent	7	G, g	New	Used

B.3.3 The analytical results are shown in Table B.3.

Table B.3. Analytical Results

Determination number	Water %
1	18.80
2	20.58
3	19.90
4	18.03
5	19.50
6	19.16
7	19.88
8	19.85

B.3.4 To determine the effect of using new reagent as opposed to used reagent, compare the average water percentage determined from the four determinations with new reagent (G) with the average determined from the four determinations with used reagent (g). Table B.1 shows that G (new reagent) was used in determinations 1, 4, 6, and 7, which gave the results 18.80, 18.03, 19.16, and 19.88. The average of these four results is 18.97. The average of the other four results with used heptane (tests 2,3,5, and 8) is 19.96. The difference in the averages is -0.99, which amounts to a 5.2-percent difference between determinations made with new reagent and used reagent. In the same manner, the effect of changing the other analytical conditions can be determined by comparing the averages of measurements performed with two different values for conditions.

APPENDIX C

PROCEDURE FOR INCLUDING SAMPLE STABILITY IN BIAS AND PRECISION EVALUATIONS

C.1 INTRODUCTION

C.1.1 The test method being evaluated must include procedures for sample storage and the time within which the collected samples shall be analyzed. For example, Tedlar bag samples may be stored at room temperature and kept in the dark to minimize photochemical degradation. Sorbent samples may be stored in the refrigerator or freezer or on dry ice. Impinger samples may be stored in refrigerators. If sorbent samples are extracted with solvent, then the solvent along with the extracted material may be stored in a refrigerator. In addition, the method may specify that the collected samples or extracted materials must be analyzed within 24 hours from the time of collection or extraction.

C.1.2 This appendix discusses the procedures for including the effect of storage time in bias and precision evaluations. The evaluation may be deleted if the test method specifies a time for sample storage.

C.2 STABILITY TEST DESIGN

The following procedures should be conducted to identify the effect of storage times on analyte samples. Store the samples according to the procedure specified in the test method.

C.2.1 For sample container (bag or canister) and impinger sampling systems set up in regards to Section 5.1 and 5.3, analyze six of the samples at the minimum storage time. Then analyze the same six samples at the maximum storage time.

C.2.2 For sorbent sampling systems set up in regards to Section 5.1 and 5.3 that require liquid extraction, extract six of the samples at the minimum storage time and extract six other samples at the maximum storage time. Analyze an aliquot of the first six extracts at both the minimum and maximum storage times. This will provide some freedom to analyze extract storage impacts.

C.2.3 For sorbent sampling systems set up in reference to Section 5.1 and 5.3 that require thermal desorption, analyze six samples at the minimum storage time. Analyze another set of six samples at the maximum storage time.

C.2.4 For systems set up in accordance with Section 5.2, the number of samples analyzed at the minimum and maximum storage times shall be half those collected (8 or 9).

APPENDIX D

PROCEDURE FOR DETERMINATION OF PRACTICAL LIMIT OF QUANTITATION

D.1 INTRODUCTION

D.1.1 The practical limit of quantitation (PLQ) is the lowest level above which quantitative results may be obtained with an acceptable degree of confidence. For this protocol, the PLQ is defined as 10 times the standard deviation, s_o, at the blank level. This PLQ corresponds to an uncertainty of ±30 percent at the 99-percent confidence level.

D.1.2 The PLQ will be used to establish the lower limit of the test method.

D.2 PROCEDURE I FOR ESTIMATING s_o

This procedure is acceptable if the estimated PLQ is no more than twice the calculated PLQ. If the PLQ is greater than twice the calculated PLQ use Procedure II.

D.2.1 Estimate the PLQ and prepare a test standard at this level. The test standard could consist of a dilution of the reference material described in Section 3.0.

D.2.2 Using the normal sampling and analytical procedures for the method, sample and analyze this standard at least seven times in the laboratory.

D.2.3 Calculate the standard deviation, s_o, of the measured values.

D.2.4 Calculate the PLQ as 10 times s_o.

D.3 PROCEDURE II FOR ESTIMATING s_o

This procedure is to be used if the estimated PLQ is more than twice the calculated PLQ.

D.3.1 Prepare two additional standards at concentration levels lower than the standard used in Procedure I.

D.3.2 Sample and analyze each of these standards at least seven times.

D.3.3 Calculate the standard deviation for each concentration level.

D.3.4 Plot the standard deviations of the three test standards as a function of the standard concentrations.

D.3.5 Draw a best-fit straight line through the data points and extrapolate to zero concentration. The standard deviation at zero concentration is s_o.

D.3.6 Calculate the PLQ as 10 times s_o.

