EPA-650/4-75-009

EVALUATION AND COLLABORATIVE STUDY OF METHOD FOR VISUAL DETERMINATION OF OPACITY OF EMISSIONS FROM STATIONARY SOURCES

by

Henry F. Hamil, Richard E. Thomas,
and Nollie F. Swynnerton

Southwest Research Institute,
8500 Culebra Road
San Antonio, Texas 78284

Contract No. 68-02-0626
ROAP No. 26AAG
Program Element No. 1HA327

EPA Project Officer: M. Rodney Midgett

Quality Assurance and Environmental Monitoring Laboratory
National Environmental Research Center
Research Triangle Park, North Carolina 27711

Prepared for

U.S. ENVIRONMENTAL PROTECTION AGENCY
OFFICE OF RESEARCH AND DEVELOPMENT
WASHINGTON, D.C. 20460

January 1975

EPA REVIEW NOTICE

This report has been reviewed by the National Environmental Research Center - Research Triangle Park, Office of Research and Development, EPA, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

RESEARCH REPORTING SERIES

Research reports of the Office of Research and Development, U. S. Environmental Protection Agency, have been grouped into series. These broad categories were established to facilitate further development and application of environmental technology. Elimination of traditional grouping was consciously planned to foster technology transfer and maximum interface in related fields. These series are:

1. ENVIRONMENTAL HEALTH EFFECTS RESEARCH
2. ENVIRONMENTAL PROTECTION TECHNOLOGY
3. ECOLOGICAL RESEARCH
4. ENVIRONMENTAL MONITORING
5. SOCIOECONOMIC ENVIRONMENTAL STUDIES
6. SCIENTIFIC AND TECHNICAL ASSESSMENT REPORTS
9. MISCELLANEOUS

This report has been assigned to the ENVIRONMENTAL MONITORING series. This series describes research conducted to develop new or improved methods and instrumentation for the identification and quantification of environmental pollutants at the lowest conceivably significant concentrations. It also includes studies to determine the ambient concentrations of pollutants in the environment and/or the variance of pollutants as a function of time or meteorological factors.

This document is available to the public for sale through the National Technical Information Service, Springfield, Virginia 22161.

SUMMARY AND CONCLUSIONS

This report presents the results of statistical analyses on data obtained in the collaborative testing of EPA Method 9 (Visual Determination of Opacity). There were 5 separate tests of Method 9 at three different sites: a training smoke generator, a sulfuric acid plant and a fossil fuel-fired steam generator.

For these tests, the opacity measure used is the average opacity, defined as the average of 25 readings made at 15-second intervals. These determinations are analyzed for their accuracy and their precision. For this report, accuracy is measured by the deviation of the observer's determination from the true opacity as measured by the in-stack transmissometer. The precision of the method is measured in terms of within-observer, observer bias and between-observer terms.

- **Training Generator**–Thirty-six runs were made over a two-day period using two training smoke generators with 9 qualified observers participating. There were 20 runs using white smoke and 16 using black smoke with plume opacity ranging from just above zero to 40 percent. The two cases are treated separately.

  - *White Smoke*–There were a total of 170 determinations made using a white plume. These values showed a low bias with respect to the transmissometer, and the deviation from the meter is estimated by

$$\text{deviation} = 3.46 - 0.33 \text{ (meter average).}$$

The within-observer standard deviation is estimated as 2.38 percent opacity with 133 degrees of freedom. The observer bias standard deviation is 0.95 percent opacity with 7 df. This gives a between-observer standard deviation of 2.56 percent opacity.

  - *Black Smoke*–There were 133 determinations of the average opacity of black smoke. The results from these values were similar to those of the white smoke. The predicted deviation from the transmissometer for this case is

$$\text{deviation} = 3.74 - 0.34 \text{ (meter average).}$$

The within-observer standard deviation is 1.84 percent opacity with 105 degrees of freedom. The observer bias standard deviation is 1.00 percent opacity with 7 df. This gives a between-observer standard deviation of 2.09 percent opacity.

- **Sulfuric Acid Plant**–Thirty observation runs were made over two days at a sulfuric acid plant. There were 11 observers on the first day for 14 runs and 9 on the second day for the remaining 16 runs. The plume opacity was varied around the compliance limit for sulfuric acid plants, ranging from 2 to 15 percent average opacity. The observers were chosen from both a local enforcement agency and from private concerns to allow a comparison of the two groups.

The predicted deviation for these data is -2.0 percent opacity for the entire range studied. The within-observer standard deviation is 2.12 percent opacity with 232 degrees of freedom. The observer bias standard deviation is 0.96 percent opacity with 8 degrees of freedom. This gives a between-observer standard deviation of 2.33 percent opacity. It is shown for this test that the enforcement observers read a higher, and more accurate, average opacity than did the private sector observers.

- **Steam Station**–Three separate tests were conducted at a power plant with plume opacity ranging up to 35 percent average opacity. On the first two days of testing, the weather conditions prevented the accurate determination of plume opacity. In the third test period, viewing conditions were ideal and the accuracy of the method is evaluated.

• *Test 1*–There were 10 runs using 10 observers at the first test. The determinations were well below the concurrent meter averages due to the sky and wind conditions. The deviation from the meter average is represented by

$$\text{deviation} = -3.01 - 0.40\ (\text{meter average}).$$

A total of 60 determinations were used to obtain the precision estimates. The within-observer standard deviation is estimated as 1.82 percent opacity with 45 df. The observer bias standard deviation is estimated as 1.39 percent opacity with 8 df, which gives a between-observer standard deviation estimate of 2.29 percent opacity.

• *Test 2*–Eighteen runs were made at the second test using 10 observers. The determinations made were more accurate than the previous test, but still were well below the meter average. The deviation from the true opacity is represented by

$$\text{deviation} = -1.02 - 0.19\ (\text{meter average}).$$

A total of 118 determinations are used to obtain the precision estimates. The within-observer standard deviation is 1.84 percent opacity with 88 degrees of freedom. The observer bias standard deviaton is estimated as 1.50 percent opacity with 8 degrees of freedom. The estimated between-observer standard deviation, then, is 2.37 percent opacity.

• *Test 3*–Twenty-four runs were made using eight observers over a two-day period. The resulting 192 determinations have a negative bias with respect to the metered opacity that can be represented by

$$\text{deviation} = 2.27 - 0.24\ (\text{meter average}).$$

The within-observer standard deviation is estimated as 1.89 percent opacity with 90 degrees of freedom. The observer bias standard deviation is 1.77 percent opacity with 6 degrees of freedom, which gives a between-observer standard deviation of 2.59 percent opacity.

Also included for this test are evaluations of two variations to the method: reading to the nearest percent, and averaging two observers' results.

• **Composite Estimation**–Using the results from the three test sites, it is possible to make composite estimates of the accuracy and precision that can be expected with field use of Method 9. The expected deviation from the average metered opacity is given by

$$\text{deviation} = 3.13 - 0.31\ (\text{meter average})$$

for the range from 5 to 35 percent average opacity. This equation is obtained from the training generator and steam station test data.

The composite precision estimates are obtained using the estimates from all tests. The estimated within-observer standard deviation is 2.05 percent opacity with 693 df. The estimated observer bias standard deviation is 1.29 percent opacity with 44 df. This gives a composite between-observer standard deviation estimate of 2.42 percent opacity.

For each case, estimates are made of the range of determinations that would be expected from a single observer, and of the maximum difference that would be expected between two observers when determining the average opacity of a plume.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

LIST OF TABLES

LIST OF TABLES (Cont'd)

I. INTRODUCTION

This report describes the work performed and results obtained on Southwest Research Institute Project 01-3462-006, EPA Contract No. 68-02-0626, which includes collaborative testing of Method 9 for visual determination of the opacity of emissions from stationary sources as given in "Standards of Performance for New Stationary Sources"(1)* and "Stationary Sources, Proposed Emission Monitoring and Performance Testing Requirements."(2)

This report describes the statistical analysis of the data from collaborative tests conducted on two EPA-type training smoke generators, and at a sulfuric acid plant and a coal-fired power plant.

The results of the data analyses and the conclusions based on these analyses are given in this report.

*Superscript numbers in parentheses refer to the List of References at the end of this report.

II. COLLABORATIVE TESTING

A. Purpose

Collaborative testing of the method for visual determination of opacity of emissions from stationary sources[1,2] was conducted in order to obtain data using certified observers which would allow statistical evaluation of the method. Results from these tests are presented in this report.

Three collaborative test sites were used: a training smoke generator, a sulfuric acid plant, and a fossil fuel-fired steam generator. Tests at the first two sites mentioned above were designed to evaluate the determination of average opacity of emissions[1], while the test at the third site was designed to allow evaluation of both the determination of minutes of noncompliance with a standard and determination of average opacity.[2] The results in this report are all based upon average opacity determinations.

The initial test on the training smoke generator was conducted to provide background information on the use of the method. The tests at the sulfuric acid plant and the fossil fuel-fired steam generator were conducted to obtain information on the use of the method on applicable sources under field conditions. At no time during any of the test were warm-up or practice runs allowed prior to the test itself.

B. Smoke Generator Test

The smoke generator test was conducted on 19-20 November 1973 at the National Environmental Research Center, Research Triangle Park, North Carolina. Two training smoke generators were used, one belonging to the EPA and one belonging to the State of North Carolina. A total of nine collaborators from enforcement agencies in the State of North Carolina were available for the tests. All nine collaborators participated in the first fifteen runs--only eight participated in the remaining twenty-one runs, since one collaborator operated the State of North Carolina's smoke generator. Time since last certification for the collaborators varied from one to thirteen weeks.

During the two-day test, a total of thirty-six runs were made, with a run consisting of twenty-five readings taken at the fifteen-second intervals by each collaborator. The collaborators read the plume on a time signal from the test supervisor in order to assure that the readings were as nearly simultaneous as possible. The opacity of the plume as indicated by the in-stack transmissometer of the smoke generator was recorded at the time signal to provide "true" opacity data. Twenty runs were made with white plumes, and sixteen runs were made with black plumes. The range of opacity studied was from just above 0 to 40 percent. During some of the runs, plume opacity was held essentially constant while on other runs the opacity of the plume was varied randomly to determine how well the observers could follow changing opacity during the course of a set of observations.

Due to the relatively short stack height on the smoke generators, in some runs the observers had objects in the background (buildings, trees, power lines, etc.), while on other runs only a sky background was available. Since use of background objects in evaluating plume opacity is permitted in most training courses, no differentiation was made in this test between runs with or without background objects. During both days of the test, viewing conditions were ideal, with blue sky and bright sunshine. On all runs, the observers were allowed to select the optimum viewing position with respect to sun and wind angles.

C. Stauffer Chemical Company Test

The sulfuric acid plant test was conducted on 29-30 January 1974 at Stauffer Chemical Company's facilities in Houston, Texas. The unit was equipped with a 300-foot-high stack. In-stack monitoring of opacity was accomplished by use of a Lear Siegler Model 610A transmissometer*.

A total of eleven collaborators were available for the first day of the test. Nine collaborators were available for the second day of the test. Four of the collaborators were from the City of Houston Department of Public Health,

*Mention of trade names or specific products does not constitute endorsement by the Environmental Protection Agency.

six were from area industries, and one was from a local health association. All collaborators' most recent certification was on 24 August 1973 at the State of Texas Training Session conducted at Houston, Texas.

During the two-day test, a total of thirty runs were made, fourteen runs on the first day and sixteen runs on the second day. Each run consisted of twenty-five readings taken at 15-second time intervals by each observer. The collaborators began each run on a time signal from the test supervisor. The average opacity as indicated by the in-stack transmissometer was obtained from the instrument strip chart recorder for the time period corresponding to each run.

Arrangements were made with plant personnel to change unit operating conditions during the test to provide plumes of varying opacity. The range of opacity studied was from slightly above 0 to 15 percent. Opacity of the plume was essentially constant during any single run. During both days of the test, good viewing conditions prevailed, with blue sky and a slight haze. Wind was light and variable throughout the test. Color of the plume was white.

The observers were allowed to select the optimum viewing position with respect to wind direction and sun angle.

D. Riverbend Steam Station

Three tests were conducted at the Duke Power Company's Riverbend Steam Station near Charlotte, N.C. Tests were conducted on 27 August, 5 September, and 30 September-1 October, 1974. The tests involved use of two units at the power plant, number 9 and number 10 boilers. Each boiler was equipped with two stacks (A&B). Readings for the first two tests and the first day of the third test were made on Stack B of boiler number 9. On 1 October 1974, the number 9 boiler was shut down, and the third test was completed using Stack A of number 10 boiler. Both stacks were fitted with Lear Siegler Model RM4 transmissometers*.

The plan for the tests at the Riverbend Steam Station varied from the plans for the previously described tests This set of tests was designed to provide information on some variations to the method as written.[1,2] The collaborators were divided into two equal groups with regard to size and experience. One group was used as a control and made their readings of opacity within the constraints of the published method, with the sun in the quadrant to their rear, assigning opacity to the plumes in 5-percent increments. The other was used as the experimental group. One variation under study by this group was reading at 45- and 90-deg angles with respect to the sun. Opacity was assigned in 5-percent increments during this portion of the experiment. The other variation was reading the plume in 1-percent increments from the same position as the control group.

From this test, the effect on a method test result induced by reading outside some of the constraints of the method may be estimated. By altering the angle of view with respect to the sun, it may be determined whether this is an important consideration for the observer. By requiring that the plume be read in 1-percent increments, it may be determined if this variation is more accurate or more precise. In addition to these, the relative performance of an inexperienced to an experienced observer may be assessed, and the possible improvement of the method by averaging two observers' results may be evaluated.

In the first two tests, a run consisted of forty observations by each observer taken at 15-second intervals. This was done for the purpose of evaluating the method when used to determine minutes of noncompliance with a standard. The average opacity presented in the data tables represents the average of the first twenty-five readings by an observer in each run. In the third test, a run consisted of twenty-five observations by each observer taken at 15-second intervals. In all three tests, the observers read the plume on a time signal from test personnel. Transmissometer readings were recorded to the nearest percent on the time signal to provide "true" opacity readings.

*Trade name.

TABLE 1. CERTIFICATION DATA ON COLLABORATORS, RIVERBEND STEAM STATION TESTS

Collaborator Code Number	Date of Most Recent Certification	Average Deviation from True Opacity During Certification, %
1	7-16-74	4.7
2	7-16-74	5.7
3*	4-5-74	4.6
4	4-5-74	5.4
5	4-5-74	5.9
6	4-5-74	6.2
7	4-5-74	5.9
8	4-5-74	4.4
9†	3-19-74	4.3
10	8-27-74	3.8
11	9-11-74	3.1
12	9-11-74	2.9
13	9-11-74	3.9

*Not qualified. One reading during qualification deviated from the opacity by more than 15 percent. Not determined until after test due to clerical error.
†Certification expired 9-19, recertified 10-22.

TABLE 2. TEST LOG, RIVERBEND STEAM STATION, TEST 1
27 August 1974

Run No.	Time Run Began	Position Angle, deg	Sky Conditions	Wind
1	1015	45	Mostly cloudy, low haze	NE, 0–5 mph
2	1055	90	Mostly cloudy, low haze	Light, variable
3	1130	90	Mostly cloudy, low haze	Light, variable
4	1330	0	Mostly cloudy, low haze	SW, 5–10 mph
5	1350	0	Mostly cloudy, low haze	SW, 5–10 mph
6	1410	0	Mostly cloudy, low haze	Light, variable
7	1430	45	Mostly cloudy, low haze	Light, variable
8	1450	45	Mostly cloudy, low haze	Light, variable
9	1510	90	Mostly cloudy, low haze	W, 3–5 mph
10	1530	90	Mostly cloudy, low haze	W, 3–5 mph

TABLE 3. TEST LOG, RIVERBEND STEAM STATION, TEST 2
5 September 1974

Run No.	Time Run Began	Position Angle, deg	Sky Conditions	Wind
1	0910	0	Overcast	NE, 5–10 mph
2	0930	0	Overcast	NE, 5–10 mph
3	0950	45	Overcast	NE, 5–10 mph
4	1010	45	Overcast	NE, 5–10 mph
5	1030	90	Overcast	NE, 5–10 mph
6	1050	90	Overcast	NE, 5–10 mph
7	1130	45	Overcast	NE, 5–10 mph
8	1150	45	Overcast	NE, 5–10 mph
9	1210	90	Overcast	NE, 5–10 mph
10	1230	90	Overcast	NE, 5–10 mph
11	1350	0	Overcast	NE, 5–10 mph
12	1410	0	Overcast	NE, 5–10 mph
13	1430	90	Overcast	NE, 5–10 mph
14	1450	90	Overcast	NE, 5–10 mph
15	1530	0	Overcast	NE, 5–10 mph
16	1550	0	Overcast	NE, 5–10 mph
17	1610	45	Overcast	NE, 5–10 mph
18	1630	45	Overcast	NE, 5–10 mph

The range of opacity studied in all three tests was from just above 0 to 40 percent. Opacity of the plume was varied by cutting off one or more stages of the electrostatic precipitator prior to the run. With the exception of occasional short-duration transient high values, plume opacity was essentially constant during a given run for the low opacities. Variation in the "true" opacity within a run increased as the average plume opacity increased.

The collaborators for each test came from State of North Carolina enforcement agencies, EPA offices, and private contractors. The first two tests utilized four state, four EPA, and two contractor personnel. The third test utilized four state, two EPA, and two contractor personnel. The date of the most recent certification of the collaborators is given in Table 1. Also included for each of the collaborators is the average deviation from the true plume opacity for the twenty-five white and twenty-five black plumes read during certification.

The three tests consisted of ten, eighteen, and twenty-four runs, respectively. Viewing conditions during the first test were poor, mostly cloudy and low haze. There was also interference from another plume during portions of the test. Conditions for the second test were marginal, with a solid overcast, while for the third test conditions were ideal, with a cloudless sky and bright sunshine. The plume color was white.

Test logs for the three tests are presented in Tables 2, 3, and 4. In Tables 2 and 3, the position of the experimental group is given relative to the control group. In Table 4, the positions of both the control and experimental groups are given relative to the sun, with the angles calculated as shown in Figure 1.

E. Collaborators and Test Personnel

The collaborators for the smoke generator test were Mr. Sammy L. Amerson, Mr. George Lawson, Mr. Dewey D. Johnson, Mr. Roy T. Gorman, Mr. J.R. Kirk, Mr. D.Y. Daniel, Jr., Mr. B.J. Foust and Mr. Bill Proctor from the State of North Carolina

TABLE 4. TEST LOG, RIVERBEND STEAM STATION, TEST 3
30 September – 1 October 1974

Run No.	Time Run Began	Viewing Angle, θ (Fig. 1) Control Group, deg	Experimental Group, deg	Sky Conditions	Wind
1	1315	0	0	Clear	Light, variable
2	1330	0	0	Clear	Light, variable
3	1345	10	45	Clear	Light, variable
4	1400	10	45	Clear	Light, variable
5	1415	15	60	Clear	Light, variable
6	1430	15	60	Clear	Light, variable
7	1445	45	45	Clear	Light, variable
8	1500	45	45	Clear	Light, variable
9	1515	0	0	Clear	Light, variable
10	1530	0	0	Clear	Light, variable
11	0930	30	30	Clear	N, 8–12 mph
12	0945	30	30	Clear	N, 8–12 mph
13	1000	25	75	Clear	N, 5–10 mph
14	1015	25	75	Clear	N, 5–10 mph
15	1130	20	90	Clear	Light, variable
16	1145	20	90	Clear	Light, variable
17	1200	10	45	Clear	Light, variable
18	1215	10	45	Clear	Light, variable
19	1230	0	0	Clear	Light, variable
20	1245	0	0	Clear	Light, variable
21	1300	0	45	Clear	Light, variable
22	1315	0	45	Clear	Light, variable
23	1330	0	90	Clear	Light, variable
24	1345	0	90	Clear	Light, variable

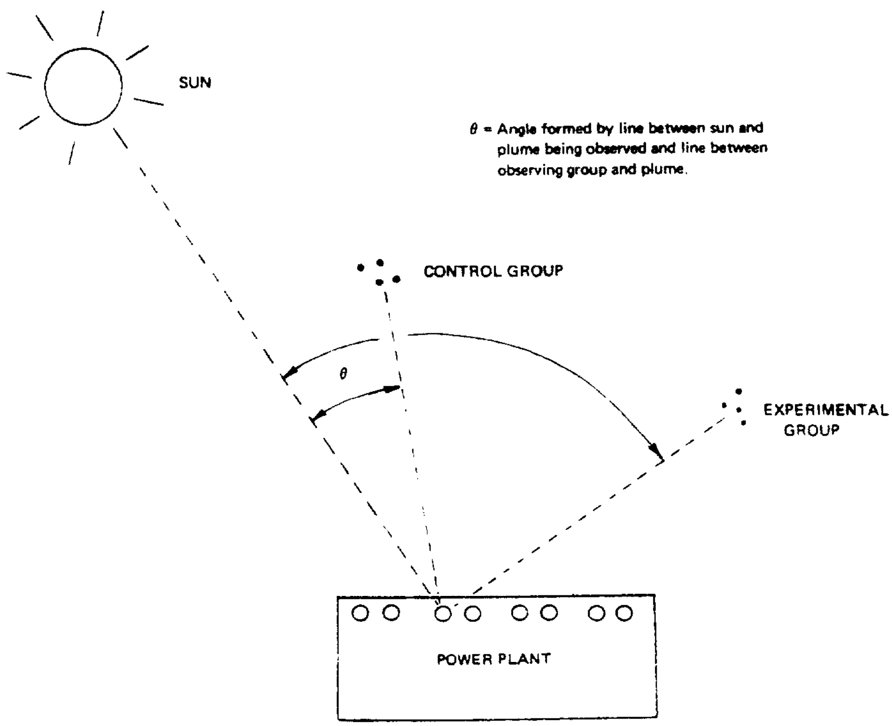

FIGURE 1. SCHEMATIC OF OBSERVER POSITIONS, RIVERBEND STEAM STATION, TEST 3

Department of Natural and Economic Resources; and Mr. William M. Edsel from the County of Forsyth Air Quality Control Division, Forsyth County, North Carolina.