APPENDIX E
EXAMPLE CALCULATIONS

E.1 INTRODUCTION

The following section provides an illustration of the principles outlined in Sections 5.0 and 6.0.

E.2 Bias and Precision Calculation for Isotopic and Analyte Spiking

E.2.1 Section 5.1 identifies an isotopic spiking approach for bias and precision determinations. Section 6.1 provides the calculation procedures for isotopic spiking. Section 5.3 identifies an analyte spiking approach for bias and precision. Sections 6.3 provides the calculation procedures for analyte spiking. Section 5.2 identifies an approach for bias and precision comparing a proposed method to a validated method. Section 6.2 provides the calculation procedures for the method comparison approach.

E.2.2 The following example deals with an isotopic spiked demonstration.

E.2.2.1 A sampling train was spiked with 100 μg of an isotope of chromium while it was sampling combustion gas. The isotope was recovered from the trains and quantified. Table E.1 depicts the data for the following example.

Table E.1 Cr(VI) Recoveries

Recovered Cr(VI), (S_i), μg	Difference from 100 μg standard	$(S_i-S_m)^2$ μg
110.2	10.2	292.07
85.9	-14.1	51.98
92.4	- 7.6	0.50
93.9	- 6.1	0.62
103.5	3.5	107.95
117.3	17.3	585.16
82.6	-17.4	110.46
102.6	2.6	90.06
79.5	-20.5	185.28
89.7	-10.3	11.63
73.1	-26.9	400.40
86.7	-13.3	41.09

E.2.2.2 Use Equation 6-1 to calculate the numerical value of the bias:

$$B = \frac{(S_1+S_2+S_3+S_4+S_5+S_6.....S_{12})}{12} - CS$$

$$B = 93.11 - 100 = -6.89\ \mu g$$

CS should remain constant; it is the amount of spiked isotope into each of the 12 samples.

E.2.2.3 Use Equation 6-2 to calculate the standard deviation (SD) of the spiked sample values.

$$SD = \sqrt{\frac{(S_1-S_m)^2 +...+ (S_{12}-S_m)^2}{n-1}}$$

$$SD = \sqrt{\frac{1877.15}{11}} = 13.06\ \mu g$$

E.2.2.4 Using the standard deviation value calculate the standard deviation of the mean (SDM) using Equation 6-3.

$$SDM = \frac{SD}{\sqrt{n}}$$

$$SDM = \frac{13.06}{\sqrt{12}} = 3.77\ \mu g$$

E.2.2.5 Test the bias for statistical significance using Eqution 6-4.

$$t = \frac{|B|}{SDM}$$

$$t = \frac{6.89}{3.66} = 1.88$$

The comparison with the critical value of the two-sided t-distribution at the 95-percent confidence level and n-1 degrees of freedom, indicates that 1.88 is less than the critical value of 2.201. The bias is not statistically significant.

E.2.2.6 To assess the acceptability of the precision, calculate the relative standard deviation using Equation 6-6.

$$RSD = \frac{SD}{S_m} \times 100$$

$$RSD = \frac{13.06}{93.11} \times 100 = 14.03\ \%$$

The precision is 14.03 percent, which is less than the 50 percent criteria specified in Section 2.2 of the protocol.

E.2.3 Bias and Precision Calculations for an Analyte Spiked Demonstration

E.2.3.1 Half of the sampling trains required by Section 5.3 were spiked with 100 μg of the analyte of concern while it was sampling combustion gas. The analyte was recovered from the trains and quantified. Table E.2 depicts the data for the following example.

Table E.2 Analyte Recoveries

Sample Run	Approximate Stack Value µg	Spike Value µg	Measured Value µg	Difference Measured Value µg, d_i	(d_i^2)
1	22	100	119.7	6.8	46.24
	22	100	112.9		
	22		24.9	-5.6	31.36
	22		30.5		
2	30	100	137.1	0.7	0.49
	30	100	136.4		
	30		32.0	10.7	114.50
	30		21.3		
3	27	100	118.0	-5.0	25.00
	27	100	123.0		
	27		35.0	-5.0	25.00
	27		32.0		
4	12	100	109.3	5.3	28.09
	12	100	104.0		
	12		5.4	-12.6	158.76
	12		18.0		
5	28	100	119.8	-4.8	23.04
	28	100	124.6		
	28		36.0	2.3	5.29
	28		33.7		
6	6	100	109.8	0.6	0.36
	6	100	109.2		
	6		11.6	-3.1	9.61
	6		14.7		

E.2.3.2 Use Equation 6-12 to calculate the numerical value of the bias:

The mean of the spiked values (S_m) = 118.65 μg
The mean of the unspiked samples (M_m) = 24.59 μg
The calculated value of the spike (CS) = 100 μg

$$B = S_m - M_m - CS$$

$$B = 118.65 - 24.59 - 100 = -5.94\ \mu g$$

CS should remain constant; it is the amount of spiked analyte into each of the 12 sampling trains.