The collaborators for the sulfuric acid plant test, all from the Houston, Texas, area, were Mr. Charles L. Owen from Armco Steel; Mr. James S. Corbin, Mr. Robert J. Stahl, Mr. Ronald F. Stockunas and Mr. Wendell La Foe from the City of Houston Department of Public Health; Mr. B. L. Bolton from Rollins Environmental Services; Mr. T. M. Walker and Mr. Klaus R. Gerlach of Stauffer Chemical Co.; Mr. Mel Remley from Jefferson Chemical Co.; Mr. William Moore from the San Jacinto Lung Association; and Mr. Larry P. Stoltz from GAF Corporation.

The collaborators for the power plant tests were Mr. Dewey D. Johnson, Mr. Ron Jernigan, Mr. John Leatherman and Mr. Don Shepherd from the State of North Carolina Department of Natural and Economic Resources; Mr. William M. Edsel from the County of Forsyth Air Quality Control Division; Mr. David Da Crema and Mr. John Yates, from Engineering Science Inc.; Mr. Alan Luther and Mr. Peter F. Burnette from Environmental Science and Engineering, Inc.; and Mr. Alfred Vervaert, Mr. J.W. Brown, Mr. Joseph Peoples, Mr. James E. Casey, Mr. Jack Siegel and Mr. John Hund from the Environmental Protection Agency.*

The tests were conducted under the general supervision of Mr. Nollie F. Swynnerton. The collaborators from the State of North Carolina were provided through the assistance of Mr. James A. McColman, Chief, Air Quality Division, State of North Carolina Department of Natural and Economic Resources. The collaborators for the sulfuric acid plant test were selected by Mr. Swynnerton from a list of people who had certified at the most recent State of Texas certification school in the Houston area. The list was provided by Mr. Thomas Jay McMickle, Environmental Health Specialist, Texas State Department of Health. The collaborators from EPA, Engineering Science, and Environmental Science and Engineering were provided by Mr. Roy Neulicht, Environmental Engineer, Emissions Standards and Engineering Division, Office of Air Quality Planning and Standards, EPA. Mr. Dennis Holzschuh, Pysical Science Technician, Engineering and Enforcement Section, Air Pollution Training Institute, Control Programs Development Division, Office of Air Quality Standards and Planning, EPA, assisted in the conduct of the smoke generator test at the National Environmental Research Center. Mr. William D. Conner, Research Physicist, Stationary Source Measurements Research Section, Emission Measurements Research Branch, Chemistry and Physics Laboratory, National Environmental Research Center, EPA, made the necessary arrangements with Duke Power Company personnel for the Riverbend Steam Station Tests.

*Throughout the remainder of this report, the observers are referred to by assigned code numbers. The code numbers used do not correspond to the above ordered listing of observers.

III. DEFINITIONS AND STATISTICAL TERMINOLOGY

To facilitate the understanding of this report and the utilization of its findings, this section explains the statistical terms used in this report and defines certain terms used in presenting the data and results.

A. Definitions

1. **Run**–A 6-minute period during which the observers determined the opacity of the plume.

2. **Determination**–The arithmetic mean of 25 consecutive opacity readings taken at 15-second intervals during a run.

3. **Meter Average**–The average opacity of the plume as determined by the in-stack transmissometer.

4. **Deviation**–The difference between a determination and the corresponding meter average; deviation = determination – meter average.

B. Statistical Terminology

1. $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} X_i$ –the sample mean. This statistic estimates the mean, or center, of a given population.

2. $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{x})^2$ –the sample variance. This statistic estimates the dispersion in a distribution around the mean value.

3. $s = \sqrt{s^2}$ –the sample standard deviation. The standard deviation is an alternative estimate of dispersion.

4. **Analysis of Variance (ANOVA)**–A statistical technique for testing whether different groups or factors perform in an equivalent manner, and for obtaining estimates of different variance terms.

5. **SS**–sums of squares. The sum of squared deviations from the mean for a component. The SS are used to estimate variability.

6. **df**–degrees of freedom. The df are an indication of the amount of confidence in the estimate. A larger number of df implies a greater degree of confidence.

7. **MS**–the mean square. The MS is the basis for hypothesis testing in the ANOVA and estimating a variance component. MS = SS/df

8. **EMS**–the expected mean square. The EMS determines which mean squares are used to test which hypotheses, and to estimate the particular variance components.

9. **F-Ratio**–the ratio of two mean squares. The F-ratio is the test statistic for determining the relative performance of two groups.

All precision components in this report are estimated using an ANOVA procedure. The statistical model is given separately for each data set, but, in general, consists of an observer term, a run term and an error term. The MS and EMS from these ANOVA tables are used to determine the significance of the various factors, and to estimate the following precision components.

- **Within-observer.** The within-observer, or error, component measures the dispersion that would be expected in replicate determinations by the same observer at the same true opacity. The within-observer variance is denoted by σ^2.

- **Observer bias.** The observer bias component measures the dispersion that would be expected in replicate determinations due to the use of the method by independent observers. These differences can result from such factors as the training generator on which the observers qualified and the relative frequency with which they use the method. The observer bias variance is denoted by σ_L^2.

- **Between-observer.** The between-observer component is estimated from the within-observer and observer bias terms. The between-observer term measures the variability associated with the difference between results by independent observers. The between-observer term is composed of within-observer and observer bias terms, and the between-observer variance, σ_b^2 is defined as

$$\sigma_b^2 = \sigma_L^2 + \sigma^2 .$$

IV. SMOKE GENERATOR TEST

The initial test of Method 9 was conducted using two EPA-approved smoke generators to produce plumes, both white and black, covering the range of opacities that could be expected by observers seeking to determine compliance with the new source performance standards. These included plumes from just above zero up to 40 percent. On each run, the collaborators recorded their observations to the nearest 5 percent for 25 readings taken at 15-second intervals. The observers read on a signal given by test personnel, so that all determinations were comparable. The 25 readings were averaged by SwRI personnel after the test to insure that the determination was correctly computed.

To determine to what extent the observers were able to follow changes in the plume opacity, the generated opacity was subject to change during a run. During the runs, the opacity was either held constant, increased or decreased according to a test plan. In addition, the number of times the plume opacity was changed, the point at which it was changed and the amount it changed all were varied from run to run so that the observers would not be able to anticipate these changes.

The observers were allowed to choose their position for observing the generated plume, consistent with the constraints of the method. On some runs, they had a background (e.g., trees, buildings) to read against, while on others, there was only blue sky. No record of the background was kept, so any differences in this regard cannot be determined.

There were a total of 36 sampling runs, 20 with white smoke and 16 with black, made during the two day period. One observer participated in only the first 15 runs then operated the second training generator for the remainder of the test. This resulted in a total of 303 determinations, 170 with white smoke and 133 with black. The analyses of the data will concern itself with two aspects: the *accuracy,* as measured by the deviation from the metered opacity, and the *precision,* expressed in terms of standard deviations. The results are presented below separately for the white and black plumes.

A. White Smoke

The 170 determinations are shown in Table 5, along with the concurrent meter averages. To evaluate the performance of the observers, the determinations will be looked at in two ways. First, they will be compared to the meter averages to determine how accurately the observers were able to read the opacity; and second, they will be compared to each other to determine how much variation can be expected both from a single observer, and between two independent observers when reading the same opacity plume.

1. Accuracy

The deviations from the plume opacity are shown in Table 6. As can be seen from the data, there is a tendency for the deviation to be greater and negative as the plume opacity increases. Because of this, a linear regression is used, with opacity as the independent variable and deviation as the dependent variable. The technique is presented in Appendix B1.

For these deviations, the slope of the least-squares fit to these points is significantly different from zero. This indicates that there is a linear relationship between the two variables. The equation of the line derived is

$$\hat{Y}_i = 3.46 - 0.33X_i$$

where

$\hat{Y}_i$ is the predicted deviation

X_i is percent opacity

TABLE 5. DETERMINATIONS OF AVERAGE OPACITY OF TRAINING GENERATOR, WHITE SMOKE

(Percent Opacity)

Run No.	Observer									Meter
	1	2	3	4	5	6	7	8	9	
1	12.8	12.0	10.8	13.2	10.6	17.4	11.0	10.6	13.0	10.0
2	10.6	9.0	11.2	11.0	8.2	10.6	8.4	12.8	14.8	13.6
3	8.8	5.8	8.6	9.6	6.0	8.6	6.6	8.6	10.0	9.2
4	27.0	28.2	23.8	22.6	13.6	22.2	14.6	25.0	24.0	23.8
5	11.2	9.6	12.0	12.0	11.4	12.8	11.8	11.4	11.2	12.3
6	5.6	4.6	5.0	6.6	5.8	5.8	6.0	5.0	6.6	4.7
7	8.8	8.4	7.0	7.2	7.6	8.2	7.2	7.6	7.4	7.6
8	31.2	22.6	24.6	24.4	20.2	21.0	21.4	27.4	27.4	34.0
9	22.6	15.4	13.6	16.6	19.6	17.4	20.6	17.0	16.0	25.4
10	9.4	7.6	8.4	7.8	10.0	7.8	9.0	7.2	8.6	12.4
27	17.4	17.4	19.2	19.6	18.8	17.4	16.4	16.4		13.1
28	8.6	6.4	8.6	7.4	8.0	6.2	5.2	5.2		8.6
29	10.0	8.8	12.4	10.6	11.4	11.4	8.7	8.2		9.0
30	13.0	14:8	15.8	19.4	15.4	12.0	13.6	12.6		11.2
31	18.2	20.8	17.4	18.4	19.6	12.8	13.8	10.8		14.3
32	32.2	29.4	30.8	28.6	27.6	18.2	23.6	23.2		31.3
33	19.0	18.8	20.8	19.6	16.6	15.0	18.8	13.0		13.2
34	9.0	8.8	13.6	9.6	7.2	7.2	7.8	7.0		16.8
35	14.8	17.0	20.4	16.4	17.0	13.0	15.8	11.0		18.7
36	12.4	10.2	9.8	10.0	9.4	7.4	9.2	8.0		14.6

TABLE 6. DEVIATION FROM TRAINING GENERATOR, WHITE SMOKE

(Percent Opacity)

Run No.	Observer									Meter
	1	2	3	4	5	6	7	8	9	
1	2.8	2.0	0.8	3.2	0.6	7.4	1.0	0.6	3.0	10.0
2	−3.0	−4.6	−2.4	−2.6	−5.4	−3.0	−5.2	−0.8	1.2	13.6
3	−0.4	−3.4	−0.6	−0.4	−3.2	−0.6	−2.6	−0.6	0.8	9.2
4	3.2	4.4	0.0	−1.2	−10.2	−1.6	−9.2	1.2	0.2	23.8
5	−1.1	−2.7	−0.3	−0.3	−0.9	0.5	−0.5	−0.9	−1.1	12.3
6	0.9	−0.1	0.3	1.9	1.1	1.1	1.3	0.3	1.9	4.7
7	1.2	0.8	−0.6	−0.4	0.0	0.6	−0.4	0.0	−0.2	7.6
8	−2.8	−11.4	−9.4	−9.6	−13.8	−13.0	−12.6	−6.6	−6.6	34.0
9	−2.8	−10.0	−11.8	−8.8	−5.8	−8.0	−4.8	−8.4	−9.4	25.4
10	−3.0	−4.8	−4.0	−4.6	−2.4	−4.6	−3.4	−5.2	−3.8	12.4
27	4.3	4.3	6.1	6.5	5.7	4.3	3.3	3.3		13.1
28	0.0	−2.2	0.0	−1.2	−0.6	−2.4	−3.4	−3.4		8.6
29	1.0	−0.2	3.4	1.6	2.4	2.4	−0.3	−0.8		9.0
30	1.8	3.6	4.6	8.2	4.2	0.8	2.4	1.4		11.2
31	3.9	6.5	3.1	4.1	5.3	−1.5	−0.5	−3.5		14.3
32	0.9	−1.9	−0.5	−2.7	−3.7	−13.1	−7.7	−8.1		31.3
33	5.8	5.6	7.6	6.4	3.4	1.8	5.6	−0.2		13.2
34	−7.8	−8.0	−3.2	−7.2	−9.6	−9.6	−9.0	−9.8		16.8
35	−3.9	−1.7	1.7	−2.3	−1.7	−5.7	−2.9	−7.7		18.7
36	−2.2	−4.4	−4.8	−4.6	−5.2	−7.2	−5.4	−6.6		14.6

and

3.34 and −0.33 are the least-squares estimates of the intercept and slope, respectively.

The t-statistic for these data is $t_c = -8.59$ with 168 degrees of freedom, which is significant at the 5-percent level. This equation holds over the range of average opacities studied, from 5 to 30 percent. By substituting values into this equation, it is possible to predict what the deviation would be at a particular true average opacity. This is done for selected opacities in Table 7.

TABLE 7. PREDICTED DEVIATIONS, TRAINING GENERATOR, WHITE SMOKE

Percent Opacity	Estimated Deviation
5	1.81
10	0.16
15	−1.49
20	−3.14
25	−4.79
30	−6.44

2. Precision

The precision of the determination of the average opacity of a white plume from a training generator is determined through use of an Analysis of Variance (ANOVA) technique. From this, it is possible to estimate both the within-observer and the observer bias components, and using these, to estimate the between-observer component.

The precision estimates are obtained using only the determinations from the eight observers who completed all 20 runs, which enables an estimate of the within-observer variance to be made without blocking the data. The details of the ANOVA are given and the estimate obtained in Appendix B.2.

The within-observer variance, σ^2, is estimated as

$$\hat{\sigma}^2 = 5.67$$

with 133 df. This gives an estimated within-observer standard deviation of

$$\hat{\sigma} = \sqrt{5.67}$$

$$= 2.38 \text{ percent opacity.}$$

The F-ratio for observers is $F_c = 4.16$. This exceeds the tabled value of 2.75, approximately, from the F distribution with 7 and 133 df, at the 5-percent significance level. Thus, there is a significant observer bias variance, σ_L^2, and it is estimated by

$$\hat{\sigma}_L^2 = 0.90$$

with 7 df. The estimated observer bias standard deviation, then, is

$$\hat{\sigma}_L = \sqrt{0.90}$$

$$= 0.95 \text{ percent opacity.}$$

The between-observer variance, σ_b^2, is estimated from the components above as

$$\hat{\sigma}_b^2 = \hat{\sigma}^2 + \hat{\sigma}_L^2$$

$$= 5.67 + 0.90$$

$$= 6.57.$$

This gives a between-observer standard deviation of

$$\hat{\sigma}_b = \sqrt{6.57}$$

$$= 2.56 \text{ percent opacity.}$$

3. Predicted Results

By combining the above results, it is possible to predict confidence limits for the average opacity determined by an observer. Let k be any true opacity level. Then the average opacity reported by the observer could be expected to be within

$$\text{expected range: } k + (a + bk) \pm (1.96)\,\hat{\sigma} \text{ percent opacity}$$

$$: k + (3.46 - 0.33k) \pm (1.96)\,(2.38) \text{ percent opacity}$$

$$: (0.67k + 3.46) \pm 4.66 \text{ percent opacity}$$

at the 95-percent confidence level. The ranges for selected values of k are given in Table 8.

TABLE 8. PREDICTED DETERMINATIONS, TRAINING GENERATOR-WHITE SMOKE

Percent Opacity	Expected Range
5	2.15–11.47
10	5.50–14.82
15	8.85–18.17
20	12.20–21.52
25	15.55–24.87
30	18.90–28.22
35	22.25–31.57

Similarly, an estimate can be made of the difference expected between two determinations of the same average plume opacity made by independent observers. The difference between determinations has a variance of $2\,\sigma_b^2$. Thus, the maximum expected difference would be no more than

$$\text{maximum difference: } \pm (1.96)\sqrt{2\,\hat{\sigma}_b^2}$$

$$: \pm (2.77)\,\hat{\sigma}_b$$

$$: \pm (2.77)\,(2.56)$$

$$: \pm 7.09 \text{ percent opacity}$$

at the 95 percent confidence level.

B. Black Smoke

The 133 determinations made using a black plume are shown in Table 9 along with the average of the transmissometer readings taken. These determinations are examined for accuracy and precision in the same manner as were the white plume determinations.

1. Accuracy

The deviations from the transmissometer average are shown in Table 10. As before, there is a tendency for the deviation to be greater as the true value increases. To obtain the best estimate of the deviation, a linear regression of the deviations on the true opacity is done.

For these data, the line that best fits these points is

$$\hat{Y}_i = 3.74 - 0.34X_i$$

TABLE 9. DETERMINATIONS OF AVERAGE OPACITY OF TRAINING GENERATOR, BLACK SMOKE

(Percent Opacity)

Run No.	Observer 1	2	3	4	5	6	7	8	9	Meter
11	24.4	20.4	22.0	23.2	19.0	21.4	21.6	21.8	25.4	35.0
12	7.4	9.6	10.0	10.4	10.6	7.6	11.2	8.4	12.0	10.8
13	20.0	20.2	22.6	24.0	22.2	20.8	23.2	21.8	23.0	26.0
14	17.6	14.4	15.6	19.6	17.2	19.6	20.0	20.2	17.0	20.0
15	7.4	11.4	12.6	14.0	12.8	12.0	13.0	13.0	12.4	12.3
16	9.6	5.2	5.8	4.6	8.8	5.6	4.8	4.8		16.0
17	13.6	11.0	12.0	11.0	10.4	12.0	11.2	12.4		11.0
18	13.2	14.0	13.4	14.2	10.0	13.8	11.6	11.8		17.6
19	12.6	10.2	16.4	16.2	11.4	14.8	11.2	15.0		17.6
20	12.0	10.8	13.8	14.0	11.6	12.4	11.8	13.6		13.6
21	18.6	18.2	17.2	21.0	17.2	23.0	17.8	19.2		18.0
22	29.2	17.0	25.2	29.4	29.4	31.2	30.0	28.6		28.2
23	14.8	11.4	15.8	17.4	15.0	17.4	13.4	13.0		13.8
24	7.8	8.0	11.4	11.4	10.0	11.6	8.4	8.6		9.0
25	15.0	12.2	15.2	20.4	15.8	16.6	15.8	15.6		14.8
26	22.0	20.2	23.2	24.2	23.0	24.4	22.0	19.8		22.2

TABLE 10. DEVIATION FROM TRAINING GENERATOR, BLACK SMOKE

(Percent Opacity)

Run No.	Observer 1	2	3	4	5	6	7	8	9	Meter
11	−10.6	−14.6	−13.0	−11.8	−16.0	−13.6	−13.4	−13.2	−9.6	35.0
12	−3.4	−1.2	−0.8	−0.4	−0.2	−3.2	0.4	−2.4	1.2	10.8
13	−6.0	−5.8	−3.4	−2.0	−3.8	−5.2	−2.8	−4.2	−3.0	26.0
14	−2.4	−5.6	−4.4	−0.4	−2.8	−0.4	0.0	0.2	−3.0	20.0
15	−4.9	−0.9	0.3	1.7	0.5	−0.3	0.7	0.7	0.1	12.3
16	−6.4	−10.8	−10.2	−11.4	−7.2	−10.4	−11.2	−11.2		16.0
17	2.6	0.0	1.0	0.0	−0.6	1.0	0.2	1.4		11.0
18	−4.4	−3.6	−4.2	−3.4	−7.6	−3.8	−6.0	−5.8		17.6
19	−5.0	−7.4	−1.2	−1.4	−6.2	−2.8	−6.4	−2.6		17.6
20	−1.6	−2.8	0.2	0.4	−2.0	−1.2	−1.8	0.0		13.6
21	0.6	0.2	−0.8	3.0	−0.8	5.0	−0.2	1.2		18.0
22	1.0	−11.2	−3.0	1.2	1.2	3.0	1.8	0.4		28.2
23	1.0	−2.4	2.0	3.6	1.2	3.6	−0.4	−0.8		13.8
24	−1.2	−1.0	2.4	2.4	1.0	2.6	−0.6	−0.4		9.0
25	0.2	−2.6	0.4	5.6	1.0	1.8	1.0	0.8		14.8
26	−0.2	−2.0	1.0	2.0	0.8	2.2	−0.2	−2.4		22.2

where

$\hat{Y}_i$ – the predicted deviation

X_i – the average opacity,

in the range from 9 to 35 percent average opacity. This equation is almost identical to that of the white smoke data. The value of t_c is −7.20, with 131 degrees of freedom which is significant at the 5-percent level.

TABLE 11. PREDICTED DEVIATION FROM METERED OPACITY, TRAINING GENERATOR–BLACK

Average Opacity	Predicted Deviation
10	0.44
15	−1.21
20	−2.86
25	−4.51
30	−6.16
35	−7.81

Using this equation, deviations may be estimated for plumes within the applicable range. This is done in Table 11.

2. Precision

The precision estimation for the black smoke determinations is similar to that done for the white smoke data. The results from the collaborator who did not complete all 16 runs are not included. The remaining 128 determinations of average opacity are used in an ANOVA to estimate the precision components. The details are presented in Appendix B.2.

The within-observer variance, σ^2, is estimated as

$$\hat{\sigma}^2 = 3.38$$

with 105 df. This gives an estimated within-observer standard deviation of

$$\hat{\sigma} = \sqrt{3.38}$$

$$= 1.84.$$

The F-ratio for observers is $F_c = 5.73$ which exceeds the critical value, $F_{.05}\ (7, 105) \approx 2.10$ at the 5-percent level. Thus the observer bias variance, σ_L^2, is significantly different from zero, and is estimated by

$$\hat{\sigma}_L^2 = 1.00$$

with 7df. Thus, the observer bias standard deviation is estimated by

$$\hat{\sigma}_L = 1.00.$$

The between-observer variance is estimated from the above as

$$\hat{\sigma}_b^2 = \hat{\sigma}^2 + \hat{\sigma}_L^2$$

$$= 3.38 + 1.00$$

$$= 4.38.$$

The estimated between-observer standard deviation, then, is

$$\hat{\sigma}_b = \sqrt{4.38}$$

$$= 2.09.$$

3. Predicted Results

Using the results above, it is possible to predict, at the 95-percent confidence level, the range of opacities that would occur in the use of the method at particular opacity levels. The predicted determination at k-percent opacity would be

$$k + (a + bk) = k + (3.74 - 0.34k)$$

$$= 0.66k + 3.74.$$

The variability in a reading would be measured by the within-observer standard deviation. For the 95-percent confidence level and 105 df, the factor 1.96 is used. This gives

$$\text{expected range: } (0.66k + 3.74) \pm 1.96\,\hat{\sigma}$$

$$: (0.66k + 3.74) \pm 1.96\,(1.84)$$

$$: (0.66k + 3.74) \pm 3.61.$$

The expected range is given for selected values of k in Table 12.

TABLE 12. PREDICTED DETERMINATIONS OF AVERAGE OPACITY, TRAINING GENERATOR–BLACK

Percent Opacity	Expected Range
5	3.43–10.65
10	6.73–13.95
15	10.03–17.25
20	13.33–20.55
25	16.63–23.85
30	19.93–27.15
35	23.23–30.45

The maximum difference between two determinations made by independent collaborators is determined from these data at the 95-percent confidence level. The factor 1.96 is again used, and this gives

$$\text{maximum difference: } \pm 1.96\sqrt{2\,\sigma_b^2}$$

$$\pm (2.77)\,\hat{\sigma}_b$$

$$\pm (2.77)\,(2.09)$$

$$\pm 5.79 \text{ percent opacity.}$$

at the 95-percent confidence level.

V. STAUFFER CHEMICAL COMPANY TEST

The second test of the method was conducted at the Stauffer Chemical Company sulfuric acid plant. This test was specifically designed to test the ability of a qualified observer to accurately read the plume opacity near the compliance limit for that source. The standard for sulfuric acid plants is 10 percent, meaning that an opacity of 10 percent or greater is in violation.[1]

The observers were chosen both from government enforcement agencies and from the private sector. This was done to investigate whether the affiliation of the individual influenced his determination of opacity.

There were 14 runs made on the first day of the test by 11 collaborators, and 16 runs on the second with 9 collaborators for a total of 298 determinations. The average opacity during each run was determined from a strip chart readout of opacity as determined by an in-stack transmissometer. Because of this, the "true" opacity could only be determined to the nearest percent, and the observers' determinations were rounded to the nearest percent opacity for comparison.

The observers made one deviation from the published method on the second day. Since the plume was light, they suggested that they read to the nearest 1 percent rather than the nearest 5 percent, and this was permitted. However, since the method specified that 5-percent increments be used, the individual readings subsequently were rounded to the nearest 5 percent prior to being averaged to obtain the determinations used. While there may be a difference between these and a Method 9 determination, it is felt that it would be slight.

The determinations are shown in Table 13 along with the average meter opacity for that range. These values are investigated for both their accuracy and precision, as for the smoke generator test.

A. Accuracy

The deviations from the transmissometer average are shown in Table 14 for the sulfuric acid plant data. The deviations show a negative bias, falling generally 1 to 3 percent below the metered opacity. To determine if there is a linear relationship between the deviations and the meter average, a regression line is fit to these points.

The equation of the best fit to these points is

$$\text{deviation} = -2.36 + (0.05)(\text{meter average}).$$

From the slope of the line, it can be seen that there is a slightly positive trend to these deviations compared to the negative trend of the smoke generator test.

To determine if the slope is significant, a t-statistic is calculated as described in Appendix B.1. The value of the statistic is $t_c = 0.68$ with 296 df. This is not significant, which implies that the slope, 0.05, is not significantly different from zero. Thus, no linear relationship can be said to exist between deviation and opacity for these data. The best estimate of the deviation, then, is the mean deviation of −2.0 percent opacity. This is valid in the range studied, from 2 to 15 percent opacity. However, since the majority of the runs were made between 5 and 10 percent average opacity, it is possible that there is a relationship between the two variables that would be noticeable if a wider range of opacities was studied.

B. Precision

The determinations are submitted to an ANOVA procedure to estimate the precision of the method at a sulfuric acid plant. The ANOVA was performed using only the nine observers who completed all 30 runs. The remaining data set is large enough to obtain the desired estimates, and the necessity for blocking is eliminated by their exclusion.

TABLE 13. DETERMINATIONS OF AVERAGE OPACITY, STAUFFER CHEMICAL

(Percent Opacity)

Run No.	Observer 1	2	3	4	5	6	7	8	9	10	11	Meter
1	7	3	7	6	5	5	4	5	5	7	3	9
2	5	1	7	5	5	5	4	4	4	7	2	8
3	5	0	7	4	3	5	4	4	4	3	3	8
4	5	0	6	3	1	5	4	3	4	3	2	8
5	7	2	6	5	4	5	5	5	6	6	4	8
6	5	3	6	5	2	5	5	3	6	6	3	8
7	5	3	6	5	5	5	5	3	6	6	2	8
8	5	3	6	5	5	5	5	4	6	5	2	8
9	18	10	15	15	12	14	11	9	14	12	13	14
10	7	5	8	7	10	8	9	5	8	9	8	7
11	1	0	3	1	1	4	2	0	5	4	1	5
12	5	2	5	5	4	5	6	3	5	5	2	6
13	5	2	4	5	3	5	5	4	5	5	1	6
14	5	3	5	3	5	5	7	4	5	5	3	6
15	0	0	1	0	0	0	3	0	3			2
16	0	0	1	1	3	4	4	1	5			2
17	0	0	1	2	1	3	1	0	3			5
18	0	0	0	0	0	1	1	0	4			4
19	9	8	8	10	12	14	7	9	9			10
20	6	9	7	12	13	14	6	10	8			10
21	5	9	5	6	10	10	5	6	6			10
22	5	7	5	5	5	6	5	5	5			10
23	3	4	4	5	6	7	2	5	5			7
24	3	5	3	3	5	5	2	5	5			7
25	4	6	5	6	7	7	2	9	5			7
26	5	9	7	10	13	13	4	11	6			9
27	8	9	7	13	13	16	5	8	7			9
28	6	9	6	16	17	17	4	9	5			9
29	6	10	8	15	13	16	5	14	7			9
30	7	10	5	16	15	15	5	14	6			9

The precision estimates are taken from the ANOVA table in Appendix B.3. The within-observer component is estimated by:

$$\hat{\sigma}^2 = 4.51$$

which has 232 df associated with it. This gives an estimated within-observer standard deviation of

$$\hat{\sigma} = \sqrt{4.51}$$

$$= 2.12 \text{ percent opacity}$$

The observer term for these data is shown in Appendix B.3 to be significantly different from zero. The observer bias variance, σ_L^2, is estimated from the ANOVA table as

$$\hat{\sigma}_L^2 = 0.93$$

with 8 df. This gives an estimated observer bias standard deviation of

$$\hat{\sigma}_L = \sqrt{0.93}$$

$$= 0.96 \text{ percent opacity}$$

TABLE 14. DEVIATIONS FROM METER AVERAGE, STAUFFER CHEMICAL

(Percent Opacity)

Run No.	Observer											Meter
	1	2	3	4	5	6	7	8	9	10	11	
1	−2	−6	−2	−3	−4	−4	−5	−4	−4	−2	−6	9
2	−3	−7	−1	−3	−3	−3	−4	−4	−4	−1	−6	8
3	−3	−8	−1	−4	−5	−3	−4	−4	−4	−5	−5	8
4	−3	−8	−2	−5	−7	−3	−4	−5	−4	−5	−6	8
5	−1	−6	−2	−3	−4	−3	−3	−3	−2	−2	−4	8
6	−3	−5	−2	−3	−6	−3	−3	−5	−2	−2	−4	8
7	−3	−5	−2	−3	−3	−3	−3	−5	−2	−2	−6	8
8	−3	−5	−2	−3	−3	−3	−3	−4	−2	−3	−6	8
9	4	−4	1	1	−2	0	−3	−5	0	−2	−1	14
10	0	−2	1	0	3	1	2	−2	1	2	1	7
11	−4	−5	−2	−4	−1	−1	−3	−5	0	−1	−4	5
12	−1	−4	−1	−1	−2	−1	0	−3	−1	−1	−4	6
13	−1	−4	−2	−1	−3	−1	−1	−2	−1	−1	−5	6
14	−1	−3	−1	−3	−1	−1	1	−2	−1	−1	−3	6
15	−2	−2	−1	−2	−2	−2	−1	−2	1			2
16	−2	−2	−1	−1	1	2	2	−1	3			2
17	−5	−5	−4	−3	−4	−2	−4	−5	−2			5
18	−4	−4	−4	−4	−4	−3	−3	−4	0			4
19	−1	−2	−2	0	2	4	−3	−1	−2			10
20	−4	−1	−3	2	3	4	−4	0	−2			10
21	−5	−1	−5	−4	0	0	−5	−4	−4			10
22	−5	−3	−5	−5	−5	−4	−5	−5	−5			10
23	−4	−3	−3	−2	−1	0	−5	−2	−2			7
24	−4	−2	−4	−4	−2	−2	−5	−2	−2			7
25	−3	−1	−2	−1	0	0	−5	2	−2			7
26	−4	0	−2	1	4	4	−5	2	−3			9
27	−1	0	−2	4	4	7	−4	−1	−2			9
28	−3	0	−3	7	8	8	−5	0	−4			9
29	−3	1	−1	6	4	7	−4	5	−2			9
30	−2	1	−4	7	6	6	−4	5	−3			9

The between-observer variance, σ_b^2, is the sum of the within-observer and observer bias components. The estimated value is

$$\hat{\sigma}_b^2 = \hat{\sigma}_L^2 + \hat{\sigma}^2$$

$$= 0.93 + 4.51$$

$$= 5.44.$$

This results in a between-observer standard deviation of

$$\hat{\sigma}_b = \sqrt{5.44}$$

$$= 2.33 \text{ percent opacity.}$$

C. Predicted Results

The above results can be used to make statements concerning the results that would be obtained by an independent observer. The range of opacities that could be expected is calculated by using the best estimate of the deviation, along with the variation that could be expected.

At k percent opacity, the expected result would be

$$(k-2)\% \pm (1.96)\hat{\sigma}$$

at the 95-percent confidence level. The factor 1.96 is used since with 232 df, the distribution of the t statistic is essentially a standard normal curve. Substituting in the value of $\hat{\sigma}$ gives:

$$\begin{aligned} \text{expected range}: &\quad (k-2)\% \pm (1.96)(2.12) \\ : &\quad (k-2) \pm 4.16 \\ : &\quad k-6.16 \text{ to } k+2.16 \end{aligned}$$

for plumes of from 2 to 15 percent average opacity. For selected values, the expected range is calculated and shown in Table 15.

TABLE 15. PREDICTED DETERMINATIONS, SULFURIC ACID PLANT

Avg Opacity	Expected Range
5	0*–7.16
10	3.84–12.16
15	8.84–17.16

*Truncated to zero.

The maximum expected difference between determinations by two independent observers is calculated using the between-observer variance term. As before, the difference is

$$\begin{aligned} \text{maximum difference} &= (2.306)\sqrt{2\hat{\sigma}_b^2} \\ &= (3.26)\hat{\sigma}_b \end{aligned}$$

at the 95-percent confidence level. The factor 2.306 comes from Student's t distribution with 8 df, due to the observer bias term. Substituting gives

$$\begin{aligned} \text{maximum difference} &= (3.26)(2.33) \\ &= 7.60 \text{ percent opacity} \end{aligned}$$

D. Comparison of Enforcement and Private Observers

The test personnel were chosen from both an enforcement agency and from private concerns to allow the determinations by each group to be compared. The nine collaborators who completed all 30 runs are used, and the mean over the 30 runs is computed. The observer means are shown in Table 16.

TABLE 16. OBSERVER MEANS, SULFURIC ACID PLANT

Observer	Mean
1	5.00
2	4.40
3	5.47
*4	6.47
*5	6.60
*6	7.63
7	4.57
*8	5.40
9	5.73

*Observer from enforcement agency.

The observers from the enforcement agency were numbers 4, 5, 6 and 8. As can be seen, this group gives the three largest overall means and the sixth largest. The indication, then, is that these people tend to read a higher opacity than the private sector people.

The hypothesis that the two groups read the same average opacity is tested in Appendix B.3. The test is a contrast among the means, and the test statistic follows Student's t distribution. The statistic is $t_c = 5.69$ with 232 df. This is a significant value and the conclusion is that the enforcement observers did indeed determine higher average opacities overall than the other group.

The average opacity for all 30 runs, however, was 7.6, so that the enforcement observers were closer to the true opacity than those from other agencies.

VI. RIVERBEND STEAM STATION TEST

The three tests at the Riverbend Steam Station are treated separately in the analysis since there were different groups of collaborators and varying sky conditions for the three tests.

The original test plan was designed to allow the evaluation of the method as used for the determination of minutes of noncompliance with a particular standard.(2) A test run was set as 10 minutes of observation, and the reported variable was to be how many observations exceeded a certain standard.

The experimental factors to be studied included the angle of observation and the relative experience of the observer. Variations to the method to be evaluated included reading in 1 percent rather than 5 percent increments and averaging the responses of two observers as opposed to a single observer's result.

In the third testing period, the test personnel were advised that the average opacity, based on 25 readings, was the test variable of interest. The other factors for the test remained unchanged. For purposes of comparison, average opacities were calculated from the first two data sets using the first 25 observations of each run.

Due to the adverse sky and wind conditions during tests 1 and 2, not all of the planned evaluations are useful. There was an inability to read the low opacity plumes against the type of background that existed, and as a result, the determinations were generally well below the concurrent meter average. The precision estimated, however, is independent of the accuracy of the determination.

The observers at each test were divided into two groups for the test, a control and an experimental. The control group observed the plume at all times from a position consistent with the method as written and read in increments of 5 percent. The experimental group either read the plume from a more extreme angle, in increments of 5 percent or from the same angle as the control but in increments of 1 percent. Each group was composed both of observers who had considerable field experience with the method and of observers who had relatively little such experience.

A. Test 1

There were 10 runs made at the first test by 10 observers. The determinations of average opacity calculated from the observer record sheets are shown in Table 17. These determinations will be evaluated with respect to their accuracy and precision. There are six missing determinations in the data set. These

TABLE 17. AVERAGE OPACITY DETERMINATIONS, STEAM STATION TEST 1

Run	Control Group					Meter	Experimental Group				
	1	2	3	4	5		1	2	3	4	5
1	6.4	4.4	6.2	5.2	5.8	16.0	5.4	5.4	1.8	5.8	2.2
2	13.4	8.2	11.2	9.6	12.8	23.5	6.6	6.2	3.8	6.4	4.4
3	11.6	11.0	10.6	13.8	12.6	26.0	10.8	9.4	7.8	13.0	8.8
4	5.4	4.4	6.0	6.8	4.8	11.6	5.2	5.5	4.1	5.8	6.5
5	8.4	6.4	7.8	6.6	11.4	15.7	7.6	5.8	5.9	5.9	9.5
6	14.8	7.6	12.4	14.6	9.0	26.1	12.2	8.5	7.5	15.0	8.9
7	17.6	9.6	12.4	14.2	11.4	24.4	7.5	9.4	4.6	8.6	7.0
8	9.4	3.2	4.0	9.6	1.8	14.2	4.4	6.0	1.0	5.0	2.2
9	1.2	0.2	2.4	5.6	0.6	13.0	1.6	3.0	0.2	1.0	0.4
10	*	*	*	*	*	13.4	2.0	3.0	0.0	0.0	*

*Incomplete data sheet.

occur when due to sky and wind conditions and interference from another plume, the observer felt that at times he was unable to make a determination of the opacity, and in these instances, a direct comparison would not be valid.

1. Accuracy

It is apparent that these determinations are generally considerably lower than the meter average. The deviations from the meter average are shown in Table 18.

TABLE 18. DEVIATIONS FROM METERED OPACITY, STEAM STATION TEST 1

Run	Control Group					Meter	Experimental Group				
	1	2	3	4	5		1	2	3	4	5
1	-9.6	-11.6	-9.8	-10.8	-10.2	16.0	-10.6	-10.6	-14.2	-10.2	-13.8
2	-10.1	-15.3	-12.3	-13.9	-10.7	23.5	-16.9	-7.3	-19.7	-17.1	-19.1
3	-14.4	-15.0	-15.4	-12.2	-13.4	26.0	-15.2	-16.6	-18.2	-13.0	-17.2
4	-6.2	-7.6	-5.6	-4.8	-6.8	11.6	-6.4	-6.1	-7.5	-5.8	-5.1
5	-7.3	-9.3	-7.9	-9.1	-4.3	15.7	-8.1	-9.9	-9.8	-9.8	-6.2
6	-11.3	-18.5	-13.7	-11.5	-17.1	26.1	-13.9	-17.6	-18.6	-11.1	-17.2
7	-6.8	-14.8	-12.0	-10.2	-13.0	24.4	-16.9	-15.0	-19.8	-15.8	-17.4
8	-4.8	-11.0	-10.2	-4.6	-12.4	14.2	-9.8	-8.2	-13.2	-9.2	-12.0
9	-11.8	-12.8	-10.6	-7.4	-12.4	13.0	-11.4	-10.0	-12.8	-12.0	-12.6
10	*	*	*	*	*	13.4	-11.4	-10.4	-13.4	-13.4	*

*Incomplete Data Sheet.

To assess the accuracy of the method under these conditions the control group data is used. These observers had taken the optimum viewing position consistent with the method, and thus are considered to best represent the accuracy that could be expected of the method.

The deviations are fit to a regression model as with the data from the previous tests. The least-squares fit to these points is given by

$$\text{deviation} = -3.01 - 0.40\,(\text{meter average}).$$

The t statistic for this fit is $t_c = 5.89$ with 43 df. This far exceeds the critical value of $t_{.05}(43) \approx 2.02$, and this equation may be said to be valid in the range up to about 25 percent opacity.

TABLE 19. PREDICTED DEVIATION FROM METERED OPACITY, STEAM STATION TEST 1

Percent Opacity	Predicted Deviation
5	-5.01
10	-7.01
15	-9.01
20	-11.01
25	-13.01

Substituting into this equation, it is possible to predict what the deviation would be at selected opacities in the applicable range. This is done for the 5 percent increments and the results shown in Table 19. As can be seen, a plume opacity up to 10 percent is barely distinguishable from zero under these conditions.

2. Precision

The precision estimates are obtained using both groups' determinations from runs where the experimental group was reading in 5-percent increments. The 1-percent increment data were intended only to evaluate that variation as a possible improvement to the published method when working on low opacity plumes, and due to the viewing conditions, no evaluation using these data is performed. The precision estimation is shown in Appendix B.4, using runs 1, 2, 3, 7, 8 and 9, omitting run 10 where there were 6 incomplete determinations.

The estimated within-observer variance is

$$\hat{\sigma}^2 = 3.31$$

with 45 df. This gives an estimated standard deviation of

$$\hat{\sigma} = 1.82 \text{ percent opacity}$$

The observers are determined to be a significant source of difference between observed values. The F-ratio is 4.50 which exceeds the critical value of 4.08 with 8 and 54 df. Thus, σ_L^2 is estimated as

$$\hat{\sigma}_L^2 = 1.93$$

with 8 df. The estimated observer bias standard deviation is

$$\hat{\sigma}_L = \sqrt{1.93}$$

$$= 1.39 \text{ percent opacity.}$$

Combining the above estimates, the between-observer component can be estimated. The between-observer variance, σ_b^2, is estimated by

$$\hat{\sigma}_b^2 = \hat{\sigma}^2 + \hat{\sigma}_L^2$$

$$= 3.31 + 1.93$$

$$= 5.24.$$

Thus, the estimated between-observer standard deviation is

$$\hat{\sigma}_b = \sqrt{5.24}$$

$$= 2.29 \text{ percent opacity}$$

3. Predicted Results

Using the results of the previous two sections, it is possible to estimate the range of opacities that could be expected from a qualified observer determining average opacity under these conditions.

The expected range at any level, k, of true opacity would be given by

$$\text{expected range: } k + [a + bk] \pm (1.96)\,\hat{\sigma}$$

at the 95-percent confidence level. Substituting gives

$$\text{expected range: } k + [-3.01 + 0.40k] \pm 1.96\,(1.82)$$

$$: [0.60k - 3.01] \pm 3.57.$$

The range is calculated for selected opacities and shown in Table 20. As expected, there are no instances where an accurate determination can be expected under these conditions.

The maximum difference that could be expected between two observers is estimated using the between-observer standard deviation. The difference would not be expected to exceed

TABLE 20. PREDICTED DETERMINATIONS, STEAM STATION TEST 1

Percent Opacity	Expected Range
5	0* - 3.56
10	0* - 6.56
15	2.42 - 9.56
20	5.42 - 12.56
25	8.42 - 15.56

*Truncated to zero.

TABLE 21. OBSERVER MEANS FOR STEAM STATION TEST 1

Group	Observer	Mean Determination
Control	1	9.93
	2	6.10
	3	7.80
	4	9.67
	5	7.50
Experimental	1	6.05
	2	6.57
	3	3.20
	4	6.63
	5	4.17

$$\text{maximum difference} = t_{0.05}\,(8)\sqrt{2}\hat{\sigma}_b$$

$$= (2.306)(1.414)(2.29)$$

$$= 7.49 \text{ percent opacity}$$

at the 95-percent confidence level.

4. Experience Effect

The means of the observers are shown in Table 21 and these are used to determine if there is any significant difference between the average opacity as determined by an experienced as opposed to an inexperienced observer. The test statistic is a t-statistic based upon a contrast among the means. The details are shown in Appendix B.4.

The calculated t-statistic is

$$t_c = 1.39$$

with 45 df. There can be said to be a difference between the two groups at the 5-percent level only if t_c exceeds the tabled value. Since $t_{0.05}\,(45) \approx 2.016$, the difference is not significant at the 5-percent level. Thus, for these conditions, the relative experience of the observer had no noticeable effect.

B. Test 2

There were 18 runs made at the second test period with 10 collaborators participating. The personnel breakdown was identical to the first test, with 4 experienced and 6 relatively inexperienced observers. One run, run 10, could not be completed. During the course of the run, the opacity meter began its automatic calibration check. During this time, the meter displays only the calibration opacities, 0 and 95 percent, and no true value is available for comparison. Only ten readings were made before the run was halted, and thus no determinations could be calculated.

The determinations made by the observers are shown in Table 22. As in the previous test, missing determinations occur when an observer is unable to make a determination of the plume opacity due to wind and sky conditions.

TABLE 22. AVERAGE OPACITY DETERMINATIONS, STEAM STATION TEST 2

Run	Control Group 1	2	3	4	5	Meter	Experimental Group 1	2	3	4	5
1	11.2	5.8	11.6	10.6	4.2	14.4	8.0	6.1	4.3	11.7	8.3
2	13.8	14.4	13.2	14.8	5.6	13.2	13.7	7.2	9.4	11.3	8.6
3	10.0	9.2	9.2	12.0	5.2	11.2	13.0	5.2	3.4	4.2	9.0
4	5.8	5.4	5.6	6.2	3.4	5.6	4.6	1.4	0.2	3.4	4.0
5	20.0	15.0	12.6	16.2	13.6	16.4	20.0	13.8	11.4	17.8	19.0
6	5.0	0.4	4.4	1.0	0.0	6.2	0.4	1.6	0.0	0.0	3.0
7	13.8	11.4	9.4	14.2	10.4	11.6	*	8.0	6.6	*	10.6
8	5.0	5.4	6.6	5.2	0.8	9.5	1.8	5.6	0.6	4.2	3.8
9	14.4	10.6	10.8	15.2	10.6	15.0	8.6	6.0	5.4	9.4	8.8
†10											
11	5.2	4.0	5.0	7.6	4.4	7.5	4.4	5.0	0.1	2.6	3.6
12	5.6	2.2	7.8	6.4	2.6	10.8	2.7	3.4	0.0	2.5	1.8
13	5.2	0.6	5.0	0.0	0.0	4.7	0.0	0.0	0.0	0.0	0.0
14	9.4	7.0	7.6	11.8	7.4	10.1	6.6	4.8	0.6	6.4	2.0
15	17.0	12.8	16.8	21.8	15.2	23.6	13.6	13.0	9.8	15.6	12.3
16	7.8	7.0	10.8	9.8	3.2	13.5	7.0	6.0	2.3	5.4	4.0
17	12.4	8.0	10.0	12.4	8.2	13.4	8.4	7.6	1.8	10.0	7.0
18	7.6	6.0	6.4	7.6	5.0	13.6	5.0	6.0	2.2	7.8	6.6

*Incomplete data sheet.
†Run 10 could not be completed.

1. Accuracy

The deviations of the determinations from the meter average are shown in Table 23. As before, the determinations are generally below the meter due in large part to the viewing conditions. A linear regression of the deviations is used to assess the accuracy, and as in the first test, only the control group data is used.

The best fit to these points is given by

$$\text{deviation} = -1.02 - 0.19\,(\text{meter average})$$

for the range from 5 to 25 percent average opacity. The t statistic for this model is $t_c = 2.56$ with 83 df, which is significant at the 5-percent level.

Using the above equation, it is possible to predict what deviation would be expected at various opacity levels. This is done in Table 24 for selected values in the applicable range. These are somewhat better than those from the first test, but still have a strong negative bias at the low range of opacities.

2. Precision

The precision components for this test will be estimated using runs 3–9, 13, 14, 17 and 18. The remaining runs were those where the experimental group was reading in 1-percent increments, and a direct comparison between the two sets of readings is not proper.

The precision components are estimated by an ANOVA procedure comparable to that for the first test. The details and ANOVA table are shown in Appendix B.5. The within-observer component is estimated from the error term of the ANOVA.

TABLE 23. DEVIATIONS FROM METERED OPACITY, STEAM STATION TEST 2

(Percent Opacity)

Run	Control Group					Meter	Experimental Group				
	1	2	3	4	5		1	2	3	4	5
1	−3.2	−8.6	−2.8	−3.8	−10.2	14.4	−6.4	−8.3	−10.1	−2.7	−6.1
2	0.6	1.2	0.0	1.6	−7.6	13.2	0.5	−6.0	−3.8	−1.9	−4.6
3	−1.2	−2.0	−2.0	0.8	−6.0	11.2	1.8	−6.0	−7.8	−7.0	−2.2
4	0.2	−0.2	0.0	0.6	−2.2	5.6	−1.0	−4.2	−5.4	−2.2	−1.6
5	3.6	−1.4	−3.8	−0.2	−2.8	16.4	3.6	−2.6	−5.0	1.4	2.6
6	−1.2	−5.8	−1.8	−5.2	−6.2	6.2	−5.8	−4.6	−6.2	−6.2	−3.2
7	2.2	−0.2	−2.2	2.6	−1.2	11.6	*	−3.6	−5.0	*	−1.0
8	−4.5	−4.1	−2.9	−4.3	−8.7	9.5	−7.7	−3.9	−8.9	−5.3	−5.7
9	−0.6	−4.4	−4.2	0.2	−4.4	15.0	−6.4	−9.0	−9.6	−5.6	−6.2
11	−2.3	−3.5	−2.5	0.1	−3.1	7.5	−3.1	−2.5	−7.4	−4.9	−3.9
12	−5.2	−8.6	−3.0	−4.4	−8.2	10.8	−8.1	−7.4	−10.8	−8.3	−9.0
13	0.5	−4.1	0.3	−4.7	−4.7	4.7	−4.7	−4.7	−4.7	−4.7	−4.7
14	−0.7	−3.1	−2.5	1.7	−2.7	10.1	−3.5	−5.3	−9.5	−3.7	−8.1
15	−6.6	−10.8	−6.8	−1.8	−8.4	23.6	−10.0	−10.6	−13.8	−8.0	−11.3
16	−5.7	−6.5	−2.7	−3.7	−10.3	13.5	−6.5	−7.5	−11.2	−8.1	−9.5
17	−1.0	−5.4	−3.4	−1.0	−5.2	13.4	−5.0	−5.8	−11.6	−3.4	−6.4
18	−6.0	−7.6	−7.2	−6.0	−8.6	13.6	−8.6	−7.6	−11.4	−5.8	−7.0

*Incomplete Data Sheet.

TABLE 24. PREDICTED DEVIATION FROM METERED OPACITY, STEAM STATION TEST 2

Percent Opacity	Predicted Deviation
5	−1.97
10	−2.92
15	−3.87
20	−4.82
25	−5.77

The within-observer variance, σ^2, is estimated by

$$\hat{\sigma}^2 = 3.37$$

with 88 df. This gives an estimated standard deviation of

$$\hat{\sigma} = \sqrt{3.37}$$
$$= 1.84 \text{ percent opacity}$$

The observer-within-group term is used to determine if there is a significant observer bias component. The F-ratio is $F_c = 8.37$, which is significant at the 5-percent level. This is equivalent to saying that the observer bias component, σ_L^2, is greater than zero. The observer bias variance is estimated from the observer mean square as

$$\hat{\sigma}_L^2 = 2.26$$

with 8 df. The observer bias standard deviation is estimated by

$$\hat{\sigma}_L = \sqrt{2.26}$$
$$= 1.50 \text{ percent opacity}$$

The between-observer component is estimated from the above. The between-observer variance, σ_b^2, is estimated as

$$\hat{\sigma}_b^2 = \hat{\sigma}_L^2 + \hat{\sigma}^2$$
$$= 2.26 + 3.37$$
$$= 5.63.$$

This gives an estimated standard deviation of

$$\hat{\sigma}_b = \sqrt{5.63}$$
$$= 2.37 \text{ percent opacity}$$

3. Predicted Results

The accuracy and precision statements above can be combined to predict the performance that could be expected when using Method 9 under these conditions.

The range of opacities that could be expected from a qualified observer at k-percent opacity, at the 95-percent confidence level, is given by

$$\text{expected range: } k + [a + bk] \pm 1.96\,\hat{\sigma}$$

$$: k + [-1.02 - 0.19k] \pm 1.96\,(1.84)$$

$$: [0.81k - 1.02] \pm 3.61$$

Values for this expected range are shown in Table 25 for the range of opacities studied. As can be seen, the observer would be expected to be biased on the low side above 10 percent opacity.

TABLE 25. PREDICTED DETERMINATIONS, STEAM STATION TEST 2

Percent Opacity	Expected Range
5	0*–6.64
10	3.47–10.69
15	7.52–14.74
20	11.57–18.79
25	15.62–22.84

*Truncated to zero.

The between-observer standard deviation can be used to estimate the difference that could be expected between two observers. The maximum expected difference is calculated as

$$\text{maximum difference} = t_{0.05}(8)\sqrt{2}\,\hat{\sigma}_b$$

at the 95-percent confidence level. The 8 df comes from the observer bias term. Substituting gives

$$\text{expected difference} = (2.306)(1.414)(2.37)$$

$$= 7.73 \text{ percent opacity}$$

4. Experience Effect

The relative experience of an observer is investigated using the data from the ANOVA in Appendix B.5. The means of the observers over the 11 runs are calculated, and these are shown in Table 26.

TABLE 26. OBSERVER MEANS FOR STEAM STATION TEST 2

Group	Observer	Mean Determination
Control	1	9.48
	2	6.76
	3	7.82
	4	8.76
	5	5.42
Experimental	1	6.84
	2	5.20
	3	2.56
	4	6.32
	5	6.32

The 1% variation runs are excluded to avoid the possibility of this variation influencing the results. Run 7 is also excluded, due to the two missing observations.

The details of the test for significance are presented in Appendix B.5. The value of the test statistic is $t_c = 3.19$. From a Student's t distribution with 88 df, the critical value is approximately 1.99, and the hypothesis of equality between experienced and non-experienced observers is rejected since t_c exceeds 1.99.

C. Test 3

There were 8 collaborators for the third test. Of these, 4 were from state and county agencies, 2 were from EPA offices and 2 were from a private contractor. The observers were divided into two equal groups as in the two previous tests. There were 10 observation runs on the first day and 14 on the second for a total of 192 individual determinations. These are shown in Table 27, along with the meter average concurrently obtained.

TABLE 27. AVERAGE OPACITY DETERMINATIONS, STEAM STATION TEST 3

Run	Control Group 1	2	3	4	Meter	Experimental Group 1	2	3	4
1	5.0	5.2	3.0	5.0	3.1	4.2	4.6	2.1	3.1
2	13.6	13.6	15.4	5.4	7.5	12.2	8.6	6.1	10.8
3	5.0	5.0	4.0	5.0	4.2	4.2	3.0	2.4	5.0
4	9.0	8.4	8.6	5.2	7.2	7.8	7.0	1.8	6.8
5	10.4	6.8	8.8	5.2	6.7	9.2	6.4	3.8	7.8
6	6.6	6.6	5.4	5.0	5.0	5.6	6.0	4.0	4.8
7	11.0	10.2	8.4	5.6	5.4	5.4	5.6	4.0	6.0
8	5.8	6.0	5.8	5.2	3.4	3.2	2.0	2.0	1.6
9	7.6	5.0	5.4	5.0	4.2	3.0	1.6	3.3	4.5
10	6.8	6.0	5.4	5.0	4.8	3.8	2.0	2.3	5.3
11	25.0	23.0	36.2	16.8	21.2	15.3	13.5	16.5	10.2
12	40.8	39.0	33.0	39.6	36.0	25.2	28.7	25.7	18.1
13	11.2	9.8	5.6	6.2	10.3	10.2	9.0	8.6	7.4
14	17.0	15.2	6.4	10.2	14.2	18.0	14.2	12.2	8.8
15	12.8	10.8	7.8	10.8	12.2	14.4	16.6	9.6	12.8
16	26.0	25.6	19.0	25.2	29.6	29.8	26.6	24.2	21.6
17	13.4	12.4	11.2	7.6	14.6	16.6	16.4	10.0	12.8
18	29.2	25.2	24.0	25.0	28.3	30.0	30.6	22.0	27.2
19	23.8	23.0	18.4	21.2	34.6	21.8	25.6	15.8	21.6
20	23.0	19.4	16.6	12.4	27.2	22.8	21.6	13.5	18.8
21	33.8	34.6	30.0	33.2	36.9	34.6	40.2	31.6	31.8
22	17.4	12.6	10.0	8.0	15.2	14.0	13.4	11.8	12.2
23	23.8	19.8	18.0	19.2	27.8	21.2	25.2	22.2	24.8
24	12.6	15.0	9.4	6.8	14.4	15.0	9.2	8.8	12.8

To evaluate the observer performance at this test, the accuracy and precision will be estimated as for the first two sites. In addition, the relative performance of the tested variations to the method will be evaluated.

1. Accuracy

The evaluation of the accuracy of the method is performed using the deviations from the meter obtained from the control group data as shown in Table 28. These deviations were all from determinations made within the constraints of the method, and thus offer the most applicable information.

TABLE 28. DEVIATIONS FROM METERED OPACITY, STEAM STATION TEST 3

Run	Control Group 1	2	3	4	Meter	Experimental Group 1	2	3	4
1	1.9	2.1	−0.1	1.9	3.1	1.1	1.5	−1.0	0.0
2	6.1	6.1	7.9	−2.1	7.5	4.7	1.1	−1.4	3.3
3	0.8	0.8	−0.2	0.8	4.2	0.0	−1.2	−1.8	0.8
4	1.8	1.2	1.4	−2.0	7.2	0.6	−0.2	−5.4	−0.4
5	3.7	0.1	2.1	−1.5	6.7	2.5	−0.3	−2.9	1.1
6	1.6	1.6	0.4	0.0	5.0	0.6	1.0	−1.0	−0.2
7	5.6	4.8	3.0	0.2	5.4	0.0	0.2	−1.4	0.6
8	2.4	2.6	2.4	1.8	3.4	−0.2	−1.4	−1.4	−1.8
9	3.4	0.8	1.2	0.8	4.2	−1.2	−2.6	−0.9	0.3
10	2.0	1.2	0.6	0.2	4.8	−1.0	−2.8	−2.5	0.5
11	3.8	1.8	15.0	−4.4	21.2	−5.9	−7.7	−4.7	−11.0
12	4.8	3.0	−3.0	3.6	36.0	−10.8	−7.3	−10.3	−17.9
13	0.9	−0.5	−4.7	−4.1	10.3	−0.1	−1.3	−1.7	−2.9
14	2.8	1.0	−7.8	−4.0	14.2	3.8	0.0	−2.0	−5.4
15	0.6	−1.4	−4.4	−1.4	12.2	2.2	4.4	−2.6	0.6
16	−3.6	−4.0	−10.6	−4.4	29.6	0.2	−3.0	−5.4	−8.0
17	−1.2	−2.2	−3.4	−7.0	14.6	2.0	1.8	−4.6	−1.8
18	0.9	−3.1	−4.3	−3.3	28.3	1.7	2.3	−6.3	−1.1
19	−10.8	−11.6	−16.2	−13.4	34.6	−12.8	−9.0	−18.8	−13.0
20	−4.2	−7.8	−10.6	−14.8	27.2	−4.4	−5.6	−13.7	−8.4
21	−3.1	−2.3	−6.9	−3.7	36.9	−2.3	3.3	−5.3	−5.1
22	2.2	−2.6	−5.2	−7.2	15.2	−1.2	−1.8	−3.4	−3.0
23	−4.0	−8.0	−9.8	−8.6	27.8	−6.6	−2.6	−5.6	3.0
24	−1.8	0.6	−5.0	−7.6	14.4	0.6	−5.2	−5.6	−1.6

As was the case with the smoke generator test, there appears to be a relationship between the magnitude and direction of the deviation and the meter average. To test this, a linear regression model is fit in the manner of Appendix B.1. The least-squares line obtained is

$$\text{deviation} = 2.27 - (0.24)(\text{meter average}).$$

The t statistic for this test is $t_c = 6.10$ with 94 df, which is significant at the 5-percent level. This equation is valid in the range from 5 to 35 percent average opacity.

By substituting into the equation, it is possible to predict what the deviation would be at different levels of true average opacity. This is done in Table 29. These predicted deviations are comparable to those obtained in the smoke generator test. The observers are fairly accurate in the low opacity ranges, but acquire a low bias as the opacity increases.

For comparison, the deviations from the meter are calculated for the experimental group as well. The three cases, 45°, 90°, and 1 percent are treated separately and regression lines calculated for each case. These are:

TABLE 29. PREDICTED DEVIATION FROM METERED OPACITY, STEAM STATION TEST 3

Percent Opacity	Predicted Deviation
5	1.07
10	−0.13
15	1.33
20	−2.53
25	−3.73
30	−4.93
35	−6.13

$$\text{deviation} = -0.68 - 0.03\ (\text{opacity})$$

$$\text{deviation} = 1.34 - 0.19\ (\text{opacity})$$

$$\text{deviation} = 1.76 - 0.39\ (\text{opacity})$$

respectively, for the three cases. The first line does not have a significant slope, while the second two do. However, it is easy to see that the three slopes are different, and thus that the observers' determinations differed when the manner and position of observation differed.

It is interesting to note that the most accurate readings were made when the group was at an approximate 45° angle to the sun. The mean deviation was −1.18 over the range from 5–32 percent average opacity.

2. Precision

The model used for estimating the precision of the method is similar to that used in the two previous tests. The variance components are estimated from an ANOVA of all determinations made by both groups on the 45° and 90° runs. The ANOVA is presented in detail in Appendix B.6. Precision estimation using the 1 percent variation runs is contained in a separate evaluation.

The estimated within-observer variance is

$$\hat{\sigma}^2 = 3.57$$

with 90 df. The estimated within-observer standard deviation, then, is

$$\hat{\sigma} = \sqrt{3.57}$$

$$= 1.89 \text{ percent opacity}$$

The F-ratio for observers/groups is $F_c = 14.99$. From a table of the F distribution, the critical value at the 5-percent significance level is $F_{0.05}\ (6,90) \approx 2.30$. Thus, the observer/group term is significant, which is equivalent to saying that an observer bias variance term, σ_L^2, does exist. The observer bias variance is estimated by

$$\hat{\sigma}_L^2 = 3.12$$

with 6 df. The estimated observer standard deviation, then, is

$$\hat{\sigma}_L = \sqrt{3.12}$$

$$= 1.77 \text{ percent opacity.}$$

The between-observer variance is estimated from the above two terms as

$$\hat{\sigma}_b^2 = \hat{\sigma}_L^2 + \hat{\sigma}^2$$

$$= 3.12 + 3.57$$

$$= 6.69.$$

This gives a between-observer standard deviation of

$$\hat{\sigma}_b = \sqrt{6.69}$$

$$= 2.59 \text{ percent opacity.}$$

3. Predicted Results

The above results can be combined to give estimates of the type of average opacities that would be expected from field use of the method. The expected range of opacities that would be reported by an independent observer are obtained by using the predicted deviation and the within-observer standard deviation.

The range of opacities, at k percent opacity, is defined as

$$\text{expected range: } [k + \text{predicted deviation}] \pm (1.96)\hat{\sigma}$$

at the 95-percent confidence level. The normal variate 1.96 is used since there are 69 df for the within-observer component. Substituting gives

$$\text{expected range: } k\% + 2.27 - (0.24)k \pm (1.96)(1.89)$$

$$(0.76)k + 2.27 \pm (3.70) \text{ percent opacity}$$

TABLE 30. PREDICTED DETERMINATIONS, STEAM STATION TEST 3

Percent Opacity	Expected Range
5	2.37–9.77
10	6.17–13.57
15	9.97–17.37
20	13.77–21.17
25	17.57–24.97
30	21.37–28.77
35	25.17–32.57

Values for the expected range are given in Table 30 for 5-percent increments in the applicable range. Similar to the previous results, these ranges contain the true opacities at the low end, but acquire a negative bias beginning at the 25-percent level.

The between-observer component is used to estimate the difference that would be expected between two observers making independent observations of the plume opacity. The maximum difference that would be expected is

$$\text{maximum difference} = t_{0.05}\,(6)\sqrt{2\hat{\sigma}_b}$$

$$= (2.447)(1.414)(2.59)$$

$$= 8.96 \text{ percent opacity}$$

at the 95-percent confidence level.

The t value has 6 df from the observer bias variance term. This maximum difference is slightly larger than the numbers previously obtained.

4. Experience Effect

The observer means shown in Table 31, are used to test the hypothesis that the experienced observers read the same average opacity as the inexperienced. The test is a contrast among the means, and the details are given in Appendix B. 6.

TABLE 31. OBSERVER MEANS FOR STEAM STATION TEST 3

Group	Observer	Mean Determination
Control	1	15.31
	2	14.00
	3	11.40
	4	11.46
Experimental	1	14.95
	2	14.46
	3	11.19
	4	12.76

The test statistic is based on Student's t distribution, and compares the difference between the four experienced observers, 1, 2, 5, and 6, and the four inexperienced, 3, 4, 7, and 8. The value of the statistic is $t_c = 8.89$ with 69 df. The tabled value is $t_{0.05}(69) \approx 2.00$, and thus the hypothesis is clearly rejected. The conclusion, then, is that the experience observers read different, and in this case higher, average opacities than the inexperienced. By inspection of the deviations, it appears that this difference occurs mainly in the high opacity range ($\geqslant 25\%$) where the inexperienced observers read much lower than the experienced. As it was previously, though, the experienced observers compare more favorably to the meter average.

5. 1-Percent Increments

On runs 1, 2, 9, 10, 11, 12, 19 and 20, the experimental group read from the same viewing position as the control, but recorded their observations to the nearest 1 percent opacity rather than the nearest 5 percent. These runs are considered separately from the others in terms of their accuracy and precision.

a. Accuracy

As the regression line in Section VI.C.1 above indicates, there was a negative bias in the determinations. Using that line, predicted deviations are calculated for the applicable range, and these are calculated and presented in Table 32.

TABLE 32. PREDICTED DEVIATIONS, 1-PERCENT INCREMENT DATA

Opacity	Expected Deviation
5	–0.19
10	–2.14
15	–4.09
20	–6.04
25	–7.99
30	–9.94
35	–11.84

In addition, the runs are broken down into single readings, and the opacity from the meter is used to group the observations. All observations made at the same opacity level are treated as a single sample, and the accuracy of each is calculated. The results are summarized in Table 33.

The meter values ranged from 3 to 62 percent opacity. For each, a mean is calculated and a standard deviation. A 95-percent confidence interval is calculated using the expression

$$CI_{0.95} : \bar{x} \pm c_{0.05} \frac{s}{\sqrt{n}}$$

TABLE 33. 1-PERCENT DATA SUMMARY

Opacity	$\bar{X}$	s	n	t	$CI_{0.95}$ UL	$CI_{0.95}$ LL	Max	Min
3	3.3	2.0	108	1.96	3.677	2.923	10	0
4	3.2	2.0	100	1.96	3.592	2.808	9	0
5	4.6	3.1	112	1.96	5.174	4.026	16	0
6	9.4	3.7	20	2.093	11.132	7.668	17	4
7	6.0	3.3	28	2.052	7.280	4.720	14	1
8	9.5	2.9	4	3.182	14.114	4.886	13	6
9	9.8	5.0	8	2.365	13.981	5.619	18	5
10	10.4	2.7	8	2.365	12.658	8.142	15	8
11	11.0	2.3	8	2.365	12.923	9.077	15	8
13	14.3	4.0	4	3.182	20.664	7.936	19	10
17	13.3	2.2	4	3.182	16.800	9.800	15	10
19	13.5	5.0	4	3.182	21.455	5.545	19	7
20	13.9	3.8	24	2.069	15.505	12.295	23	8
21	13.3	3.4	44	2.021	14.336	12.264	26	8
22	15.1	4.5	16	2.131	17.497	12.703	24	8
23	14.9	2.7	12	2.201	16.616	13.184	20	12
24	20.0	1.6	4	3.182	22.546	17.454	22	18
25	18.3	4.0	16	2.131	20.431	16.169	28	12
26	19.1	6.2	32	2.042	21.338	16.862	28	5
27	17.3	4.2	20	2.093	19.266	15.334	25	11
28	18.6	4.6	32	2.042	20.260	16.940	25	9
29	20.0	4.2	12	2.201	22.669	17.331	25	12
30	20.3	4.0	16	2.131	22.431	18.169	26	12
31	20.3	6.6	8	2.365	25.819	14.781	28	8
32	21.4	3.9	16	2.131	23.478	19.322	25	16
33	22.2	6.7	20	2.093	25.336	19.064	30	7
34	20.7	6.8	32	2.042	23.155	18.245	34	5
35	20.4	6.1	36	2.031	22.465	18.335	36	5
36	22.8	3.6	8	2.365	25.810	19.790	28	19
37	23.5	4.4	4	3.182	30.500	16.500	28	18
38	27.0	8.8	4	3.182	41.001	12.999	39	21
39	21.3	6.1	8	2.365	26.401	16.199	28	12
40	22.8	1.7	4	3.182	25.505	20.095	25	21
44	32.3	5.4	4	3.182	40.891	23.709	38	25
45	31.0	4.7	4	3.182	38.478	23.522	36	25
46	30.5	4.9	4	3.182	38.296	22.704	37	25
50	32.0	5.4	4	3.182	40.591	23.409	38	25
58	36.3	4.8	4	3.182	43.937	28.663	40	30
62	36.5	6.6	4	3.182	47.001	25.999	44	30

where $c_{0.05}$ is the tabled value, taken either from the standard normal or from Student's t distribution, and s is the calculated standard deviation. The observations are accurate when the "true value" lies within the interval. As can be seen from the table, the 1-percent data are generally accurate up to 20 percent opacity, but beyond 20 percent the accuracy standard is met only once. Also included in the tables are the maximum and minimum observations at each opacity level.

b. Precision

The precision of the 1-percent variation is investigated using an ANOVA, as before. The details are contained in Appendix B.7. The estimated within-observer variance is

$$\hat{\sigma}^2 = 8.29$$

with 21 df. This gives an estimated within-observer standard deviation of

$$\hat{\sigma} = \sqrt{8.29}$$

$$= 2.88 \text{ percent opacity.}$$

This value is somewhat larger than that obtained from the 5-percent variation. To test whether or not these two variance terms are equivalent, an F-ratio is calculated comparing this estimate to the variance estimate from the 5-percent increment data. The statistic is $F_c = 2.32$ with 21 and 90 df. From the tables, the critical value for this test is $F_{0.05}$ (21,90) $\approx$ 1.75. Since the calculated F exceeds the tabled F, the two variances may be said to be different.

The F-ratio for observers is $F_c = 1.84$ with 3 df. This is not significant at the 5-percent level, which indicates that there was no observer effect in these runs. Thus, σ_L^2 is taken to be zero, and

$$\hat{\sigma}_b^2 = 8.29,$$

with 21 df as above.

c. Predicted Results

Using the results obtained above, the range of opacities that would be expected from the use of this variation can be predicted. As before, the range is calculated as

$$\text{expected range: } k + \text{predicted deviation} \pm t_{0.05}\,(21)\,\hat{\sigma}$$

$$: \; k + [1.76 - 0.39k] \pm (2.080)(2.88)$$

$$: \; [0.71k + 1.76] \pm (5.99)$$

Values of this range are listed in Table 34 for selected opacity levels. The data indicate that there is fairly good agreement at the lower opacities, but that a negative bias appears above 25 percent. This is a comparable result to the 5-percent increment data.

The maximum difference that would be expected between observations by two independent observers is estimated using the between-observer standard deviation. As before, the equation is

$$\text{maximum difference} = t_{0.05}\,(21)\sqrt{2\hat{\sigma}_b^2}$$

$$= (2.080)\sqrt{16.58}$$

$$= (2.080)(4.07)$$

$$= 8.47 \text{ percent opacity}$$

As expected, this value is greater than that for the 5-percent increment data. The apparent conclusion is that there was more variability in the 1-percent data than the 5 percent. A contributing factor to this was undoubtedly that the observers had not been trained to read to the nearest percent, but it is impossible to determine how much effect this actually had. From a second ANOVA shown in Appendix B.7, the two groups of observers cannot be shown to be reading different average opacities on these runs.

TABLE 34. PREDICTED DETERMINATIONS, 1-PERCENT DATA

Opacity	Expected Range
5	0*–11.30
10	2.87–14.85
15	6.42–18.40
20	9.97–21.95
25	13.52–25.50
30	17.07–29.05
35	20.62–32.60
40	21.17–36.15

*Truncated to zero.

6. Two Observers

A second variation to be studied involved taking the average of two observers results as opposed to a single observer's average. The 6 possible pairings among the control group are made, and the determinations are shown in Table 35 with the deviations shown in Table 36.

TABLE 35. OPACITY DETERMINATIONS USING PAIRED OBSERVERS

Run	Observer Pair (1,2)	(1,3)	(1,4)	(2,3)	(2,4)	(3,4)	Meter
1	5.1	4.0	5.0	4.1	5.1	4.0	3.1
2	13.6	14.5	9.5	14.5	9.5	10.4	7.5
3	5.0	4.5	5.0	4.5	5.0	4.5	4.2
4	8.7	8.8	7.1	8.5	6.8	6.9	7.2
5	8.6	9.6	7.8	7.8	6.0	7.0	6.7
6	6.6	6.0	5.8	6.0	5.8	5.2	5.0
7	10.6	9.7	8.3	9.3	7.9	7.0	5.4
8	5.9	5.8	5.5	5.9	5.6	5.5	3.4
9	6.3	6.5	6.3	5.2	5.0	5.2	4.2
10	6.4	6.1	5.9	5.7	5.5	5.2	4.8
11	24.0	30.6	20.9	29.6	19.9	26.5	21.2
12	39.9	36.9	40.2	36.0	39.3	36.3	36.0
13	10.5	8.4	8.7	7.7	8.0	5.9	10.3
14	16.1	11.7	13.6	10.8	12.7	8.3	14.2
15	11.8	10.3	11.8	9.3	10.8	9.3	12.2
16	25.8	22.5	25.6	22.3	25.4	22.1	29.6
17	12.9	12.3	10.5	11.8	10.0	9.4	14.6
18	27.2	26.6	27.1	24.6	25.1	24.5	28.3
19	23.4	21.1	22.5	20.7	22.1	19.8	34.6
20	21.2	19.8	17.7	18.0	15.9	14.5	27.2
21	34.2	31.9	33.5	32.3	33.9	31.6	36.9
22	15.0	13.7	12.7	11.3	10.3	9.0	15.2
23	21.8	20.9	21.5	18.9	19.5	18.6	27.8
24	13.8	11.0	9.7	12.2	10.9	8.1	14.4

There are 15 deviations of absolute magnitude greater than 7.5 percent out of the 144 pairings, and also 15 out of the original 96 determinations. Thus, there was a decrease in the frequency of the large deviation. The percentages are 0.10 as opposed to 0.16. In Appendix B.8, it is shown that this is a significant decrease in frequency.

The precision that would be expected from this variation is estimated by taking one-half of the between-observer variance term from the previous analysis. The variance between two observations, then, would be

$$\hat{\sigma}_b^2 = 1/2\,(6.69)$$

$$= 3.35$$

and a standard deviation of

$$\hat{\sigma}_b = 1.83 \text{ percent opacity.}$$

TABLE 36. DEVIATIONS FROM OPACITY, PAIRED OBSERVERS

Run	Observer Pair						Meter
	(1,2)	(1,3)	(1,4)	(2,3)	(2,4)	(3,4)	
1	2.0	0.9	1.9	1.0	2.0	0.9	3.1
2	6.1	7.0	2.0	7.0	2.0	2.9	7.5
3	0.8	0.3	0.8	0.3	0.8	0.3	4.2
4	1.5	1.6	−0.1	1.3	−0.4	−0.3	7.2
5	1.9	2.9	1.1	1.1	−0.7	0.3	6.7
6	1.6	1.0	0.8	1.0	0.8	0.2	5.0
7	5.2	4.3	2.9	3.9	2.5	1.6	5.4
8	2.5	2.4	2.1	2.5	2.2	2.1	3.4
9	2.1	2.3	2.1	1.0	0.8	1.0	4.2
10	1.6	1.3	1.1	0.9	0.7	0.4	4.8
11	2.8	9.4	−0.3	8.4	−1.3	5.3	21.2
12	3.9	0.9	4.2	0.0	3.3	0.3	36.0
13	0.2	−1.9	−1.6	−2.6	−2.3	−4.4	10.3
14	1.9	−2.5	−0.6	−3.4	−1.5	−5.9	14.2
15	−0.4	−1.9	−0.4	−2.9	−1.4	−2.9	12.2
16	−3.8	−7.1	−4.0	−7.3	−4.2	−7.5	29.6
17	−1.7	−2.3	−4.1	−2.8	−4.6	−5.2	14.6
18	−1.1	−1.7	−1.2	−3.7	−3.2	−3.8	28.3
19	−11.2	−13.5	−12.1	−13.9	−12.5	−14.8	34.6
20	−6.0	−7.4	−9.5	−9.2	−11.3	−12.7	27.2
21	−2.7	−5.0	−3.4	−4.6	−3.0	−5.3	36.9
22	−0.2	−1.5	−2.5	−3.9	−4.9	−6.2	15.2
23	−6.0	−6.9	−6.3	−8.9	−8.3	−9.2	27.8
24	−0.6	−3.4	−4.7	−2.2	−3.5	−6.3	14.4

The 95-percent confidence limit on the difference between two determinations made this way would be

$$\text{maximum difference: } t_{0.05}(6)\sqrt{2}\sigma_b$$

$$: (2.447)(1.414)(1.83)$$

$$: 6.33 \text{ percent opacity.}$$

This is a 29-percent reduction from the single observer determination result.

VII. COMPOSITE ESTIMATION

As a final evaluation, it is investigated whether the precision and accuracy from all the tests may be combined to produce statements that are applicable to any site. The accuracy statements are derived from the training generator and Steam Station Test 3 data, where good background was available. The sulfuric acid test data is not used since it applies to a much narrower range of opacities.

A. Accuracy

The slopes of the regression lines for the training generator and steam station test are tested for equality. A test statistic is calculated for comparing the three estimated slopes using their standard deviations and an uncertainty factor. The details are given in Appendix B.9.

The slopes can be said to be estimating the same true value if they differ from each other by no more than 0.12. As can be seen, the difference between the largest and the smallest slope estimate is 0.10, and thus the three values are insignificantly different from one another.

Because of this, a composite line is calculated using all the points from the three sets of data. The equation thus calculated is

$$\text{deviation} = 3.13 - (0.31)\ (\text{percent opacity})$$

for the range from 5 to 35 percent average opacity. The expected deviation calculated using this equation is shown in Table 37 for values in that range.

TABLE 37. PREDICTED DEVIATION FROM METERED OPACITY, COMPOSITE ESTIMATE

Opacity	Expected Deviation
5	1.58
10	0.03
15	−1.52
20	−3.07
25	−4.62
30	−6.17
35	−7.72

B. Precision

The six estimates of the within-observer variance, σ^2, are compared using Bartlett's test[3]. The test statistic is a chi-square, and is calculated to be 13.77 with 5 df, and the significance level for this value is approximately 0.02. This is sufficient to accept the hypothesis that all are estimating the same true variance.

The combined estimate is obtained by "pooling" the variances, as described in Appendix B.9. The pooled variance is

$$\hat{\sigma} = 4.22$$

with 693 df. This gives an estimated within-observer standard deviation of

$$\hat{\sigma} = \sqrt{4.22}$$

$$= 2.05 \text{ percent opacity.}$$

Similarly, the 6 values of σ_L^2 are tested for equality using Bartlett's test. The test statistic is 4.72 with 5 df, which is not significant. Thus, for these components as well, the hypothesis of equality is accepted. The pooled estimate of the observer bias variance is

$$\hat{\sigma}_L^2 = 1.66$$

with 44 df. This gives an estimated observer bias standard deviation of

$$\hat{\sigma}_L = 1.29 \text{ percent opacity.}$$

Combining the above estimates, the between-observer variance is estimated as

$$\hat{\sigma}_b^2 = \hat{\sigma}_L^2 + \hat{\sigma}^2$$

$$= 4.22 + 1.66$$

$$= 5.88.$$

This gives a between-observer standard deviation of

$$\hat{\sigma}_b = \sqrt{5.88}$$

$$= 2.42 \text{ percent opacity.}$$

C. Predicted Results

Using the results of the previous two sections, the expected range and maximum expected difference can be calculated for the composite values. The expected range would be

$$\text{expected range: } k + [3.13 - 0.31k] \pm 1.96\,\hat{\sigma}$$

$$: (0.69k + 3.13) \pm 1.96\,(2.05)$$

$$: (0.69k + 3.13) \pm 4.02$$

for any true opacity k, at the 95-percent confidence level. The expected range is presented for selected values of k in Table 38. As before, the results show accuracy up to about 25 percent opacity, then acquire a low-side bias.

TABLE 38. PREDICTED DETERMINATIONS OF AVERAGE OPACITY, COMPOSITE ESTIMATE

Percent Opacity	Expected Range
5	2.56–10.60
10	6.01–14.05
15	9.46–17.50
20	12.91–20.95
25	16.36–24.40
30	19.81–27.85
35	23.26–31.30

The maximum difference that could be expected between two independent observers is calculated using the between-observer standard deviation. At the 95 percent confidence level, the difference could not be expected to exceed

$$\text{maximum difference: } t_{0.05}\,(44)\sqrt{2\,\hat{\sigma}_b^2}$$

$$: (2.021)\,(1.414)\,(2.42)$$

$$: 6.92 \text{ percent opacity.}$$

The 44 df are due to the observer bias term.

APPENDIX A. EPA METHOD 9 FOR VISUAL DETERMINATION OF OPACITY

Rom, Jerome J., Maintenance, Calibration, and Operation of Isokinetic Source Sampling Equipment, Environmental Protection Agency, Air Pollution Control Office Publication No. APTD-0576.

Shell Development Co. Analytical Department, Determination of Sulfur Dioxide and Sulfur Trioxide in Stack Gases, Emeryville Method Series, 4516/59a.

METHOD 9—VISUAL DETERMINATION OF THE OPACITY OF EMISSIONS FROM STATIONARY SOURCES

1. *Principle and applicability.*

1.1 Principle. The relative opacity of an emission from a stationary source is determined visually by a qualified observer.

1.2 Applicability. This method is applicable for the determination of the relative opacity of visible emissions from stationary sources only when specified by test procedures for determining compliance with the New Source Performance Standards.

2. *Procedure.*

2.1 The qualified observer stands at approximately two stack heights, but not more than a quarter of a mile from the base of the stack with the sun to his back. From a vantage point perpendicular to the plume, the observer studies the point of greatest opacity in the plume. The data required in Figure 9-1 is recorded every 15 to 30 seconds to the nearest 5% opacity. A minimum of 25 readings is taken.

3. *Qualifications.*

3.1 To certify as an observer, a candidate must complete a smokereading course conducted by EPA, or equivalent; in order to certify the candidate must assign opacity readings in 5% increments to 25 different black plumes and 25 different white plumes, with an error not to exceed 15 percent on any one reading and an average error not to exceed 7.5 percent in each category. The smoke generator used to qualify the observers must be equipped with a calibrated smoke indicator or light transmission meter located in the source stack if the smoke generator is to determine the actual opacity of the emissions. All qualified observers must pass this test every 6 months in order to remain certified.

4. *Calculations.*

4.1 Determine the average opacity.

5. *References.*

Air Pollution Control District Rules and Regulations, Los Angeles County Air Pollution Control District, Chapter 2, Schedule 6, Regulation 4, Prohibition, Rule 50, 17 p.

Kudluk, Rudolf, Ringelmann Smoke Chart, U.S. Department of Interior, Bureau of Mines, Information Circular No. 8333, May 1967.

SEC / MIN	0	15	30	45	SEC / MIN	0	15	30	45
0					30				
1					31				
2					32				
3					33				
4					34				
5					35				
6					36				
7					37				
8					38				
9					39				
10					40				
11					41				
12					42				
13					43				
14					44				
15					45				
16					46				
17					47				
18					48				
19					49				
20					50				
21					51				
22					52				
23					53				
24					54				
25					55				
26					56				
27					57				
28					58				
29					59				

Observation data

Plant ________

Stack location ________

Observer ________

Date ________

Time ________

Distance to stack ________

Wind direction ________

Wind speed ________

Sum of numbers recorded ________

Total number of readings ________

$$\text{Opacity} = \frac{\text{Sum of nos. recorded}}{\text{Total no. readings}}$$

Figure 9-1. Field data.

[FR Doc.71-13624 Filed 12-22-71;8:45 am]

A.2. Federal Register, Vol 39, No. 177—Wednesday, September 11, 1974

(e) For the purpose of reports pursuant to § 60.7(c), periods of excess emissions that shall be reported are defined as any two consecutive hourly periods during which the average nitrogen oxides emissions exceed the limits of § 60.72(a).

Subpart H—Standards of Performance for Sulfuric Acid Plants

§ 60.83 [Amended]

11. Paragraph (a)(2) of § 60.83 is amended by deleting the second sentence.

12. Section 60.84 is amended by revising paragraphs (a), (b), (c), (d) and (e) to read as follows:

§ 60.84 Emission monitoring.

(a) A continuous monitoring system for the measurement of sulfur dioxide shall be installed, calibrated, maintained, and operated in any affected facility by the owner or operator.

(b) The owner or operator shall establish a conversion factor for the purpose of converting monitoring data into units of the applicable standard (kg/metric ton). The conversion factor shall be established by measuring emissions with the continuous monitoring system in accordance with § 60.13(d) at the same time as emissions are determined for the applicable reference method tests (§60.85). Using emission data from the performance test and only that portion of the continuous monitoring emission data that represents emission measurements concurrent with the reference method test periods, the conversion factor shall be determined by dividing the performance test data integrated averages by the continuous emission monitoring test data averages to obtain a ratio expressed in units of the applicable standard to units of the monitoring data, i.e., kg/metric ton per ppm (lb/short ton per ppm).

(c) The owner or operator shall record the daily production rates and hours of operation.

(d) Paragraph (m) of § 60.13 does not apply to this subpart.

* * * * *

(e) For the purpose of reports pursuant to § 60.7(c), periods of excess emissions that shall be reported are defined as any two consecutive hourly periods during which the average sulfur dioxide emissions exceed the limits of § 60.82.

Subpart I—Standards of Performance for Asphalt Concrete Plants

§ 60.92 [Amended]

13. Paragraph (b) of § 60.92 is amended by deleting the second sentence.

Subpart J—Standards of Peformance for Petroleum Refineries

§ 60.102 [Amended]

14. Paragraph (a)(2) of § 60.102 is amended by deleting the second sentence.

15. Section 60.105 is amended by revising paragraphs (a) and (b) to read as follows:

§ 60.105 Emission monitoring.

(a) Continuous monitoring systems shall be installed, calibrated, maintained, and operated in any affected facility by the owner or operator as follows:

(1) A continuous monitoring system for the measurement of the opacity of emissions discharged into the atmosphere from the fluid catalytic cracking unit catalyst regenerator.

(2) A continuous monitoring system for the measurement of carbon monoxide in gases discharged into the atmosphere from fluid catalytic cracking unit catalyst regenerators, except where the requirements of paragraph (a)(3) of this section are met.

(3) Continuous monitoring systems for the measurement of firebox temperature and oxygen in exhaust gases from any incinerator-waste heat boiler which combusts the exhaust gases from a fluid catalytic cracking unit catalyst regenerator, except where the requirements of paragraph (a)(2) of this section are met.

(4) A continuous monitoring system for the measurement of hydrogen sulfide in fuel gases burned in any fuel gas combustion device, except where the requirements of paragraph (a)(5) of this section are met. Fuel gas combustion devices having a common source of fuel gas may be monitored at one location if sampling at this location produces results representative of the hydrogen sulfide concentration in the fuel gas burned.

(5) A continuous monitoring system for the measurement of sulfur dioxide in the gases discharged into the atmosphere from the combustion of fuel gases except where the requirements of paragraph (a)(4) of this section are met.

(b) The owner or operator of an affected facility required to install a continuous monitoring system under paragraph (a)(5) of this section shall establish daily a conversion factor (mg H_2S/dscm of fuel gas per ppm SO_2 in flue gas) for the purpose of converting monitoring data into units of the applicable standard. The conversion factor shall be established by measuring ppm SO_2 with the continuous monitoring system in accordance with paragraph (a)(5) of this section, conducting analyses daily of any fuel gas combusted using A.S.T.M. methods D1071–55 (sampling) and D1946–67 (analysis), and computing the factor as follows:

$$CF=(0.0162)\left(\sum_{i=1}^{n} X_i Y_i\right),$$

where:
- CF = daily conversion factor (mg H_2S/dscm per ppm)
- X_i = stoichiometric constant, and
- Y_i = volumetric percent of fuel gas component, dry basis
- n = number of components.

TABLE OF VALUES FOR X_i

Component:	X_i
H_2	1.88
H_2S	6.64
CO	2.88
CH_4	8.52
C_2H_6	15.16
C_3H_8	21.8
C_4H_{10}	28.4
C_5H_{12}	35.1
C_6H_{14}	41.7
C_2H_4	13.3
C_3H_6	19.9
C_4H_4	18.0
C_4H_8	26.6
C_6H_6	34.2
C_9H_{12}	39.8
Inerts	1.0

(1) For facilities which operate at levels greater than 5% oxygen in the flue gases, the oxygen content of the flue gases shall be determined at least monthly by sampling at the same point in the duct specified by § 60.106(d) and by analyzing the sample for oxygen by Orsat analysis or by use of the oxygen continuous monitoring instrument under § 60.105(a)(3). The results shall be used to adjust the daily conversion factors monthly (Orsat analysis) or daily (continuous monitor) as follows:

$$(CF)_{adjusted}=\left(\frac{CF}{1.17}\right)\left(\frac{20.90}{20.9-\%O_2}\right).$$

where:
- CF = daily conversion factor (mg H_2S/dscm per ppm), and
- $\%O_2$ = percent oxygen by volume (expressed as percent), dry basis.

* * * * *

Subpart L—Standards of Performance for Secondary Lead Smelters

§ 60.122 [Amended]

16. Section 60.122 is amended by deleting paragraph (c).

Subpart M—Standards of Performance for Secondary Brass and Bronze Ingot Production Plants

§ 60.132 [Amended]

17. Section 60.132 is amended by deleting paragraph (c).

Subpart O—Standards of Performance for Sewage Treatment Plants

§ 60.152 [Amended]

18. Paragraph (a)(2) of § 60.152 is amended by deleting the second sentence.

19. Appendix A is amended by revising Method 9 as follows:

METHOD 9—VISUAL DETERMINATION OF THE OPACITY OF EMISSIONS FOR STATIONARY SOURCES

1. *Principle and applicability.*

1.1 Principle. The opacity of emissions from stationary sources is determined visually by a qualified observer.

1.2 Applicability. This method is applicable for the determination of the opacity of emissions from stationary sources pursuant to § 60.11(b) and for qualifying observers for visually determining opacity of emissions.

2. *Procedures.* The observer qualified in accordance with paragraph 3 of this method shall use the following procedures for visually determining the opacity of emissions:

2.1 Position. The qualified observer shall stand at a distance sufficient to provide a clear view of the emissions with the sun oriented in the quadrant to his back. Consistent with maintaining the above requirement, the observer shall, as much as possible, make his observations from a position such that his line of vision is approximately perpendicular to the plume direction, and when observing opacity of emissions from rectangular outlets (e.g. roof monitors, open baghouses, noncircular stacks), approximately perpendicular to the longer axis of the outlet.

2.2 Field records. The observer shall record the name of the plant, emission location, type facility, observer's name and affiliation, and the date on a field data sheet (Figure 9-1). The time, estimated distance to the emission location, approximate wind direction, estimated wind speed, description of the sky condition (presence and color of clouds), and plume background are recorded on a field data sheet at the time opacity readings are initia ed and completed.

2.3 Observations. Opacity observations shall be made at the point of greatest opacity in that portion of the plume where condensed water vapor is not present. The observer shall not look continuously at the plume, but instead shall observe the plume momentarily at 15-second intervals.

2.3.1 Attached steam plumes. When condensed water vapor is present within the plume as it emerges from the emission outlet, opacity observations shall be made beyond the point in the plume at which condensed water vapor is no longer visible. The observer shall record the approximate distance from the emission outlet to the point at which the observations are made.

2.3.2 Detached steam plume. When water vapor in the plume condenses at a distinct distance from the emission outlet, the opacity of emissions may be evaluated at the emission outlet prior to the condensation of water vapor and the formation of the steam plume.

2.4 Recording observations. Opacity observations shall be recorded to the nearest 5 percent at 15-second intervals on an observational record sheet (Figure 9-2). Each momentary observation recorded shall represent the opacity of emissions for that 15-second time period.

2.5 Data reduction. Determine the opacity standard applicable to the affected facility being observed and compute minutes of noncompliance using the appropriate procedure as follows:

2.5.1 For all opacity standards except average opacity, mark on the observation record sheet (Figure 9-2) all observations which exceed the opacity level specified by the standard. Compute the aggregate number of minutes during which the opacity of emissions was observed to exceed the applicable standard. Exclude from this computation the number of minutes of operation exempted by the applicable opacity standard or by § 60.11(c) (if any). Record the minutes of noncompliance on Figure 9-1.

2.5.2 For average opacity standards, sum on the observation record sheet (Figure 9-2) all observations for any specific time period (use time periods specified by the standard when applicable), and compute the arithmetic mean average by dividing by the total number of observations for the time period evaluated. Exclude from this computation any minutes of operation exempted by the applicable opacity standard or § 60.11(c) (if any). Record the minutes of each time period of noncompliance on Figure 9-1.

3. *Qualifications and testing.*

3.1 Certification requirements. To receive certification as a qualified observer, a candidate must be tested and demonstrate the ability to assign opacity readings in 5-percent increments to 25 different black plumes and 25 different white plumes, with an error not to exceed 15-percent opacity on any one reading and an average error not to exceed 7.5-percent opacity in each category. Candidates shall be tested according to the procedures described in paragraph 3.2. Smoke generators used pursuant to paragraph 3.2 shall be equipped with a smoke meter which meets the requirements of paragraph 3.3.

The certification shall be valid for a period of 6 months, at which time the qualification procedure must be repeated by any observer in order to retain certification.

3.2 Certification procedure. The smoke test consists of showing the candidate a complete run of 50 plumes—25 black plumes and 25 white plumes—generated by a smoke generator. Plumes within each set of 25 black and 25 white runs shall be presented in random order. The candidate assigns an opacity value to each plume and records his observation on a suitable form. At the completion of each run of 50 readings, the score of the trainee is determined. If a trainee fails to qualify, the complete run of 50 readings must be repeated in any retest. The smoke test may be administered as part of a smoke school or training program, and may be preceded by training or familiarization runs of the smoke generator during which candidates are shown black and white plumes of known opacity.

3.3 Smoke generator specifications. Any smoke generator used for the purposes of paragraph 3.2 shall be equipped with a smoke meter installed to measure opacity across the diameter of the smoke generator stack. The smoke meter output shall display in-stack opacity based upon a pathlength equal to the stack exit diameter, on a full 0 to 100 percent chart recorder scale. The smoke meter optical design and performance shall meet the specifications shown in Table 9-1. The smoke meter shall be calibrated as prescribed in paragraph 3.3.1 prior to the conduct of each smoke reading test. At the completion of each test, the zero and span drift shall be checked and if the drift exceeds ±1 percent opacity, the condition shall be corrected prior to conducting any subsequent test runs. The smoke meter shall be demonstrated, at the time of installation, to meet the specifications listed in Table 9-1. This demonstration shall be repeated following any subsequent repair or replacement of the photocell or associated electronic circuitry including the chart recorder or output meter, or every 6 months, whichever occurs first.

TABLE 9-1 SMOKE METER DESIGN AND PERFORMANCE SPECIFICATIONS

Parameter:	*Specification*
a. Light source	Incandescent lamp operated at nominal rated voltage.
b. Spectral response of photocell.	Photopic (Daylight spectral response of the human eye: reference 4.3).
c. Angle of view	15° maximum total angle.
d. Angle of projection.	15° maximum total angle.
e. Calibration error	±3% opacity maximum.
f. Zero and span drift.	±1% opacity, 30 minutes.
g. Response time	≤5 seconds.

3.3.1 Calibration. The smoke meter is calibrated after allowing a minimum of 30 minutes warmup by alternately producing simulated opacity of 0 percent and 100 percent. When stable response at 0 percent or 100 percent is noted, the smoke meter is adjusted to produce an output of 0 percent or 100 percent, as appropriate. This calibration shall be repeated until stable 0 percent and 100 percent readings are produced without adjustment. Simulated 0 percent and 100 percent opacity values may be produced by alternately switching the power to the light source on and off while the smoke generator is not producing smoke.

3.3.2 Smoke meter evaluation. The smoke meter design and performance are to be evaluated as follows:

3.3.2.1 Light source. Verify from manufacturer's data and from voltage measurements made at the lamp, as installed, that the lamp is operated within ±5 percent of the nominal rated voltage.

3.3.2.2 Spectral response of photocell. Verify from manufacturer's data that the photocell has a photopic response; i.e., the spectral sensitivity of the cell shall closely approximate the standard spectral-luminosity curve for photopic vision which is referenced in (b) of Table 9-1.

3.3.2.3 Angle of view. Check construction geometry to ensure that the total angle of view of the smoke plume, as seen by the photocell, does not exceed 15°. The total angle of view may be calculated from: $\theta = 2 \tan^{-1} d/2L$, where θ = total angle of view, d = the sum of the photocell diameter + the diameter of the limiting aperture; and L = the distance from the photocell to the limiting aperture. The limiting aperture is the point in the path between the photocell and the smoke plume where the angle of view is most restricted. In smoke generator smoke meters this is normally an orifice plate.

3.3.2.4 Angle of projection. Check construction geometry to ensure that the total angle of projection of the lamp on the smoke plume does not exceed 15°. The total angle of projection may be calculated from: $\theta = 2 \tan^{-1} d/2L$, where θ = total range of projection; d = the sum of the length of the lamp filament + the diameter of the limiting aperture; and L = the distance from the lamp to the limiting aperture.

3.3.2.5 Calibration error. Using neutral-density filters of known opacity, check the error between the actual response and the theoretical linear response of the smoke meter. This check is accomplished by first calibrating the smoke meter according to 3.3.1 and then inserting a series of three neutral-density filters of nominal opacity of 20, 50, and 75 percent in the smoke meter pathlength. Filters calibrated within ±2 percent shall be used. Care should be taken when inserting the filters to prevent stray light from affecting the meter. Make a total of five nonconsecutive readings for each filter. The maximum error on any one reading shall be 3 percent opacity.

3.3.2.6 Zero and span drift. Determine the zero and span drift by calibrating and operating the smoke generator in a normal manner over a 1-hour period. The drift is measured by checking the zero and span at the end of this period.

3.3.2.7 Response time. Determine the response time by producing a series of five simulated 0 percent and 100 percent opacity values and observing the time required to reach stable response. Opacity values of 0 percent and 100 percent may be simulated by alternately switching the power to the light source off and on while the smoke generator is not operating.

4. *References.*

4.1 *Air Pollution Control District Rules and Regulations*, Los Angeles County Air Pollution Control District, Regulations IV, Prohibitions, Rule 50.

4.2 Weisburd, Melvin I., *Field Operations and Enforcement Manual for Air*, U.S. Environmental Protection Agency, Research Triangle Park, N.C., APTD-1100, August 1972, pp. 4.1-4.36.

4.3 Condon, E. U., and Odishaw, H., *Handbook of Physics*, McGraw-Hill Co., New York, N.Y., 1958, Table 3.1, p. 6-52.

Figure 9-1 Page____of____

RECORD OF VISUAL DETERMINATION OF OPACITY

COMPANY________________ HOURS OF OBSERVATION________________
LOCATION________________ OBSERVER________________
TEST NUMBER________________ OBSERVER CERTIFICATION DATE________________
DATE________________ OBSERVER AFFILIATION________________
TYPE FACILITY________________ POINT OF EMISSIONS________________
CONTROL DEVICE________________ HEIGHT OF DISCHARGE POINT________________

Record the following information prior to and upon completion of observations at each source. If observations are made over an extended period of time, additional recordings should be made as applicable.

CLOCK TIME INITIAL __:__a.m. __:__p.m. FINAL __:__ a.m. __:__ p.m.

OBSERVER LOCATION
Distance to Discharge________________
Direction from Discharge________________
Height of Observation Point________________

BACKGROUND DESCRIPTION________________

WEATHER CONDITIONS
Wind Direction________________
Wind Speed________________
Ambient Temperature________________

SKY CONDITIONS (clear, overcast, %clouds, etc.)________________

PLUME DESCRIPTION
Color________________
Distance Visible________________

MINUTES OF NONCOMPLIANCE________________

FIGURE 9-2 OBSERVATION RECORD PAGE ___ OF ___

COMPANY________________ OBSERVER________________
LOCATION________________ TYPE FACILITY________________
TEST NUMBER________________ POINT OF EMISSIONS________________
DATE________________

Hr.	Min.	Seconds				STEAM PLUME (check if applicable)		COMMENTS
		0	15	30	45	Attached	Detached	
	0							
	1							
	2							
	3							
	4							
	5							
	6							
	7							
	8							
	9							
	10							
	11							
	12							
	13							
	14							
	15							
	16							
	17							
	18							
	19							
	20							
	21							
	22							
	23							
	24							
	25							
	26							
	27							
	28							
	29							

FIGURE 9-2 OBSERVATION RECORD (Cont.) PAGE ___ OF ___

COMPANY ______________ OBSERVER ______________
LOCATION ______________ TYPE FACILITY ______________
TEST NUMBER ______________ POINT OF EMISSIONS ______________
DATE ______________

		Seconds				STEAM PLUME (check if applicable)		
Hr.	Min.	0	15	30	45	Attached	Detached	COMMENTS
	30							
	31							
	32							
	33							
	34							
	35							
	36							
	37							
	38							
	39							
	40							
	41							
	42							
	43							
	44							
	45							
	46							
	47							
	48							
	49							
	50							
	51							
	52							
	53							
	54							
	55							
	56							
	57							
	58							
	59							

20. Appendix B is added as follows:

APPENDIX B—PERFORMANCE SPECIFICATIONS

PERFORMANCE SPECIFICATION 1—PERFORMANCE SPECIFICATIONS AND SPECIFICATION TEST PROCEDURES FOR TRANSMISSOMETER SYSTEMS FOR CONTINUOUS MEASUREMENT OF THE OPACITY OF STACK EFFLUENTS

1. *Principle and Applicability.*

1.1 Principle. Transmissometry is a direct measurement of attenuation of visible radiation (opacity) by particulate matter in a stack effluent. Light from a lamp is projected across the stack of a pollution source to a light sensor. The light is attenuated due to absorption and scatter by the particulate matter in the effluent. The percentage of light attenuated is defined as the opacity of the emission. Transparent stack emissions that do not attenuate light will have a transmittance of 100 or an opacity of 0. Opaque stack emissions that attenuate all of the light will have a transmittance of 0 or an opacity of 100 percent.

1.2 Applicability. This method is applicable to the instrument systems specified in the subparts for continuously monitoring the opacity of emissions. Specifications for continuous measurement of visible emissions are given in terms of design, performance, and installation parameters. Performance specifications and test procedures are given for transmissometer systems to test their capability before approving the systems installed by an affected facility.

2. *Apparatus.*

2.1 Neutral Density Filters. Filters with neutral spectral characteristics and known optical densities to visible light. One each low, mid, and high range filters, 5- or 6-inch square or 6-inch diameter, with nominal optical densities of 0.1, 0.2, and 0.3 (20, 37, and 50 percent opacity) are required. Calibrated filters with accuracies certified by the manufacturer to within 3 percent shall be used. It is recommended that filter calibrations be checked with a well-collimated photopic transmissometer of known linearity prior to use.

2.2 Chart Recorder. Analog chart recorder with input voltage range and performance characteristics compatible with the measurement system output.

2.3 Opacity Measurement System. An in-stack transmissometer (folded or single path) with the optical design specifications designated below, associated control units and apparatus to keep optical surfaces clean.

3. *Definitions.*

3.1 Measurement System. The total equipment required for the continuous determination of a pollutant concentration in a source effluent. The system consists of three major subsystems:

3.1.1 Sampling Interface—That portion of the measurement system that performs one or more of the following operations: delineation, acquisition, transportation, and conditioning of a sample of the source effluent, or protection of the analyzer from the effluent.

3.1.2 Analyzer—That portion of the system which senses the pollutant and generates a signal output that is a function of the pollutant concentration.

3.1.3 Data Recorder—That portion of the measurement system that processes the analyzer output and provides a permanent record of the output signal in terms of pollutant concentration.

3.2 Span. The value of opacity at which the measurement system is set to produce the maximum data display output. For the purpose of this method, the span shall be set at an opacity of 50 percent for a pathlength equal to the stack exit diameter of the source.

3.3 Calibration Error. The difference between the opacity reading indicated by the measurement system and the known values of a series of test standards. For this method the test standards are a series of calibrated neutral density filters.

3.4 Zero Drift. The change in measurement system output over a stated period of time of normal continuous operation when the pollutant concentration at the time of the measurements is zero.

3.5 Calibration Drift. The change in measurement system output over a stated period of time of normal continuous operation when the pollutant concentration at the time of the measurements is the same known upscale value.

3.6 System Response. The time interval from a step change in opacity in the stack at the input to the measurement system to the time at which 95 percent of the corresponding final value is reached as displayed on the measurement system data presentation device.

3.7 Operational Test Period. A minimum period of time over which a measurement system is expected to operate within certain performance specifications without unscheduled maintenance, repair, or adjustment.

3.8 Transmittance. The fraction of incident light that is transmitted through an optical medium of interest.

3.9 Opacity. The fraction of incident light that is attenuated by an optical medium of interest. Opacity (O) and transmittance (T) are related as follows:

$$O = 1 - T$$

3.10 Optical Density. A logarithmic measure of the amount of light that is attenuated by an optical medium of interest. Optical density (D) is related to the transmittance and opacity as follows:

$$D = -\log_{10} T$$
$$D = -\log_{10} (1-O)$$

3.11 Mean Spectral Response. The wavelength which bisects the total area under the curve obtained pursuant to paragraph 9.2.1.

3.12 Angle of View. The maximum (total) angle of radiation that is seen by the photodetector assembly of an optical transmissometer.

3.13 Angle of Projection. The maximum (total) angle of radiation that is projected by the lamp assembly of an optical transmissometer.

3.14 Pathlength. The depth of effluent in the light beam between the receiver and the transmitter of the single-pass transmissometer, or the depth of effluent between the transceiver and reflector of a double-pass transmissometer.

4. *Installation Specification.*

4.1 Location. The transmissometer must be located across a section of duct or stack that will provide a particulate matter flow through the optical volume of the trans-

APPENDIX B. STATISTICAL METHODS

APPENDIX B. STATISTICAL METHODS

B.1 Linear Regression and Test for Significance

In order to obtain the best estimate of the average deviation from the metered opacity, the relationship between the deviation and the meter average is investigated. To investigate this, the method of least-squares is used to give the equation of the best straight line that fits the data obtained. The line is defined by its slope and intercept, estimated by:

$$b = \frac{\Sigma X_i Y_i / n - \overline{xy}}{\Sigma X_i^2 / n - \overline{x}^2}$$

and

$$a = \overline{y} - b\overline{x}$$

where

X_i – the i^{th} value of the independent variable

Y_i – the i^{th} value of the dependent variable

$\overline{x}, \overline{y}$ – the sample means of X and Y, respectively

n – the number of paired observations.

To determine if the line is a good estimate of the relationship between the two variables, the slope of the line is tested. A linear relationship exists if the slope is non-zero. To test this, a statistic is used that is based on Student's t distribution. The statistic is

$$t_c = \frac{b}{s_b}$$

where

t_c – the calculated value

and

s_b – the standard deviation of the slope, b. This statistic has $n - 2$ df associated with it for n pairs of observations.

The standard deviation of the slope, s_b, is calculated by first calculating the standard deviation around the regression line, $s_{y.x}$. The equation for the variance, $S^2{}_{y.x}$ is:

$$\begin{aligned} s^2{}_{y \cdot x} &= \frac{1}{n-2} \sum (Y_i - \hat{Y}_i)^2 \\ &= \frac{1}{n-2} \sum [Y_i - (a + bX_i)]^2 \\ &= \frac{1}{n-2} [\Sigma Y_i^2 - a \, \Sigma Y_i - b \, \Sigma X_i Y_i] \end{aligned}$$

The variance of the slope, b, is then given by

$$s_b^2 = \frac{s^2_{y \cdot x}}{[\Sigma X_i^2 - n\bar{x}^2]}$$

Taking the ratio of the slope to its standard deviation, gives the desired statistic. To determine if this is significant, the value is compared to a Student's t distribution with $n - 2$ df.

B.2 Precision Estimation for Training Generator Test

1. White Smoke

An analysis of variance was performed on the average opacities determined by the 8 collaborators who completed all 20 runs. The model is a random effects model represented by

$$Y_{ij} = \mu + \rho_i + \gamma_j + \epsilon_{ij}$$

where

Y_{ij} – the determination by observer j in run i

μ – overall mean

ρ_i – the effect of run i

γ_j – the effect of collaborator j

and

ϵ_{ij} – the random error associated with Y_{ij}.

This is an additive or no-interaction model. This means that the collaborator and run effects are assumed to be independent of each other, or that an observer reads consistently low or high with respect to the other observers regardless of the opacity level.

TABLE B1. ANALYSIS OF VARIANCE FROM TRAINING GENERATOR TEST, WHITE SMOKE

Source	df	SS	MS	F	EMS
Runs	19	5545.06	–*	–*	–*
Observers	7	165.04	23.58	4.16†	$\sigma^2 + 20\,\sigma_L^2$
Error	133	754.51	5.67		σ^2
Total	159	6464.61			

*Not of interest.
†Significant. $F_c > F_{0.05}(7,133) \approx 2.09$.

The ANOVA results are shown in Table B1. The expected mean squares indicate the proper F-ratios and the means of obtaining estimates of the desired precision components. The EMS of the error term is σ^2, the within-observer variance. Thus, the error mean square is the estimated within-observer variance, and this gives

$$\hat{\sigma}^2 = 5.67$$

with 133 df. This gives an estimated within-observer standard deviation of

$$\hat{\sigma} = \sqrt{5.67}$$

$$= 2.38 \text{ percent opacity.}$$

The significance of the observer term is determined by comparing the observer mean square to the error mean square. The F-ratio is $F_c = 4.16$ which exceeds the critical value, $F_{0.05}$ (7, 133) $\approx$ 2.09. Thus there are differences among the observers, and an observer bias variance can be estimated .

The EMS of the observer term is $\sigma^2 + 20\,\sigma_L^2$, so that the observer bias variance estimate is

$$\hat{\sigma}_L^2 = \frac{MS_L - \hat{\sigma}^2}{20}$$

$$= \frac{23.58 - 5.67}{20}$$

$$= \frac{17.91}{20}$$

$$= 0.90$$

with 7 df. The estimated observer bias standard deviation is

$$\sigma_L^2 = \sqrt{0.90}$$

$$= 0.95 \text{ percent opacity.}$$

2. Black Smoke

The ANOVA table for the black smoke determinations is in Table B2. The model is identical to that for the white smoke.

TABLE B2. ANALYSIS OF VARIANCE FROM TRAINING GENERATOR TEST, BLACK SMOKE

Source	df	SS	MS	F	EMS
Runs	15	3963.43	–*	–*	–*
Observers	7	135.58	19.37	5.73†	$\sigma^2 + 16\,\sigma_L^2$
Error	105	354.81	3.38		σ^2
Total	127	4453.82			

*Not of interest.
†Significant. $F_{0.05}(7,105) \approx 2.12$.

The expected value of the error mean square is σ^2, the within-observer variance. Thus, the estimated value is

$$\hat{\sigma}^2 = 3.38,$$

the error mean square. There are 105 df associated with this estimate. The within-observer standard deviation is

$$\hat{\sigma} = \sqrt{3.38}$$

$$= 1.84 \text{ percent opacity.}$$

The ratio of the observer mean square to the error mean square tests for differences among the observers. The calculated F-statistic is $F_c = 5.73$, which exceeds the critical value of 2.13, approximately, with 7 and 105 df. Thus, there can be said to be differences among the observers.

The observer bias variance is estimated as

$$\hat{\sigma}_L^2 = \frac{MS_L - \hat{\sigma}^2}{16}$$

as indicated by the EMS for observers. This gives

$$\hat{\sigma}_L^2 = \frac{19.37 - 3.38}{16}$$

$$= \frac{15.99}{16}$$

$$= 1.00$$

with 7 df. Thus the observer bias standard deviation is

$$\hat{\sigma}_L = 1.00 \text{ percent opacity.}$$

B.3 Precision Estimation for Sulfuric Acid Plant and Tests of Hypotheses

The precision estimates for use of the method at a sulfuric acid plant are obtained from an analysis of variance. The model for the data is a random effects model,

$$Y_{ij} = \mu + \gamma_i + \rho_j + \epsilon_{ij}$$

where

Y_{ij} – the observation by observer i in run j

μ – the overall mean

γ_i – the effect due to collaborator i

ρ_j – the effect due to run j

and

ϵ_{ij} – the random error associated with Y_{ij}.

The model assumes that there is no interaction between the run and the collaborator. This is equivalent to saying that an observer's tendency to read higher or lower than other observers is independent of the true opacity level. The error term, then, is taken from the interaction between the observers and the runs.

The ANOVA is summarized in Table B3. There are 29 df from the 30 runs and 8 df for the observer term. This leaves a remainder of 232 for error. The expected mean square (EMS) column gives the basis for forming F-ratios to test for significance and for estimating the individual variance components.

The EMS of the error term is σ^2, the within-observer variance term. Thus, the estimated within-observer variance is

$$\hat{\sigma}^2 = 4.51,$$

the error mean square. This gives an estimated standard deviation of

$$\hat{\sigma} = \sqrt{4.51}$$

$$= 2.12 \text{ percent opacity}$$

TABLE B3. ANALYSIS OF VARIANCE FROM SULFURIC ACID PLANT TEST

Source	df	SS	MS	F	EMS
Runs	29	2636.68	90.92	–*	–*
Observers	8	259.06	32.38	7.18†	$\sigma^2 + 30\sigma_L^2$
Error	232	1046.56	4.51		σ^2
Total	269	3942.30			

*Not of interest.
†Significant. $F_{0.05}$ (8,232) ≈ 2.00.

The EMS of the observer term is $\sigma^2 + 30\sigma_L^2$. The factor 30 results from the fact that each observer mean is the average over the 30 runs. The ratio of the observer mean square to the error mean square provides the proper ratio for determining if there is a significant observer term The value of F_c is 7.18, which exceeds the tabled value for the F distribution at the 95-percent confidence level. The tabled value is $F_{0.05}$ (8,232) ≈ 2.00 and, thus, there is a significant observer effect.

Solving the EMS for σ_L^2 gives

$$\sigma_L^2 = \frac{\text{EMS}_{\text{Observer}} - \sigma^2}{30}$$

so that an estimate is given as

$$\hat{\sigma}_L^2 = \frac{\text{MS}_{\text{Observer}} - \hat{\sigma}^2}{30}$$

$$= \frac{32.38 - 4.51}{30}$$

$$= \frac{27.87}{30}$$

$$= 0.93$$

with 8 df. The estimated observer bias standard deviation, then, is

$$\hat{\sigma}_L = \sqrt{0.93}$$

$$= 0.96 \text{ percent opacity.}$$

Combining these two estimates gives

$$\hat{\sigma}_b^2 = \hat{\sigma}_L^2 + \hat{\sigma}^2$$

$$= 0.93 + 4.51$$

$$= 5.44$$

and

$$\hat{\sigma}_b = \sqrt{5.44}$$

$$= 2.33 \text{ percent opacity.}$$

The test for significance of σ_L^2 is a test for the equality among the observers of the mean opacity determined. A significant value of σ_L^2 is equivalent to saying that not all the means are equal. The next problem is to investigate where the differences lie.

For this test, the only comparison of interest concerns the difference between the observers from the enforcement agency and those from the private sector. The enforcement personnel were observers 4, 5, 6 and 8. The means for the 9 collaborators who participated in all 30 runs are shown in Table 16.

To determine if there is a significant difference between the two groups, a contrast among the means is tested. A contrast is of the form

$$\sum c_i \mu_i = 0$$

where

μ_i — the i^{th} true mean

c_i — the coefficient of the i^{th} mean

and

$$\sum c_i = 0.$$

The null hypothesis to be tested is that there is no difference between the two groups of observers. The contrast that tests this hypothesis is written as

$$H_o\text{: } 5(\mu_4 + \mu_5 + \mu_6 + \mu_8) - 4(\mu_1 + \mu_2 + \mu_3 + \mu_7 + \mu_9) = 0$$

where 5 and 4 are chosen so that the sum of the coefficients is 0.

The test statistic follows Student's t distribution and is defined by

$$t_c = \frac{\Sigma c_i \bar{x}_i}{\hat{\sigma}\sqrt{\frac{\Sigma c_i^2}{n}}}$$

where

$\hat{\sigma}$ — the standard deviation for error

$\bar{x}_i$ — the i^{th} sample mean

and

n — the number of determinations used to obtain $\bar{x}_i$

The degrees of freedom for this statistic are the df associated with the estimated variance.

For this test,

$$\Sigma c_i \bar{x}_i = 5(6.47 + \ldots + 5.40) - 4(5.00 + \ldots + 5.73)$$

$$= 29.54.$$

$$\hat{\sigma} = 2.12, \text{ from Table B3}$$

$$\Sigma c_i^2 = 4(5^2) + 5(4^2) = 180$$

$$n = 30$$

so that

$$t_c = \frac{29.54}{2.12\sqrt{\frac{180}{30}}}$$

$$= \frac{29.54}{2.12\sqrt{6}}$$

$$= \frac{29.54}{5.19}$$

$$= 5.69.$$

A t value at the 0.05 level with 232 df is essentially a standard normal value or 1.96. Since t_c exceeds 1.96, we reject the null hypothesis, and conclude that the observers from the enforcement agency read a higher opacity than did the observers from the private sector.

B.4 Analysis of Variance and Tests of Hypotheses from Riverbend Steam Station Test 1

The determinations from the first test are used to obtain precision estimates for within- and between-observer and observer bias components. The determinations used are taken from runs 1, 2, 3, 7, 8 and 9. Runs 4, 5, and 6 are eliminated since the experimental group was reading in 1-percent increments. Run 10 was eliminated since there were no determinations from 6 of the 10 observers, including the entire control group. This was due to the fact that sky conditions and interference from another plume at times prevented the observer from making a reading.

The model for the ANOVA is represented by

$$Y_{ijk} = \mu + \beta_i + \gamma_{j/i} + \rho_k + \epsilon_{ijk}$$

a completely random model, where

Y_{ijk} – the k^{th} observation by collaborator j in group i

μ – the overall mean

β_i – the effect due to group i

$\gamma_{j/i}$ – the effect due to observer j, nested within group i

ρ_k – the effect due to run k

and

ϵ_{ijk} – the random error associated with Y_{ijk}

The observers are treated as a nested factor since the observers differ from one group to another.

TABLE B4. ANALYSIS OF VARIANCE FROM STEAM STATION TEST 1

Source	df	SS	MS	F	EMS
Groups	1	124.42	124.42	8.36*	–†
Runs	5	662.03	–†	–†	–†
Observers/Group	8	119.05	14.88	4.50‡	$\sigma^2 + 6\sigma_L^2$
Error	45	149.10	3.31		
Total	59	1054.60			

*Significant. $F_{0.05}$ (1,8) = 5.32.
†Not of interest.
‡Significant. $F_{0.05}$ (8,45) ≈ 4.08.

The ANOVA table is shown in Table B4. The expected mean squares (EMS) determine the proper F-ratio for each factor. The group mean square is divided by the observer mean square to test for a difference between groups. The F-ratio is $F_c = 8.36$ with 1 and 8 degrees of freedom. The critical value from the F-distribution is $F_{0.05}\,(1,8) = 5.32$, so that the group term is significant. The means for the two groups are 8.20 and 5.32 for the control and experimental, respectively, so that the control group was reading significantly higher opacities during the test. Since the determinations were generally biased low, however, this indicates that the control group was able to judge the opacity more accurately than the experimental.

The EMS of the error term is σ^2, the within-observer component. This gives an estimated within-observer variance of

$$\hat{\sigma}^2 = 3.31$$

with 45 df. The estimated within-laboratory standard deviation, then is

$$\hat{\sigma} = \sqrt{3.31}$$

$$= 1.82 \text{ percent opacity}$$

To determine if there is a significant observer effect, the observer mean square is compared to the error mean square. The ratio is $F_c = 4.50$. The critical value for this test is $F_{0.05}\,(8,45) \approx 4.08$, and the observer bias term, σ_L^2, is determined to be greater than zero.

The observer bias is estimated from the observer group mean square. The EMS is $\sigma^2 + \sigma_L^2$, so the estimated variance is

$$\hat{\sigma}_L^2 = \frac{MS_{\text{Observer}} - \hat{\sigma}^2}{6}$$

$$= \frac{14.88 - 3.31}{6}$$

$$= 1.93$$

with 8 df. The df are calculated as (5-1) df for each group.

Since there are differences among the observers, the experience levels of the observers are important considerations. The test is based on the means of the collaborators, and a contrast among the means is tested. The hypothesis is

$$H_o: \sum c_i \mu_i = 0 \text{ vs } H_A: \sum c_i \mu_i \neq 0$$

where c_i is the coefficient of the i^{th} mean, chosen so that

$$\sum c_i = 0.$$

Since there are 4 experienced observers and 6 inexperienced observers, the c_i are chosen as

$$H_o\text{: } 6(\mu_1 + \mu_2 + \mu_6 + \mu_7) - 4(\mu_3 + \mu_4 + \mu_5 + \mu_8 + \mu_9 + \mu_{10})$$

is the hypothesis tested. The test statistic is based on Student's t distribution, and given by

$$t_c = \frac{\Sigma\, c_i \bar{x}_i}{\hat{\sigma} \sqrt{\frac{\Sigma c_i^2}{n}}}$$

where n is the number of determinations comprising each mean, in this case 6. Substituting gives

$$t_c = \frac{6(\bar{x}_1 + \ldots + \bar{x}_7) - 4\,(\bar{x}_3 + \ldots + \bar{x}_{10})}{1.82 \sqrt{\frac{6^2 + \ldots + 6^2 + 4^2 + \ldots + 4^2}{6}}}$$

$$t_c = \frac{6(28.65) - 4\,(38.97)}{1.82 \;\; \frac{144 + 96}{6}}$$

$$= \frac{171.90 - 155.88}{1.82\sqrt{40}}$$

$$= \frac{16.02}{11.50} = 1.39$$

This statistic has 45 df associated with it. These df come from the estimated variance, $\hat{\sigma}$, used in the calculation. A t-value with 45 df must exceed 2.016, approximately, to be considered significant at the 0.05 level. Thus, no difference is determined between the experienced and the inexperienced observer for these data.

B.5 Precision Estimation and Tests of Hypotheses for Riverbend Steam Station Test 2

The deviations shown in Table 23 are submitted to an ANOVA procedure to estimate the precision of the method. The model selected is a completely random model represented by

$$Y_{ijk} = \mu + \beta_i + \gamma_{j/i} + \rho_k + \epsilon_{ijk}$$

where

Y_{ijk}	–	the observed value in run k by observer j in group i
μ	–	the overall mean
β_i	–	the effect due to group i
$\gamma_{j/i}$	–	the effect due to collaborator j in group i
ρ_k	–	the effect due to run k

and

ϵ_{ijk}	–	the random error associated with Y_{ijk}

The observer term is nested within groups, since a different set of observers comprised each group. The run term is crossed with both the group and the observer terms. The error term is estimated from the interaction of the run and group, and of the run and observer, with the interactions assumed to be zero. The resulting ANOVA table is in Table B5.

The EMS column tells which mean squares provide the proper denominator for the F-ratio, and how the resulting variance components may be estimated.

The within-observer variance is estimated from the error term. Since the EMS is σ^2, the estimated variance is

$$\hat{\sigma}^2 = 3.37,$$

the mean square for error. There are 88 df for this estimate. The standard deviation is estimated as

$$\hat{\sigma} = \sqrt{3.37}$$
$$= 1.84 \text{ percent opacity.}$$

TABLE B5. ANALYSIS OF VARIANCE FROM STEAM STATION TEST 2

Source	df	SS	MS	F	EMS
Groups	1	156.77	156.77	5.56*	–†
Runs	10	1866.06	186.61	–†	–†
Observers/Groups	8	225.71	28.21	8.37‡	$\sigma^2 + 11\sigma_L^2$
Error	88	296.21	3.37		
Total	107	2544.75			

*Significant. $F_{0.05}$ (9,8) = 5.32.
†Not of interest.
‡Significant. $F_{0.05}$ (8,88) ≈ 2.10.

The observer bias term is estimated from the observer term in the ANOVA. The denominator for the observer F-ratio is the error term, and the test statistic is F_c = 8.37 with 8 and 88 df. The tabled F value at the 5-percent level is approximately 2.10 and the observer term is determined to be significant. The expected value of the mean square is $\sigma^2 + 11\sigma_L^2$. Thus, we estimate the variance by

$$\hat{\sigma}_L^2 = \frac{MS_L - \hat{\sigma}^2}{11}$$

where MS_L is the observer/group mean square. Substituting gives

$$\hat{\sigma}_L^2 = \frac{28.21 - 3.37}{11}$$
$$= \frac{24.84}{11}$$
$$= 2.26$$

with 8 df. The degrees of freedom are calculated as $n - 1 = 5 - 1 = 4$ from each group. The estimated observer bias standard deviation is

$$\hat{\sigma}_L = \sqrt{2.26}$$
$$= 1.50 \text{ percent opacity.}$$

The F-ratio for groups is the ratio of the group mean square to the observer/group mean square. The statistic is F_c = 5.56. To determine if this value is significant, F_c is compared to a tabled F value at the 5-percent significance level. The tabled value is $F_{0.05}(1.8)$ = 5.32, so that a significant difference can be said to exist between the two groups in the determination of average opacity. This is an indication that the angle of view is an important consideration in the visual estimation of opacity.

Since it has been determined that there are differences among the observers, the hypothesis that the experienced observers differ from the inexperienced can be tested. This is done by use of a contrast among the means. A contrast is represented by $\Sigma c_i \mu_i$, where the c_i are coefficients applied to the means, constructed so that $\Sigma c_i = 0$. For this test, the null hypothesis is stated as

$$H_o\colon 6(\mu_1 + \mu_2 + \mu_6 + \mu_7) - 4(\mu_3 + \mu_4 + \mu_5 + \mu_8 + \mu_9 + \mu_{10}) = 0$$

where

μ_i — the mean for observer i.

Observers 1, 2, 6 and 7 are experienced, and 3, 4, 5, 8, 9 and 10 are relatively inexperienced. The test statistic is given by

$$t_c = \frac{\Sigma c_i \bar{x}_i}{\hat{\sigma}\sqrt{\dfrac{\Sigma c_i^2}{n}}}$$

where

$\bar{x}_i$ — the sample mean for observer i

$\hat{\sigma}$ — the estimate of the within-observer standard deviation

and

n — the number of determinations that comprise $\bar{x}_i$.

For these data, the statistic is:

$$t_c = \frac{(6)(29.60) + (-4)(37.20)}{1.84\sqrt{\dfrac{240}{10}}}$$

$$= \frac{177.60 - 148.80}{1.84\sqrt{24}}$$

$$= \frac{28.80}{9.02}$$

$$= 3.19$$

There are 88 df for the statistic taken from the estimated variance. By comparing this statistic to a tabled t with 88 df, $t_c = 3.19$ exceeds $t_{0.05}(88) \approx 1.99$ and thus a significant difference does exist. The conclusion is that the experienced observers did determine a higher average opacity than the inexperienced.

B.6 Precision Estimation and Tests of Hypotheses for Riverbend Steam Station Test 3

The determinations of average opacity made by the collaborators as shown in Table 27 are submitted to an ANOVA procedure to estimate the precision components. The model for these data is a completely random model,

$$Y_{ijk} = \mu + \beta_i + \gamma_{j/i} + \rho_k + \rho\beta_{ik} + \epsilon_{ijk}$$

where

Y_{ijk} — the observation on run k by observer j in group i

μ — the overall mean

β_i — the effect of the i^{th} group

$\gamma_{j/i}$ — the effect of the j^{th} collaborator in group

ρ_k — the effect due to run k

$\rho\beta_{ik}$ — the interaction between run k and group i

and

ϵ_{ijk} — the random error associated with

The two groups are considered together with respect to precision under the assumption that a more direct sun angle would not affect the precision of the individual observer. To test this further, separate estimates of the error variance are made using the run by observer interaction within each group. The estimated values of σ^2 are

$$\hat{\sigma}_c^2 = 3.49$$

$$\hat{\sigma}_E^2 = 4.09$$

for the control and experimental groups, respectively. To test whether the experimental estimate is significantly larger, an F-ratio is used. The calcualted statistic is

$$F_c = \frac{4.09}{3.49}$$

$$= 1.17$$

with 45 and 45 df. The critical value is $F_{0.05}\ (45, 45) \approx 1.69$, and the two values may be said to be estimating the same true value.

The ANOVA is summarized in Table B6 for the third test. The evaluation is made using the 16 runs where the experimental group read in 5-percent increments. The EMS column provides the basis for testing for significant effects and estimating the variance components.

TABLE B6. ANALYSIS OF VARIANCE FROM STEAM STATION TEST 3

Source	df	SS	MS	F	EMS
Groups	1	2.88	2.88	0.31*	$\sigma^2 + 16\sigma_L^2 + 32\sigma_{RG}^2 + 64\sigma_G^2$
Runs	15	9440.54	629.37	–†	–†
Group X Run	15	138.39	9.23	–†	$\sigma^2 + 16\sigma_L^2 + 32\sigma_{RG}^2$
Observer/Group	6	321.05	53.51	14.99‡	$\sigma^2 + 16\sigma_L^2$
Error	90	313.73	3.57		σ^2
Total	127	10216.59	698.56		

*Not significant.
†Not of interest
‡Significant. $F_{0.05}\ (6,90) \approx 2.30$.

The error term comes from the observer-run interaction. The assumption made is that this effect is equal to zero, which is equivalent to saying that if a particular observer reads higher opacities than the others on one run, he would not read lower on other runs. The EMS of the error term is σ^2, which gives an estimated within-observer variance of

$$\hat{\sigma}^2 = 3.57$$

with 90 df. This results in a within-observer standard deviation estimate of

$$\hat{\sigma} = \sqrt{3.57}$$

$$= 1.89.$$

The EMS for observers/group is $\sigma^2 + 16\sigma_L^2$, since each observer mean is the average over the 16 runs. The F-ratio is the ratio of the observer mean square to the error mean square. This gives $F_c = 14.99$ with 6 and 90 degrees of freedom. The critical value is $F_{0.05}(6,90) \approx 2.30$. Thus, there is a significant observer effect at the 5-percent level.

To estimate the observer variance component, the EMS is again used.

$$\hat{\sigma}_L^2 = \frac{\mathrm{MS}_L - \hat{\sigma}^2}{16}$$

where MS_L is the observer mean square, and $\hat{\sigma}^2$ is the error mean square. Substituting gives

$$\hat{\sigma}_L^2 = \frac{53.51 - 3.57}{16}$$

$$= \frac{49.94}{16}$$

$$= 3.12.$$

There are 6 df for laboratories calculated as (4-1) for each group. The estimated observer bias standard deviation is

$$\hat{\sigma}_L = \sqrt{3.12}$$

$$= 1.77 \text{ percent opacity}$$

To test whether there is a significant group effect, an F-ratio is constructed for the group term. The denominator is the group by run interaction, and the statistic is $F_c = 0.31$ with 1 and 15 df. The critical value is $F_{0.05}\ (1, 15) = 4.54$, and no group effect is found. This indicates that the angle of observation had essentially no effect in this test.

The means of the observers over the 16 runs are used to determine if there was a significant difference in this test between the experineced and the inexperienced observers. A contrast among the means is used for the test. The hypothesis is of the form

$$H_O: \Sigma c_i \mu_i = 0 \quad \text{versus} \quad H_A: \Sigma c_i \mu_i \neq 0$$

where

$$\Sigma c_i = 0$$

and

μ_i is the mean of collaborator i.

For this test, the c_i are chosen to be 1 for the experienced and −1 for the inexperienced observer means. The test statistic is given by

$$t_c = \frac{\Sigma c_i \bar{x}_i}{\hat{\sigma}\sqrt{\frac{\Sigma c_i^2}{n}}}$$

where

$\bar{x}_i$ is the sample mean for collaborator i

$\hat{\sigma}$ is the estimated within-observer standard deviation

and

n is the number of determinations used to obtain $\bar{X}_i$.

Substituting from Table 31 gives

$$t_c = \frac{(15.31 + \ldots + 14.46) - (11.40 + \ldots + 12.76)}{1.89\sqrt{\frac{8}{16}}}$$

$$= \frac{(58.72) - (46.81)}{(1.89)(0.71)}$$

$$= \frac{11.91}{1.34}$$

$$= 8.89.$$

There are 90 df associated with this statistic, and this gives a critical value of $t_{0.05}\,(90) \approx 2.00$. Thus, there was a significant difference between the experienced and inexperienced observers.

B.7 Precision Estimation for Riverbend Steam Station Test 3, 1-Percent Increment Runs

The determinations made in increments of 1 percent are submitted to an ANOVA procedure according to the completely random model

$$Y_{ij} = \mu + \rho_i + \gamma_j + \epsilon_{ij}$$

where

Y_{ij} – the observation in run i by observer j

μ – the overall mean

ρ_i – the effect due to run i

γ_j – the effect due to observer j

and

ϵ_{ij} — the random error associated with Y_{ij}

The model assumes that there is no interaction between run and observer. This assumption is equivalent to saying that the observers will be consistently high or low with respect to each other, regardless of the density of the plume. By making this assumption, blocking the runs in order to make an estimate of the within-observer variance is not necessary.

TABLE B7. ANALYSIS OF VARIANCE FROM STEAM STATION TEST 3, 1% VARIATION

Source	df	SS	MS	F	EMS
Runs	7	2105.62	–*	–*	–*
Observers	3	45.75	15.25	1.84†	
Error	21	174.07	8.29		
Total	31	2325.44			

*Not of interest.
†Not significant. $F_{0.05}$ (3.21) = 3.07.

There were 8 runs by 4 the observers in the control group for a total of 32 determinations. The ANOVA is summarized for these data in Table B7. The EMS of the error term is σ^2, the within-observer variance. Thus, the estimated value of σ^2 is

$$\hat{\sigma}^2 = 8.29$$

with 21 df. This gives a within-observer standard deviation of

$$\hat{\sigma} = \sqrt{8.29}$$
$$= 2.88.$$

The EMS of the observer term is $\sigma^2 + 8\sigma_L^2$. Thus, the ratio of the observer term to the error term tests for significance. The resultant F-ratio is F_c = 1.84. The critical value for this test is $F_{0.05}$ (3,21) = 3.07, and thus, the observer term is not significant. This implies that the between-observer term, σ_b^2, is identical to σ^2, and gives

$$\hat{\sigma}_b = \hat{\sigma} = 2.88 \text{ percent opacity}$$

as before.

The within-observer term for these runs is higher than that for the previous runs. An F-test is conducted to determine if a significant difference exists between the true variances of the two sets. The F-ratio is

$$F_c = \frac{8.29}{3.57}$$
$$= 2.32$$

with 21 and 90 df. The critical value for this test is $F_{0.05}$ (21, 90) ≈ 1.75. Thus, the hypothesis of equality is rejected, and we can say that there is more variability reading in 1 percent increments.

To determine if there was any difference between the performance of the two groups, a second ANOVA is done in which the control group data is included. The model is

$$Y_{ijk} = \mu + \beta_i + \gamma_{j:i} + \rho_k + \rho\beta_{ik} + \epsilon_{ijh}$$

where

Y_{ijk} — the determination by observer j of group i in run k

μ — the overall mean

β_i – the effect due to group i

$\gamma_{j/i}$ – the effect due to observer j in group i

ρ_k – the effect due to run k

$\rho\beta_{ik}$ – the run by group interaction

and

ϵ_{ijk} – the random error associated with Y_{ijk}

As before, the run by observer interaction is assumed to be zero and used to estimate the error term.

TABLE B8. ANALYSIS OF VARIANCE FROM STEAM STATION TEST 3, 1% RUNS

Source	df	SS	MS	F	EMS
Groups	1	268.96	268.96	4.61*	$\sigma^2 + 8\sigma_L^2 + 16\sigma_{RG}^2 + 32\sigma^2 G$
Runs	7	5549.59	792.80	–†	–†
Group X Run	7	408.16	58.31	–†	$\sigma^2 + 8\sigma_L^2 + 16\sigma_{RG}^2$
Observer/Group	6	127.69	21.28	–†	$\sigma^2 + 8\sigma_L^2$
Error	42	476.61	11.35		σ^2
Total	63	6831.01			

*Not significant. $F_{0.05}(1,7) = 5.59$.
†Not of interest.

The ANOVA is summarized in Table B8. The only term of interest is the group term. The F-ratio is formed by dividing the group mean square by the group by run mean square. The calculated value is $F_c = 4.61$ with 1 and 7 df. The critical value is $F_{0.05}(1,7) = 5.59$, and so no group effect can be determined.

B.8. Significance of Reduction of Frequency of Large Deviations

To determine if a significant decrease is noted in the frequency of large deviations, a confidence interval is set up around the observed proportion. The confidence interval is constructed using

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

where

$\hat{p}$ – the sample proportion

n – the number of items in the sample

and

1.96 – the standard normal deviate at the 95-percent confidence level.

The proportion of large deviations (absolute value greater than 7.5) in the sample is obtained by taking the number of large deviations, and dividing by the size of the sample. This gives

$$\hat{p} = \frac{15}{144}$$

$$= 0.10.$$

Substituting this into the above formula gives

$$CI_{0.95} : 0.10 \pm 1.96\sqrt{\frac{(0.10)(1-0.10)}{144}}$$

$$: 0.10 \pm 1.96(0.03)$$

or from 0.05 to 0.15

The frequency of large deviations in the original data set is 15/96 or 0.16. Thus, since this falls outside the above confidence interval, the two may be said to represent different true proportions of occurrence. Thus, the frequency of large deviations was reduced by averaging two observers' result.

B.9. Multiple Comparison Tests and Composite Estimation

The slopes of the three regression lines obtained as representative of observer deviation from plume opacity are compared for equality using the least significant difference (LSD) test. The test statistic provides a number which must be exceeded in order for two sample values to be said to represent different true values.

The standard deviations of the slopes is calculated as in Appendix B.1 for each regression line. For the white training generator smoke, black training generator smoke and steam station data, these values are 0.04, 0.05, and 0.04, respectively. Since these are close to each other, it is reasonable to assume that there is a common variance, s_b^2, for all three slopes, and to estimate it by pooling the three estimates. A pooled variance is of the form

$$s_p^2 = \frac{\sum_{i=1}^{k} df_i s_i^2}{\sum_{i=1}^{k} df_i}$$

where

df_i – the number of degrees of freedom in the i^{th} sample

s_i^2 – the variance of the i^{th} sample

k – the number of samples.

Pooling the variance gives an estimated value of $s_p^2 = 0.0019$ with 393 df. The df are obtained by summing the df for the individual variance estimates.

The LSD at the 5-percent significance level is given by

$$LSD_{0.05} = 1.96\sqrt{2s_p^2}$$

$$= 1.96\sqrt{0.0038}$$

$$= 0.12.$$

All three slopes can be said to be equivalent if they differ from each other by no more than 0.12. The difference between the greatest and least values is 0.34 – 0.24 = 0.10, so that all three slopes can be said to be estimating the same true value.

LIST OF REFERENCES

1. Environmental Protection Agency, "Standards of Performance for New Stationary Sources," *Federal Register*, Vol. 36, No. 247, December 23, 1971, pp. 24876-24893.

2. Environmental Protection Agency, "Stationary Sources, Proposed Emission Monitoring and Performance Testing Requirements," *Federal Register*, Vol. 39, No. 177, September 11, 1974, pp. 32852-32874.

3. Dixon, W.J. and Massey, F.J., Jr., **Introduction to Statistical Analysis**, 3rd Edition, McGraw-Hill, New York, 1969.

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

1. REPORT NO. EPA-650/4-75-009	2.	3. RECIPIENT'S ACCESSION NO.
4. TITLE AND SUBTITLE "Evaluation and Collaborative Study of Method for Visual Determination of Opacity of Emissions from Stationary Sources."		5. REPORT DATE January 1975 (date of approval)
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Henry F. Hamil, Richard E. Thomas and Nollie F. Swynnerton		8. PERFORMING ORGANIZATION REPORT NO. SwRI #01-3462-006
9. PERFORMING ORGANIZATION NAME AND ADDRESS Southwest Research Institute, 8500 Culebra Road, San Antonio, Texas 78284		10. PROGRAM ELEMENT NO. 1HA327
		11. CONTRACT/GRANT NO. 68-02-0626
12. SPONSORING AGENCY NAME AND ADDRESS U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C. 20460		13. TYPE OF REPORT AND PERIOD COVERED Task Order
		14. SPONSORING AGENCY CODE
15. SUPPLEMENTARY NOTES		

16. ABSTRACT

This study presents the results of statistical analyses of determinations of the average opacity of emissions. The determinations are based on 25 readings made by qualified observers at three types of sources. The tests were conducted 1) using training smoke generators, 2) at a sulfuric acid plant and 3) at a coal-fired power plant. Observers used in the tests came from both enforcement agencies and private companies, and varied in the amount of field experience in the use of the method.

For each test, the accuracy and precision of the determinations is investigated. Accuracy is measured by the deviation of the determination from a concurrent meter average, used as true opacity for these tests. The precision of the method is measured as standard deviations for within-observer, observer bias and between-observer terms. For each test, the expected range of determinations by a single observer and the maximum difference expected between two observers are calculated.

The results of all the tests are used to estimate the accuracy and precision of the method independent of the nature of the site or color of the plume. These estimates show a reasonable degree of accuracy and precision for the method when used in the probable range of opacities that would occur in field use, 0 to 35 percent opacity.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Emission Field tests Statistical analysis Plume Opacity	EPA Method 9 Stationary Sources Visual method Smoke generator Collaborative testing	13B 14B

18. DISTRIBUTION STATEMENT Unlimited	19. SECURITY CLASS (This Report) Unclassified	21. NO. OF PAGES 70
	20. SECURITY CLASS (This page) Unclassified	22. PRICE

EPA Form 2220-1 (9-73)

ERRATA

EPA-650/4-75-009

EVALUATION AND COLLABORATIVE STUDY OF METHOD FOR VISUAL DETERMINATION OF OPACITY OF EMISSIONS

by

Henry F. Hamil
Richard E. Thomas
Nollie F. Swynnerton

The following correction is applicable to the above cited report which was issued under Project No. 01-3462-006, dated January, 1975.

Page 32

Paragraph 2 is deleted in its entirety and the following paragraph is substituted:

This value is comparable to that for the 5-percent increment data. From a second ANOVA shown in Appendix B.7, the two groups of observers cannot be shown to be reading different average opacities on these runs.