E.2.3.3 Use Equation 6-13 to calculate the standard deviation (SD) of the spiked sample values.

$$SD_S = \sqrt{\frac{\sum (d_i)^2}{24}} = 3.204$$

E.2.3.4 Using the standard deviation calculated above calculate the standard deviation of the mean (SDM) for the spiked samples using Equation 6-3.

$$SDM = \frac{3.204}{\sqrt{12}} = 0.926$$

E.2.3.5 Test the bias for statistical significance using Equation 6-4.

$$t = \frac{|B|}{SDM}$$

$$t = \frac{5.94}{0.926} = 6.41$$

The comparison with the critical value of the two-sided t-distribution at the 95-percent confidence level and n-1 degrees of freedom, indicates that 6.41 is greater than the critical value of 2.201. The bias is statistically significant.

E.2.3.6 If the bias is statistically significant, a correction factor is calculated using Equation 6-5.

$$CF = \frac{1}{(1 + \frac{B}{CS})}$$

$$CF = \frac{1}{1 + \frac{-5.94}{100}} = 1.06$$

The correction factor is between Section 2.1 criteria of .70 and 1.3, and is therefore acceptable. All of the analyte recovery results should be multiplied by this factor for the final recovery results.

E.2.3.7 To assess the acceptability of the precision calculate the relative standard deviation using Equation 6-6.

$$RSD = \frac{SD}{S_m} \times 100$$

$$RSD = \frac{3.204}{118.65} \times 100 = 2.70\ \%$$

The precision is 2.70 percent, which is less than the 50 percent criteria specified in Section 2.2 of the protocol.

E.2.3.8 From the differences between the pairs of the unspiked method measurements, determine the standard deviation using Equation 6-13.

$$SD_U = \sqrt{\sum \frac{(d_{iu})^2}{24}} = 0.66$$

Complete the evaluation by calculating the relative standard deviation of the mean of the unspiked samples using Equation 6-6.

$$RSD = \frac{0.66}{24.95} \times 100 = 2.64\ \%$$

The precision is 2.64 percent, which is less than the 50 percent criteria specified in Section 2.2 of the protocol.

E.2.4 Bias and Precision Calculations for Method Comparisons

E.2.4.1 A source is proposing to reduce the cost of testing through the use of a new method developed by the company. Table E.3 provides the results of a paired train sampling validation demonstration.

Table E.3 Paired Sample Recoveries

Test	Validated Value μg	Proposed Value μg	Difference d_i μg	d_i^2
1	14.7	14.0	-0.7	0.49
2	14.5	15.0	0.5	0.25
3	14.7	14.6	-0.1	0.01
4	14.6	14.9	0.3	0.09
5	14.5	15.0	0.5	0.25
6	14.8	15.4	0.6	0.36
7	14.3	14.9	0.6	0.36
8	15.0	14.4	-0.6	0.36
9	14.4	14.5	0.1	0.01
V_m = 14.6 μg		P_m = 14.7 μg	d_m = 0.13	

E.2.4.1.1 Determine the acceptance of the proposed method's variance with respect to the variability of the validated method results. The procedure provides a method for testing whether the scatter of two sets of data is such as would be expected from two samples from the same population. The variance provided with the validated method is 0.046.

E.2.4.1.2 Calculate the variance of the proposed method, S_p^2, and the validated method, S_v^2, using Equation 6-7.

$$S_v^2 = 0.046$$

$$S_p^2 = 0.1717$$

E.2.4.1.3 Test to compare if the experimental variance of the proposed method is significantly different from that of the validated method by calculating the F-value using Equation 6-8.

$$F = \frac{0.1717}{0.046} = 3.73$$

The calculated value of 3.73 is greater than the critical value of 1.0. The proposed method is considered to be less precise than the validated method. The data is not as precise as would be obtained by the validated method and therefore, the method and data should be rejected. A bias analysis should be made even though the precision is not acceptable.

E.2.4.1.4 Test the bias for statistical significance. This illustration is provided to evaluate, in this example, whether the bias is significant and the resulting correction factor would also cause rejection of the data and the method. The results do not nullify the results of the F-test. The d_m equals 0.13 μg. The standard deviation of the d_i's is, using Equation 6-2:

$$S_d = \sqrt{\frac{2.017}{8}} = 0.502$$

The SD of the proposed method using Equation 6-9a is:

$$SD_p = \sqrt{SD_d - SD_v}$$

$$SD_p = 0.1717$$

Calculate the t-statistic using Equation 6-9. The calculated t-value is:

$$t = \frac{0.13}{\left[\frac{0.1717}{\sqrt{9}}\right]} = 2.28$$

E.2.4.1.5 Compare the calculated t-statistic to the critical t-value at the 80-percent confidence level and n-1 degrees of freedom. The critical value is 1.397 for eight degrees of freedom. The calculated value, 2.28, is greater than the criticle value. A correction factor determination is necessary. **The data and method rejection in this example is principly due to variability. The cause of the unacceptable variability should be ascertained by the analyst.**

E.2.4.2 Table E.4, Quad Sample Recoveries, provides the results of a quadruplet test conducted in a proposed method demonstration test following the procedures in Section 5.2.2.

Table E.4 Quad Sample Recoveries

Test	Method	Value ppm	$(V_1+V_2)/2$	$(P_1+P_2)/2$	d_i	$(d_i-d_m)^2$
A	V	365	368		7.5	83.26
	V	372				
	P	366		360.5		
	P	355				
B	V	381	379		4.0	31.64
	V	377				
	P	370		375		
	P	380				
C	V	349	364.5		-39.5	1434.51
	V	380				
	P	330		325		
	P	320				
D	V	362	363.5		21.5	534.77
	V	365				
	P	338		342		
	P	346				
					d_m= - 1.625	

E.2.4.2.1 Determine the acceptance of the proposed method's variance with respect to the variability of the validated method results. If a significant difference is determined the proposed method and the results are rejected. Calculate the variance of the validated method, S_v^2, and the proposed method, S_p^2, using Equation 6-7 and the results in Table E.5.

Table E.5 Method Mean Data

Test	e_i	$(e_i-e_m)_v^2$	e_i	$(e_i-e_m)_p^2$
A	365	49	366	225
	372	16	355	16
B	381	169	370	841
	377	81	380	361
C	349	361	330	441
	380	144	320	961
D	362	36	338	169
	365	9	346	25
	$S_V^2 = 108.75$		$S_P^2 = 378.87$	

E.2.4.2.2 Determine if the variance of the proposed method is significantly different from that of the validated method by calculating the F-value using Equation 6-8. Compare the calculated F-value with the critical value. The critical value is 1.0 when the procedure specified in 5.2.2 for quadruplet trains is followed.

$$F = \frac{379.87}{108.75} = 3.49$$

The calculated F is not less than the critical value, this indicates that the precision of the proposed method is not as good as the validated method.

E.2.4.2.3 Test the bias for statistical significance by calculating the t-statistic. Determine the mean, d_m, and the standard deviation, S_d, of the d_i's. Calculate the standard deviation of the mean of the differences, SD_d, using Equation 6-7.

$$SD_d = \sqrt{\frac{2084.18}{3}} = 26.35$$

Calculate the t-statistic using Equation 6-9.

$$t = \frac{1.625}{\left(\frac{26.35}{\sqrt{4}}\right)} = 0.123$$

For the procedure in Section 5.2.2, n equals four.

Compare the calculated t-statistic with the critical t-value from the table of the t-statistic and determine if the mean is significant at the 80-percent confidence level. When four runs are conducted, as specified in Section 5.2.2, the critical value of the t-statistic is 1.638 for three degrees of freedom. The calculated t-value is less than the critical value and therefore the results should not be corrected for bias.

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
450/4-90-015		

4. TITLE AND SUBTITLE	5. REPORT DATE
Protocol For The Field Validation of Emission Concentrations From Stationary Sources	April 1991
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT NO.
	11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS U.S. Environmental Protection Agency Research Triangle Park, NC 27711	13. TYPE OF REPORT AND PERIOD COVERED
	14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

The purpose of the validation protocol is to determine bias and precision of the test method at the level of concentration in the gas stream. Procedures involve (a) introducing known concentrations of an analyte or comparing the method against a validated test method to determine bias and (b) using multiple sampling trains to determine precision.

The protocol lists a number of important requirements for the validation of the test method. They include; use of EPA audit material; documenting and reporting results; procedures for determining bias and precision by means of isotopic and analyte spiking of multiple train samples or comparison to validated methods; and, procedures for calculating precision, bias and correction factors. The protocol also defines the acceptance criteria in terms of percent bias and precision and how the determined precision dictates future testing with the validated method.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Emission Testing Early Reduction Title III	Emission Testing	

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
Release Unlimited	Unclassified	
	20. SECURITY CLASS *(This page)*	22. PRICE
	Unclassified	

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE