

Final Report

DEVELOPMENT OF PROCEDURES FOR ESTABLISHING THE UNCERTAINTIES OF EMISSION ESTIMATES

ARB Contract No. A5-184-32
VRC Project No. 1055

by:

Yuji Horie, Ph.D.
Arthur L. Shrope

Valley Research Corporation
15904 Strathern Street, Suite 22
Van Nuys, California 91406

LIBRARY
CALIFORNIA AIR RESOURCES BOARD
P.O. BOX 2015
SACRAMENTO, CA 95812

Prepared for:

CALIFORNIA AIR RESOURCES BOARD

Research Division
1800 15th Street
Sacramento, California 95814
Project Officer - Joseph Pantalone

February, 1988

DISCLAIMER

The statements and conclusions in this report are those of the Contractor and not necessarily those of the State Air Resources Board. The mention of commercial products, the source of their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
EXECUTIVE SUMMARY	i
1.0 INTRODUCTION	1-1
1.1 Uncertainty: The Conceptual Basis	1-1
1.2 Subjective Perception of Uncertainty	1-2
1.3 Scope of This Report	1-3
2.0 REVIEW OF EARLIER UNCERTAINTY STUDIES	2-1
2.1 AP-42 Emission Factor Rating	2-1
2.2 SCAQMD's Delphi Approach	2-2
2.3 BAAQMD's Lognormal Approach	2-3
2.4 NAPAP Uncertainty Study	2-4
3.0 EXPLORATION OF IDEAL METHODOLOGY	3-1
3.1 Attributes for Ideal Methodology	3-1
3.2 Standard Method	3-2
3.3 Scatter Plot Method	3-3
3.4 Handbook Method	3-5
4.0 RECOMMENDATIONS FOR FUTURE STUDIES	4-1
4.1 Development of Procedural Manual	4-1
4.2 Development of Tutorial Material	4-2
4.3 Uncertainty Study on Sample Inventory	4-3
5.0 REFERENCES	5-1
APPENDIX SCATTER PLOT METHOD QUESTIONNAIRE	

EXECUTIVE SUMMARY

Uncertainties in both individual emission estimates and a compiled emission inventory as a whole are not well characterized. Since most emission estimates are not based on repeated measurements of the quantities, the conventional statistical formula for a confidence interval around the estimated mean of reported measurements does not apply. Therefore, an alternative formulation for quantifying uncertainty of an estimated quantity must be devised. To this end, an extensive review of earlier uncertainty studies on emission estimates and other studies on subjectively evaluated uncertainties has been conducted.

The review revealed that all earlier studies failed to address bias uncertainties which are considered to be very important for quantifying overall uncertainty in emission estimates, because of the use of models and engineering analyses in those estimates instead of repeated measurements. Another finding is that there exists no established method of quantifying subjectively evaluated uncertainties in emission estimates.

The absence of an adequate method for quantifying uncertainties in emission estimates led the authors to explore and develop a new method that will facilitate assessing uncertainties in individual emission estimates and in emission inventory as a whole. In exploring such a method, the necessary attributes for an "Ideal Method" were first identified. Although an "Ideal Method" does not exist, its attributes are definable from the lessons learned from the failures and flaws found in the earlier studies. These attributes are:

Attribute 1 Ideal Methodology must be capable of quantifying both precision and bias uncertainties in original estimates.

Attribute 2 Ideal Methodology must be applicable to any levels of data and information availability in original estimates.

Attribute 3 Ideal Methodology must be compatible with the statistical theories for normal and lognormal variates.

Attribute 4 Ideal Methodology must be capable of tracking both random and systematic errors that propagate through aggregations of emission variables and source categories.

Attribute 5 Ideal Methodology must be technically sound and easy to use.

Since a new methodology has to be applicable not only for quantifying uncertainties in individual estimates but also for calculating those uncertainties that propagate through various aggregations of individual estimates into a broader source category, the methodology must be based on the same probability scale as that used in statistically defined uncertainty parameters, such as a 95 percent confidence interval and relative (random) error. Three new methods were devised under this study: the standard method (SM), the scatter plot method (SP), and the handbook method (HP). All three methods base their uncertainties on the same probability scale as used in the statistical uncertainty so that they are all compatible with the statistical formulas for calculating propagated errors.

The SM relies on the concept of betting in quantifying subjectively evaluated uncertainties in the estimated quantities. Although the concept of betting is widely used in cognitive psychology studies and decision making analysis, the SM's connotation of gambling was disliked by the reviewing members. Since the SM has other flaws as well, the method was abandoned in favor of the scatter plot method (SP), which relies on the familiarity of most emission inventory analysts with scatter plots in showing variability of the data and uncertainty of the estimate derived from such data. In the SP, a set of six typical scatter plots showing the state of data variability was used as a tutorial device for subjectively evaluating uncertainties of the estimated quantities. It was found that this method did not work consistently from application to application or from evaluator to evaluator.

After the two unsuccessful attempts for a new methodology, the handbook method (HP) was devised. This method is used and fully described in the handbook entitled, "Procedures for Establishing the Uncertainties of Emission Estimates". The HP can be summarized in the following four steps:

- Step 1 An inventory analyst evaluates a promulgated emission estimate and provides two alternative estimates: upper plausible estimate and lower plausible estimate.
- Step 2 Based on the three estimates, the analyst computes upper and lower threshold levels (i.e., UL and LL) such that UL is one standard deviation above the mean of the three estimates and LL is one standard deviation below the mean.

Step 3 Evaluators are asked to estimate odds that the true mean emission is below UL and odds that the true mean is below LL.

Step 4 Based on the evaluator's responses, subjectively evaluated precision and bias uncertainties in the promulgated emission estimate are then graphically determined using normal and lognormal probability papers.

Unlike SM, the Handbook Method does not require an evaluator to undertake a betting game involving a conceptual urn filled with red and blue balls. In HM, an evaluator is given a concise, clear narrative describing rationale and reasons for each of the three estimates so that he has a sounder base to think about the odds that the true mean is found below UL and LL. Unlike SP, HM does not require an evaluator to speculate about imaginary data points around the emission estimate in order to attach his perceived uncertainty measure to the estimate.

As to the five attributes of Ideal Methodology, HM appears to possess all five attributes.

1.0 INTRODUCTION

This final report is a supplement to the handbook entitled "Procedures for Establishing the Uncertainties of Emission Estimates". Since the uncertainty handbook (hereafter called "Handbook") discusses all practical aspects of the ways to obtain both objectively calculated and subjectively evaluated uncertainties of emission estimates, no further discussions on the same subject are presented in this final report. Instead, the report provides overviews of the emissions uncertainty studies conducted in the past and of the background work from which the methodology presented in "Handbook" evolved. It also provides a few recommendations for further work on the uncertainties of emission estimate.

1.1 UNCERTAINTY: THE CONCEPTUAL BASIS

As a matter of simple fact, uncertainty exists in any measurement, observation, engineering estimate, quantitative guess, or mathematical model simulation. When a person estimates the distance to a nearby building "by eyeball", his estimate will be more uncertain than if he uses a measuring device. In contrast, a tool and die machinist should be able to estimate the thickness of a particular piece of paper more accurately than he can measure it with a foot ruler. A plant engineer using mass balance principles together with fuel consumption and sulfur content data may be able to estimate monthly or annual SO₂ emissions from a power generating unit as accurately as (and perhaps more accurately than) by calculating it from data generated by a continuous emission monitor.

An engineering estimate is not always inferior to a measured quantity nor is an expert estimate always inferior to an emission value calculated from some emission model. Uncertainty of an engineering estimate depends on the quality of information available, whereas uncertainty of a measured quantity depends on the precision and accuracy of the measurement device. When a model is used, uncertainty depends on the accuracy of the model and of the data input to the model.

In statistical terms, uncertainty of an estimated quantity may be thought of as analogous to random errors in repeated measurements. "Precision" indicates how close those measured values are to each other.

The term "uncertainty" is inversely related to precision. If discrepancies between the individual results in a series of measurements are very small, the precision of measurement is said to be high, while the uncertainty of the measurement is small.

On the other hand, not all error associated with measurement is referable to defects in precision of measurement. An engineer can easily measure a flat piece of material to a precision of ± 0.5 mm using a meter rule; but, if the rule he uses is not accurate, his result will be inaccurate, even though precise. Accordingly, if he is uncertain as to the accuracy of his measuring stick, he must also be uncertain about the accuracy of his result, regardless of its precision. This second type of uncertainty is ascribed to systematic error, also called bias. It is "systematic" in that it is characteristic of the particular tool or system used to generate the estimate; it cannot be accounted for or corrected without a specific investigation of that system.

It should be noted that most emissions estimation systems are not based on repeated measurements, but rather on engineering understanding of the problem, exemplified by an emission model. It has been observed, however, that most, if not all, published studies of uncertainties in emission inventory work have addressed only the precision component of uncertainty (e.g., EPRI 1981; SCAQMD 1982; Mangat et al. 1984; and Chun 1986).

1.2 SUBJECTIVE PERCEPTION OF UNCERTAINTY

When an engineer estimates a quantity by measurement, engineering analysis, or mathematical simulation, he ordinarily has a notion or belief as to the accuracy of his estimate. This subjective notion of uncertainty of the estimated value often finds expression in phrases such as "within ± 10 percent", "about right", "perhaps an underestimate", and so on. The "Handbook" is intended to distill this subjectively perceived uncertainty onto a consistent, measurable probability scale.

Emission estimates are seldom made by direct measurement. Instead, most emission estimates are usually arrived at through emission models or algorithms which are supposed to show how emissions are related to process variables. These commonly consist of emission factor equations relating

emissions to activity levels of particular emission sources or source categories. Emission estimates are therefore subject to both random error (also called "precision uncertainty"), and systematic error (also called "bias uncertainty" or simply "bias").

Since practically no emissions in an inventory are directly measured, the uncertainty due to possible model inaccuracies may be very important. It can be adequately assessed only by experts who are familiar with the physical principles which operate to cause pollutant emissions and with the models and algorithms which have been used to simulate those principles. Although detailed analysis of this sort can be time-consuming and costly, the same experts who have the necessary understanding to achieve such an analysis will often have useful insights into the possible sources of bias in particular model applications. In evaluating uncertainties of emission estimates for inventory purposes, this type of uncertainty must always be considered and, where possible, the opinions of qualified experts should be obtained.

Therefore, the cruxes of uncertainty evaluation are:

1. How to project subjectively perceived uncertainty onto a uniform and consistent uncertainty scale;
2. How to link this subjectively perceived uncertainty to statistically determinable uncertainty; and
3. How to track uncertainties in individual estimates of emissions and emission model parameters through various multiplicative and additive processes, which are involved in emission inventory calculations.

1.3 SCOPE OF THIS REPORT

In "Handbook", plausible procedures for assessing the uncertainties of emission estimates are developed and discussed with a few illustrative examples. In this final report, a brief history of the development of such procedures will be discussed, together with some recommendations for testing the procedures through a pilot study or for promoting their acceptance and use by preparing tutorial materials. Before developing these procedures, VRC conducted a thorough review of earlier studies on uncertainties of emission estimates, then attempted two unsuccessful methods: the "standard method" (SM) and the "scatter plot method" (SP). This report describes major findings from the review and the key features

of the unsuccessful methods in comparison with the method presented in "Handbook".

2.0 REVIEW OF EARLIER UNCERTAINTY STUDIES

This section outlines four earlier studies on uncertainty assessment for emission estimates: the AP-42 emission factor rating system (EPA 1985); SCAQMD's "Delphi" approach (SCAQMD 1982); BAAQMD's "Lognormal" approach (Mangat et al. 1984); and the NAPAP uncertainty study (EPA 1986). Major features of the methods used in these studies, as well as some of their deficiencies, are discussed in subsections that follow.

2.1 AP-42 EMISSION FACTOR RATING

This EPA document (EPA 1985), which compiles emission factors of all stationary source categories, uses a rating scheme for indicating the reliability and accuracy of AP-42 emission factors. A rating of A through E, with A being the best, is given for each emission factor to reflect the quality and the extent of data on which the factor is based. In general, factors based on many observations or on more widely accepted test procedures are assigned higher rankings. For instance, an emission factor based on ten or more source tests on different plants would likely get an A rating, whereas a factor based on a single observation of questionable quality, or one extrapolated from another factor for a similar process, would probably be labeled D or E.

The AP-42 rating scheme may be reasonable for indicating the reliability of a promulgated emission factor. However, it is not very useful for assessing the uncertainty of the emission factor or the uncertainty of an emission estimate obtained from an application of the emission factor to a particular emission situation. The reasons are as follows:

1. The rating system used in AP-42 is coherent on a plausibility scale in that an emission factor with an A rating may be more trustworthy than one with a B rating in the same source category, but it is not coherent on a probability scale since a factor with an A rating in one category may still be less accurate than the one with a B rating in another category;
2. The rating system used in AP-42 is not coherent in another sense: an emission factor with a C rating may be more reliable than one with an A rating because the former is more applicable to a particular situation under which the emission is estimated; and

3. The rating system A through E does not provide any quantitative measure of how uncertain or how reliable the emission factor, or an emission estimate derived from it, is.

2.2 SCAQMD's "DELPHI" APPROACH

For the 1979 base year inventory, the South Coast Air Quality Management District (SCAQMD) used a panel of emission inventory experts to assign an uncertainty to the emission estimate of each major source category. Seventeen individuals closely involved in estimating 1979 Basin emissions were asked to estimate the relative uncertainty of the emission estimate of each emittant in 47 source categories having similar processes. Based on twelve responses received, SCAQMD first calculated a composite relative uncertainty for each source category and then estimated the uncertainty in the basin total emission using the following equation:

$$CV_{total} = \text{SQRT} [\text{SUM}(u_i CV_i)^2] / u_{total} \quad (2-1)$$

where

- CV_{total} = relative uncertainty of the basin total emissions,
- CV_i = relative uncertainty of the i-th source category emissions,
- u_{total} = estimated emissions of the basin total, and
- u_i = estimated emissions of the i-th source category.

The resulting uncertainty values appear to be reasonable. The composite relative uncertainties for individual categories range from 20 percent for fuel combustion sources to 60 percent for miscellaneous processes whereas those for the basin totals range from 9 percent for reactive organic gas emissions to 19 percent for particulate emissions.

Despite the reasonable uncertainty values arrived at, the method used by the SCAQMD has the following problems:

1. Although the relative uncertainty, CV, is not theoretically limited to 100 percent, even for the "missing sources" category it was estimated to be 100 percent;
2. All the uncertainties in emission estimates are assumed to be random errors and mutually independent, although most emission

estimates are expected to contain both random and systematic errors because of the heavy reliance on emission models and engineering analysis instead of repeated measurements; and

3. The uses of intimately involved individuals and direct questioning of relative uncertainties in producing their uncertainty estimates appear to be highly conducive to cognitive biases in their responses for justifying their original emission estimates.

2.3 BAAQMD's "LOGNORMAL" APPROACH

In assessing the uncertainties of emission estimates, the Bay Area Air Quality Management District (BAAQMD) assumed that every emission-related parameter (such as activity level or emission factor) was lognormally distributed and that all were mutually independent. Under such assumptions, the statistical distribution of a source category emission, which is usually estimated by the sum of products of activity level, emission factor and other emission related parameters, is also expected to be lognormal. Based on this theoretically derived conclusion, BAAQMD devised a scheme of first estimating the distribution parameter for each emission estimate and then computing the uncertainty of an aggregated emission estimate at various aggregation levels.

In the above scheme, BAAQMD assumes that every estimate of either a source category emission or an emission-related parameter in the emission model is an unbiased estimate of the true median (i.e., geometric mean), which is never known and is always lower than the true arithmetic mean because of the assumed lognormal distribution. A lognormal distribution for the source emission or the emission-related parameter is then specified by the estimated value, supposedly representing the geometric mean, and the 90th percentile value, which is subjectively estimated by an inventory analyst.

Once every source emission and every emission-related parameter are specified by the unbiased estimate of the true geometric mean and the subjectively assessed 90th percentile value, a specially designed computer program takes those estimates as input and computes the corresponding geometric mean and 90th percentile for an aggregated, broader source category and eventually for an entire emissions inventory.

Although this method is unique and has some good features, it contains the following flaws:

1. Given the geometric means for all emission-related parameters and all source emissions, then the method would work. In reality, however, the means usually estimated for those variables are arithmetic means instead.
2. This method requires 90th percentile values, as well as geometric means, to work. Since most engineers do not have a ready intuition about the 90th percentile value on a lognormal distribution, their guesses on that value for a given geometric mean would be doomed to be highly inaccurate.
3. Regardless of the underlying distribution, many emission estimates are subject to both random and systematic errors because of the use of emission models and engineering analysis instead of repeated measurements. This method neglects the presence of any systematic errors (i.e., bias) in any emission estimates.

2.4 NAPAP UNCERTAINTY STUDY

In assessing the uncertainties of emission estimates in the nationwide emission inventory under the National Acid Precipitation Assessment Program (NAPAP), a panel of emission inventory experts were asked to provide illustrative uncertainty estimates on every emission record on a process- and pollutant-specific basis. By assuming the normal distribution for every emission variable and mutual independence among all variables, the relative uncertainty of each source emission, X_i , was first computed from those of the emission-related factors U_{ij} ($j = 1, 2, \dots, m$) for the following multiplicative relationship:

$$X_i = \prod_{j=1}^m U_{ij} \quad (2-2)$$

Then, the relative uncertainty of a broader source category, X , was computed from those of individual source emissions, X_i , ($i = 1, 2, \dots, n$) for the following additive relationship:

$$X = \sum_{i=1}^m X_i \quad (2-3)$$

The equations used for calculating error propagations through the multiplicative and additive processes are the same as that used by SCAQMD and those presented in "Handbook" (Horie 1988).

This NAPAP study represents the most exhaustive effort to date, to quantify the uncertainties in original emission estimates and those propagated through various levels of source aggregations. The study employed the following assumptions:

- o Emission calculation parameters are independent, that is, they do not covary.
- o Emission factors represent true mean values.
- o All estimates are unbiased.
- o The emission parameters can be treated as random variables which are approximately normally distributed.
- o No coding or transcription errors are present.
- o The data are complete; no emissions data, emissions sources, or emissions source categories are missing.

The study drew the following conclusions:

The values of uncertainty estimates for national levels of aggregation appear to be unreasonably small. The emission inventory data do not warrant this degree of confidence. Therefore, the assumptions, data, and methodology must be reexamined for reasonableness. Preliminary indications are that the methodology must be expanded to incorporate other sources of error (e.g., bias, coding, and emission errors). Other elements of the uncertainty estimates are also counterintuitive and should be corrected by improved assumptions, data, and methodology.

In spite of the drawbacks stated above, this NAPAP study provided the authors with many valuable lessons for developing procedures for establishing the uncertainties in emission estimates. These lessons led to the list of necessary attributes for "ideal uncertainty assessment methodology" which are discussed in the next section.

3.0 EXPLORATION OF IDEAL METHODOLOGY

This section explores an "ideal methodology" for assessing the uncertainties of emission estimates. This exploration starts with a specification of the attributes that are necessary for such an "ideal methodology". A few uncertainty assessment methods that were attempted by the authors under this project are next examined in light of those attributes.

3.1 ATTRIBUTES FOR IDEAL METHODOLOGY

From the review of the four earlier studies and other related work found elsewhere, the following observations are arrived at:

1. Statistical formulas for computing propagated errors are well developed and can easily be adapted for a particular set of assumptions that may be made for conducting any uncertainty assessment study;
2. Emission estimates are generally not obtained from an analysis of repeated measurements but rather are made using an emission model or accepted engineering analysis method;
3. Emission estimates thus derived must be expected to contain systematic errors (e.g., model presentation errors) as well as random errors;
4. Except for a few special cases, even the random errors in any estimated emission records (e.g., process throughput and emission factor) are not accurately assessable from observational data;
5. Instead, both precision and bias uncertainties of any estimated quantities are perceived by the analyst who made such estimates or by other experts familiar with the subject area in the form of "confidence" in those estimates;and
6. As is the nature of random errors, the uncertainty of a broad source category (e.g., whole inventory) is strongly influenced by the number of original source categories constituting the broad category as well as the magnitudes of uncertainties in the original category emissions.

In addition to the above six observations derived from the review of earlier uncertainty studies, discussions among the members who were intimately involved in this project led to the following requirements for an acceptable methodology, whatever that methodology is:

- (i) The methodology must be capable of estimating both the precision and bias uncertainties in emission estimates;

- (ii) The methodology should be able to reveal whether the underlying distribution of data (or a variable) is normal or lognormal; and
- (iii) The methodology must be easy enough for engineers and analysts engaged in inventory works to comprehend it.

By considering the above three requirements and the six observations of the earlier uncertainty studies, the following attributes were formulated for an "Ideal Methodology", regardless of what methodology is considered and whether such a methodology is feasible.

Attribute 1. Ideal Methodology must be capable of quantifying both precision and bias uncertainties in original estimates.

Attribute 2. Ideal Methodology must be applicable to any levels of data and information availability in original estimates.

Attribute 3. Ideal Methodology must be compatible with the statistical theories for normal and lognormal variates.

Attribute 4. Ideal Methodology must be capable of tracking both random and systematic errors that propagate through aggregations of emission variables and source categories.

Attribute 5. Ideal Methodology must be technically sound and easy to use.

The following sections provide synopses of three attempted methods, of which the first two were eventually abandoned as impractical.

3.2 STANDARD METHOD

In the Standard Method (SM), it was proposed that an evaluator might estimate the uncertainties of emission estimates by a procedure involving subjective ratings of probabilities. The subjective rating would be done by introspective judgment, comparing the evaluator's confidence in finding a true mean within a given interval around an emission estimate with his confidence in a standard, easily quantified, odds-rating procedure: the probability (i.e., odds) of drawing a ball of a given color from a urn containing balls of two colors.

Further details concerning this approach can be found in Interim Report (Horie & Shrope 1986).

In light of the five attributes for "Ideal Methodology", the Standard Method meets Attributes 1, 2 and 4 satisfactorily but Attributes 3 and 5 rather unsatisfactorily. The underlying distribution in SM is assumed to be normal. Although SM can be applied to a lognormal distribution, SM's ability of detecting a bias is greatly reduced. As to Attribute 5, several reviewers of this study expressed their dissatisfaction of the SM's explicit linkage with the concept of gambling. They complained that most engineers engaged in emission inventory work are more familiar with "statistical probability" than with "odds" exemplified by red balls in an imaginary urn. Although nearly all reports on Delphi method and decision making, which require a determination of subjectively perceived probability of the occurrence of a certain event, stress the need of using a tutorial device such as an urn game or a spinning wheel for visualizing a perceived probability, the authors took their complaints so seriously that further elaboration of the SM was abandoned.

3.3 SCATTER PLOT METHOD

The scatter plot method (hereafter called "SP") was designed to utilize the engineer's familiarity with scatter plots in considering the distribution of and the confidence interval around an estimated mean of the variable of interest. It was presumed that given a scatter plot of the variable, most engineers would be able to determine whether the underlying distribution of the variable is approximately normal or log-normal and whether the confidence interval around an estimated mean of the variable is narrow or wide. Under this presumption, a standard set of six scatter plots (see Figure 3-1) were prepared and used to judge which of the six scatter plots is the closest to the distribution of the variable under consideration. Each scatter plot was characterized by the arithmetic mean, the four upper (i.e., 50) percentiles, a through d, the four lower percentiles, p through s, and the zero level (see Appendix).

To test the workability of SP, a questionnaire packet for self-evaluating expertise in emissions inventory, five training problems for evaluating probabilistic skill, and an example related to dry cleaning emissions was administered to a panel of the 13 technical members selected from three air pollution control agencies: the ARB, the SCAQMD and the BAAQMD. The entire questionnaire packet is presented in Appendix. In

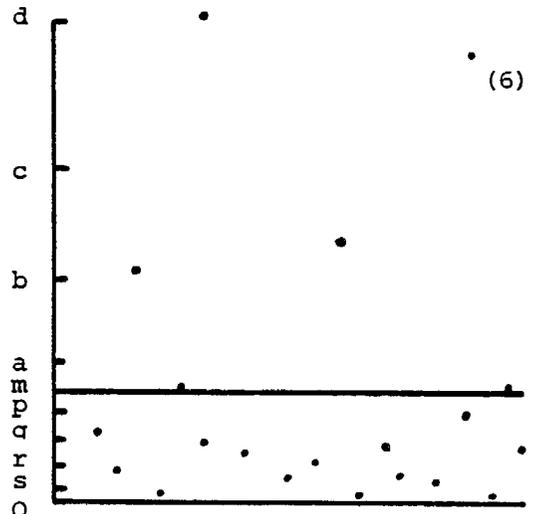
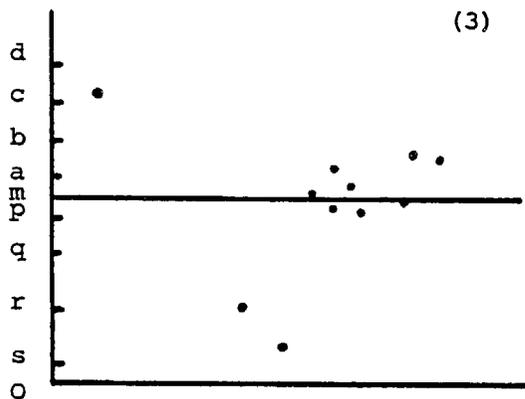
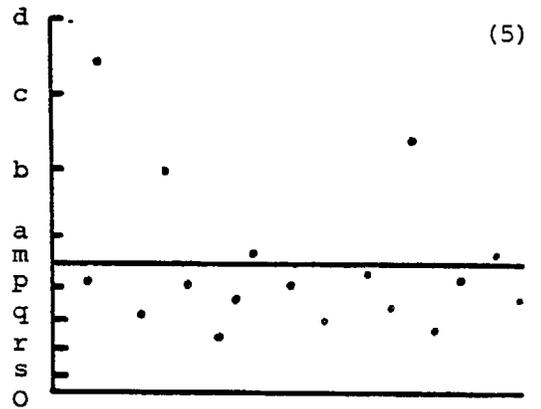
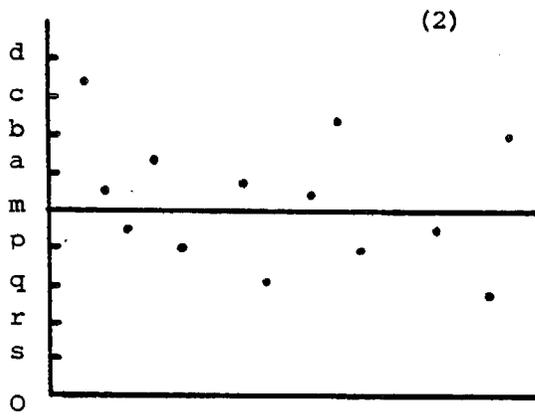
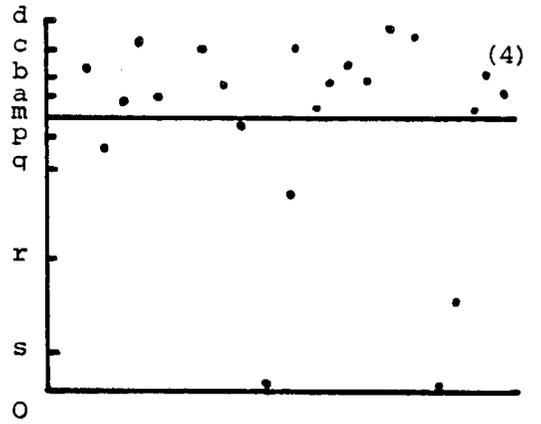
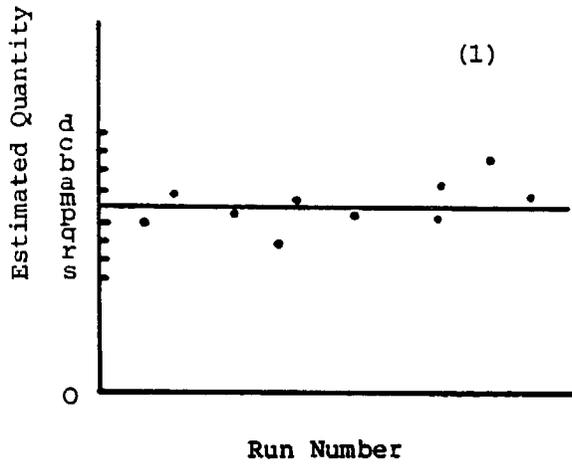


Figure 3-1. Sample Scatter Plots
(m = sample mean, 0 = zero)

this exercise, the authors expected to find a strong positive correlation either between panel members' scores in expertness and those in the emission example, or between scores in probabilistic skill and those in the example, or both. Contrary to this expectation, no significant correlation was found in either case, as seen from the scatter plots in Figures 3-2 and 3-3.

The lack of significant correlations in the above two cases is quite disturbing:

- Q1. Does expertise in emission inventory assessment have nothing to do with a person's ability to correctly assess the uncertainty in an emission estimate?
- Q2. Does probabilistic skill have nothing to do with a person's ability to correctly assess the uncertainty in an emission estimate?

Although the crudity of the experiment represented by the questionnaire packet might have caused a failure in accurately measuring expertise and probabilistic skill, this experiment, nevertheless, implies a danger of casually asking an emission expert an uncertainty question. It appears that an expert's response to such a question may be no more trustworthy than that of a lay person.

In the example emission problem, subjectively evaluated uncertainties in the presented emission estimate were asked for in terms of both SM and SP methods. Panel members' scores in SM and SP on this example were highly correlated, as can be seen from Figure 3-4. This good correlation between scores in SM and those in SP means that the potential danger pointed out above may be applicable to either method.

The authors concluded that the SP had many unresolved problems requiring further testing and refinement and was not worth pursuing any further at least under this project. Despite the problems outlined above, the experiment using the SP has revealed that a procedure for questioning emission experts about the uncertainties in emission estimates must be highly sophisticated to draw trustworthy responses.

3.4 HANDBOOK METHOD

The methodology described in the uncertainty handbook (Horie 1988, hereinafter called "Handbook Method" or "HM") was evolved late in the

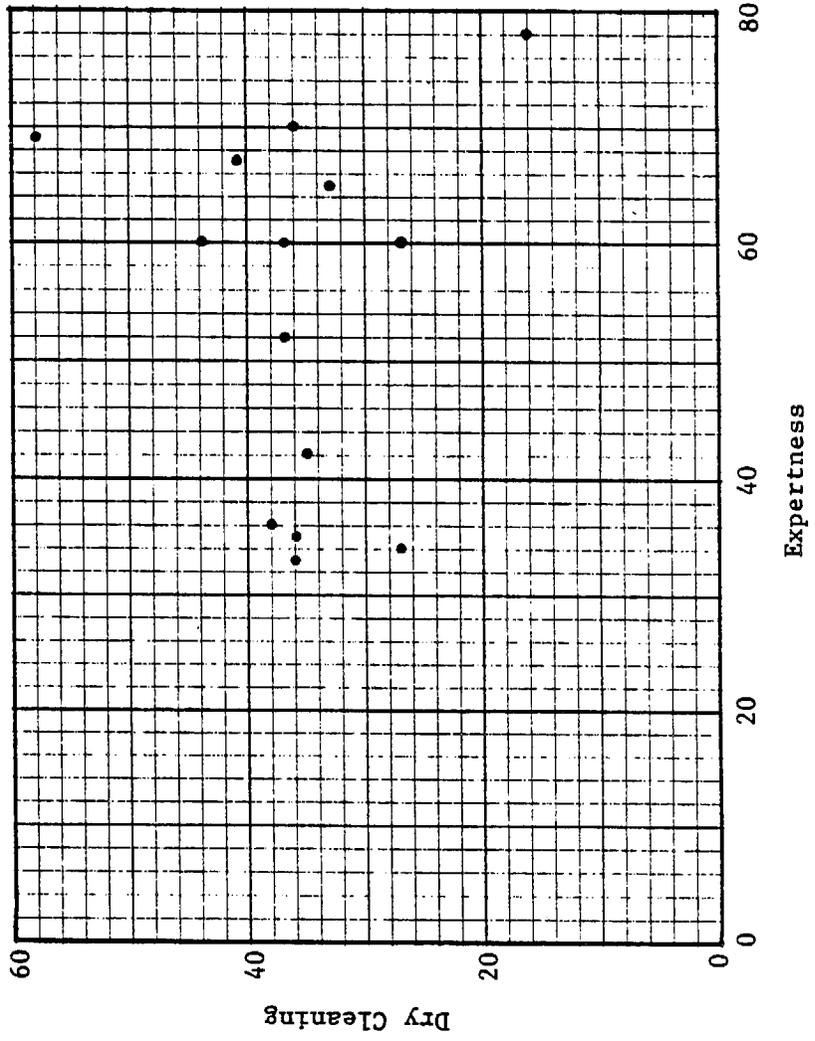


Figure 3-2. Association Between Scores in Self-Evaluated Expertness and Dry Cleaning Example

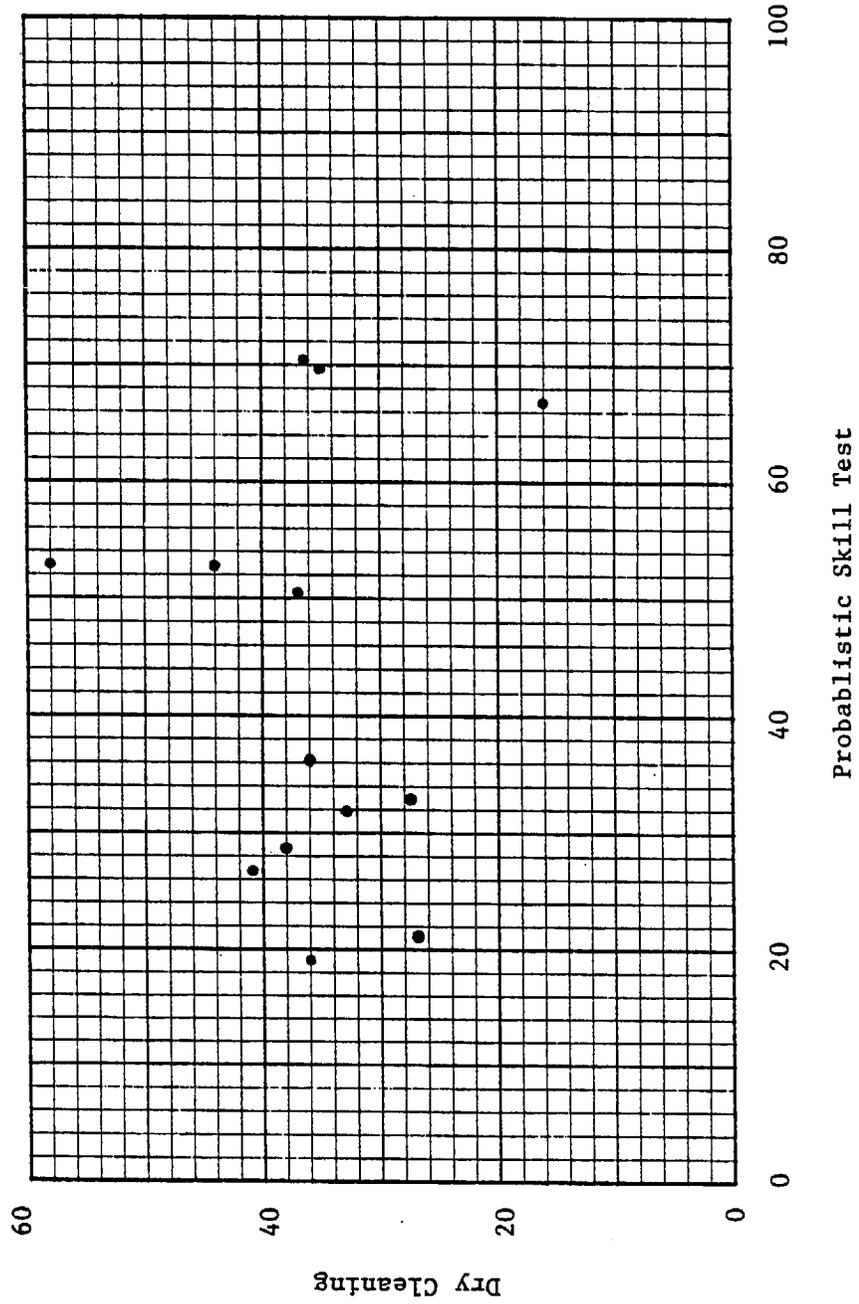


Figure 3-3. Association Between Scores in Probabilistic Skill Test and Dry Cleaning Example

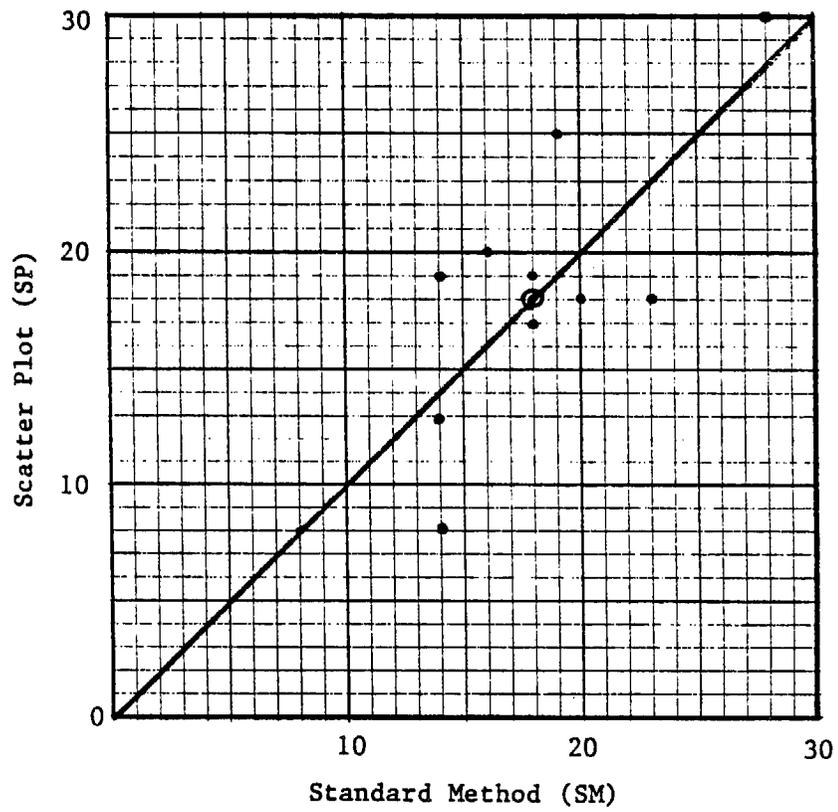


Figure 3-4. Association Between Scores in Standard Method and Scatter Plot Method

project period from the lessons learned from the failures of the NAPAP method (EPA 1986), and the SM and SP methods. The Handbook Method is summarized as follows:

- Step 1 An inventory analyst evaluates a promulgated emission estimate and provides two alternative estimates: upper plausible estimate and lower plausible estimate.
- Step 1 An inventory analyst evaluates a promulgated emission estimate and provides two alternative estimates: upper plausible estimate and lower plausible estimate.
- Step 2 Based on the three estimates, the analyst computes upper and lower threshold levels (i.e., UL and LL) such that UL is one standard deviation above the mean of the three estimates and LL is one standard deviation below the mean.
- Step 3 Evaluators are asked to estimate odds that the true mean emission is below UL and odds that the true mean is below LL.
- Step 4 Based on the evaluator's responses, subjectively evaluated precision and bias uncertainties in the promulgated emission estimate are then graphically determined using normal and lognormal probability papers.

Unlike SM, the Handbook Method does not require an evaluator to undertake a betting game involving a conceptual urn filled with red and blue balls. In HM, an evaluator is given a concise, clear narrative describing rationale and reasons for each of the three estimates so that he has a sounder base to think about the odds that the true mean is found below UL and LL. Unlike SP, HM does not require an evaluator to speculate about imaginary data points around the emission estimate in order to attach his perceived uncertainty measure to the estimate.

As to the five attributes of Ideal Methodology, HM appears to possess all five attributes. On Attribute 5, HM requires a preparation of the narrative rationalizing upper and lower plausible estimates as well as the promulgated emission estimate. However, this requirement is imposed upon an inventory analyst, not upon an evaluator. In HM, an evaluator is asked to estimate the likelihood of finding the true mean below UL and LL in terms of odds only.

4.0 RECOMMENDATION FOR FUTURE STUDIES

Although the Handbook Method appears to be reasonable, it is an unproven method, as stated in the uncertainty handbook (Horie 1988). Since HM is so new and unique, many additional efforts will be required before it can be routinely applied to uncertainty studies. This section discusses three work areas which must be explored prior to a wide dissemination of HM:

- o Procedural Manual
- o Tutorial Material
- o Model Inventory

Each of these work areas is discussed in the subsections that follow.

4.1 DEVELOPMENT OF PROCEDURAL MANUAL

A complete characterization of the uncertainties in an emission inventory requires a substantial organized effort and a set procedure to follow. Although the uncertainty handbook discussed many of the technical aspects of uncertainty assessment, it did very little on procedural and administrative aspects of uncertainty assessment efforts. The latter must be addressed separately in the so-called "Procedural Manual" for uncertainty assessment efforts.

If ARB intends to disseminate HM to the ARB's Inventory Divisions and local air quality management districts, a comprehensive procedural manual must be prepared (either internally or externally through its procurement process). This procedural manual should include guidance on and illustrations of:

1. A format for the narrative describing the method, assumptions and approximations employed in deriving the promulgated emission estimate, and the reasons of and rationale for upper and lower plausible estimates;
2. A format for reporting the procedures used for the results of assessed uncertainties in the promulgated emission estimates;
3. Procedures for developing upper and lower plausible estimates and implementing assessment efforts for both subjectively evaluated and objectively calculated uncertainties of the promulgated emission estimates; and

4. An organizational structure of the team that would be best suited to carry out an uncertainty assessment project.

The first two subject areas can be dealt with rather easily by preparing a manual similar to the manual (ARB 1982) entitled "Methods for Assessing Area Source Emissions in California". In this manual, each area source category is described concisely on the following items:

- o Source Number and Description
- o Method and Sources
- o Assumptions
- o Comments and Recommendations
- o Changes in Methodology
- o Temporal Activity
- o Sample Calculations
- o References
- o Preparer's Name(s)
- o Date of Last Update

The third subject area is incorporated into the Handbook (Horie 1988) with neither adequate explanation nor any standardized procedure. In the Handbook, each of the three example categories is described in terms of promulgated emission estimate and the two alternative estimates, with some discussion of reasons. These descriptions were prepared to merely work out the HM examples. The procedures used for estimating the upper and lower plausible estimates and for preparing the uncertainty questionnaire (for subjective assessment) can be standardized into a few variations.

The fourth subject area was not discussed much in the Handbook because of a complete lack of working experience. However, a specific organizational structure seems to be required in order to avoid unnecessary conflicts among the project members and (possibly) between the uncertainty project team effort and other on-going inventory efforts in the same agency.

4.2 DEVELOPMENT OF TUTORIAL MATERIAL

The uncertainty handbook (Horie 1988) is a sort of tutorial material for uncertainty assessment. However, the Handbook may be too technically involved for some inventory analysts with little statistical background because it is written so as to fully describe and rationalize the HM, which has been developed entirely de novo in the Handbook.

In working out the three emission examples presented in the Handbook, the authors noted that some evaluators had difficulty in representing the uncertainties of their promulgated emission estimates on a probability scale. Even with this limited experience, it is evident that these evaluator's responses will be more credible if they are given some tutorial material on uncertainties, probabilities and odds prior to their evaluations.

This tutorial material should be limited to what will be useful for thinking about the uncertainty on a probabilistic scale. Relationships between odds and probability and those between probabilities and points on normal and lognormal probability papers must be explained in an easy-to-understand manner using various examples and pictorial illustrations of the relationships.

4.3 UNCERTAINTY STUDY ON SAMPLE INVENTORY

As stated in the uncertainty handbook, HM is a plausible, yet unproven method. Therefore, if HM is to be used widely, a pilot study to prove the viability of HM must be carried out prior to public dissemination of the method. As with any new method, HM will encounter numerous difficulties and require modification to cope with particular situations that arise in an existing emission inventory system. These modifications and revisions of HM can not be worked out without identifying the potential problem and limitations that HM may experience in actual applications.

To enhance HM into a proven, practical method, HM must be applied to an actual emission inventory system on a trial basis. The scope of this pilot study might range from a limited test on a small, simple inventory of a rural air pollution control district to an extensive test on a comprehensive emission inventory like the one being developed for the South Coast Air Quality Study (STI 1986).

In either case, the pilot study will provide an opportunity to work out not only the revisions and modifications that may be deemed to be necessary for HM, but also the procedural manual and tutorial material through actual applications of HM to the inventory system.

5.0 REFERENCES

ARB, "Methods for Assessing Area Source Emissions in California", Stationary Source Control Division, Air Resources Board, Sacramento, CA, December 1982.

Chun, K. C., "Uncertainty Data Base for Emissions Estimation Parameters: Interim Report," Argonne National Laboratory, Prepared for the U.S. Department of Energy, Office of Fossil Energy, September 1986.

EPA, "Estimation of Uncertainty for the 1980 NAPAP Emissions Inventory", EPA-600/7-86-055, U.S. Environmental Protection Agency, Air and Energy Research Laboratory, Research Triangle Park, NC, December 1986.

EPA, "Compilation of Air Pollutant Emission Factors, Volume I: Stationary Point and Area Sources", AP-42 Fourth Edition, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, September 1985.

EPRI, "Emissions Inventory for the SURE Region", EA-1913 Research Project 862-5, Prepared by GCA Corp. for the Electric Power Research Institute, Palo Alto, CA, April 1981.

Horie, Y., and A. Shrope, "Development of Procedures for Establishing the Uncertainties of Emission Estimates", Interim Report to ARB Contract No. A5-184-32, Prepared by Valley Research Corporation, Van Nuys, CA, November 1986.

Horie, Y., "Handbook on Procedures for Establishing the Uncertainties of Emission Estimates," ARB Contract No. A5-184-32, Valley Research Corporation, Van Nuys, CA, February 1988.

Mangat, T. S., L. H. Robinson and P. Switzer, "Medians and Percentiles for Emission Inventory Totals", Paper No. 84-72.3, Presented at the 77th APCA Annual Meeting, San Francisco, CA, June 1984.

SCAQMD, "Uncertainty of 1979 Emissions Data, Chapter IV: 1983 Air Quality Management Plan Revision, Appendix IV-A; 1979 Emissions Inventory for the South Coast Air Basin, Revised October 1982", Planning Division of the South Coast Air Quality Management District, El Monte, CA, October 1982.

STI, "Southern California Air Quality Study (SCAQS): Suggested Program Plan", STI Ref. 10-95050-605, Prepared for California Air Resources Board, Sonoma Technology, Inc., Santa Rosa, CA, June 1986.

APPENDIX

SCATTER PLOT METHOD QUESTIONNAIRES

- o Expertness Evaluation
- o Probabilistic Skill Test
- o Area Source Emissions

SURVEY ON SUBJECTIVE
DETERMINATION OF EMISSIONS UNCERTAINTY

Name: _____ Telephone: _____
Title: _____
Company/Agency: _____
Division/Department: _____

EXPERTNESS EVALUATION

1. My knowledge of air pollution in general may be rated in the scale of 1 (least) to 10 (best) at:
(1 2 3 4 5 6 7 8 9 10)
2. My knowledge of emission inventories in general may be rated at:
(1 2 3 4 5 6 7 8 9 10)
3. My experience in emission inventories in general may be rated at:
(1 2 3 4 5 6 7 8 9 10)
4. My knowledge and experience in mobile source emissions may be rated at:
(1 2 3 4 5 6 7 8 9 10)
5. My knowledge and experience in stationary source emissions may be rated at:
(1 2 3 4 5 6 7 8 9 10)
6. My knowledge and experience in VOC emissions from dry cleaning operations may be rated at:
(1 2 3 4 5 6 7 8 9 10)
7. My knowledge and experience in NO_x and SO₂ emissions from power plants may be rated at:
(1 2 3 4 5 6 7 8 9 10)
8. My knowledge and experience in NO_x and VOC emissions from motor vehicles may be rated at:
(1 2 3 4 5 6 7 8 9 10)

PROBABILISTIC SKILL TEST

Problem 1. A seven-year old boy got a thermometer as his birthday present. He hung it outside of his room window and recorded a temperature around 8 a.m. every morning for about two weeks.

Q1a. Among the six scatter plots shown in Figure 1, please pick the one that seems to best represent the temperature data:

(1 2 3 4 5 6)

Q1b. In the scatter plot you selected, please indicate a pair of lines between which 90 percent of the data points appear to lie by circling two alphabets:

(a, b, c, d, m, p, q, r, s, O)

Q1c. Knowing that the thermometer was hung at the upper edge of the window, please indicate a line that best represents the true mean temperature by circling an appropriate alphabet:

(a, b, c, d, m, p, q, r, s, O)

Problem 2. An air pollution control district conducted a survey of NO_x emissions from industrial boilers in the district. A team of source testing engineers conducted a series of five tests to cover the range of operating conditions at each surveyed facility. Based on the test results, an emission factor was determined for each facility.

Q2a. Among the six scatter plots shown in Figure 1, please pick the one that seems to best represent the scatter of emission factors for the surveyed boilers:

(1 2 3 4 5 6)

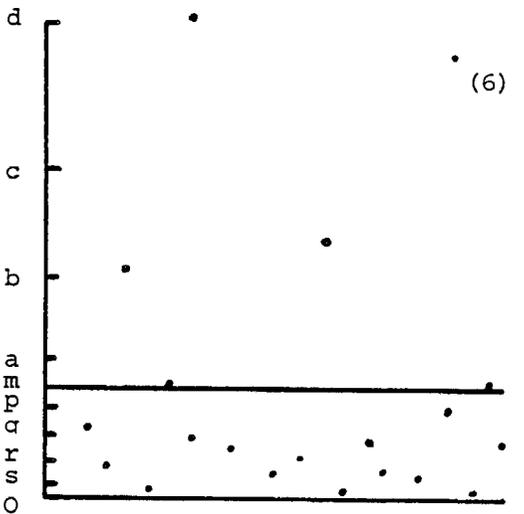
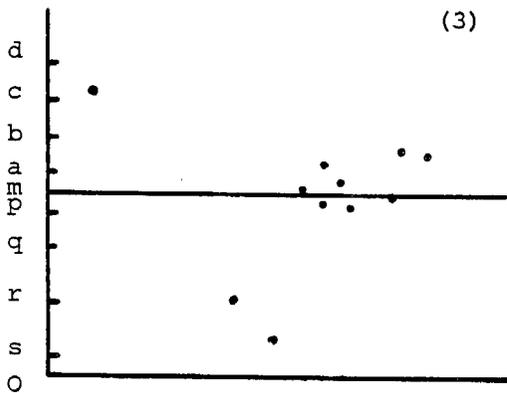
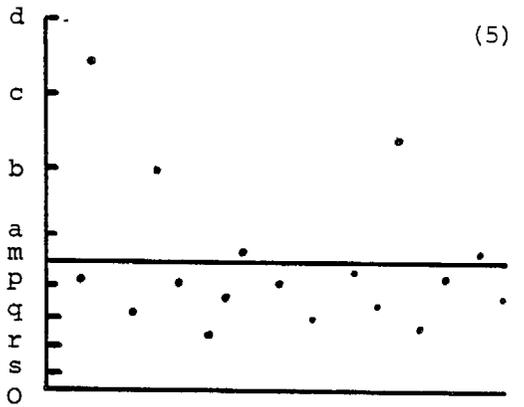
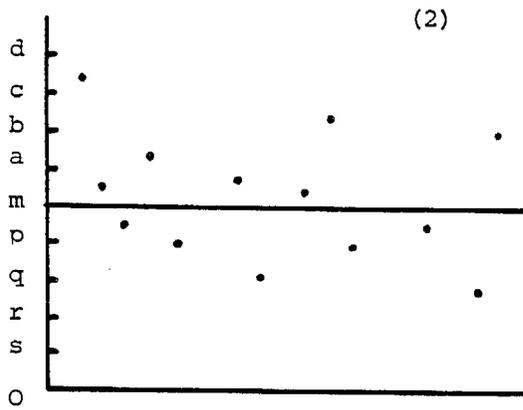
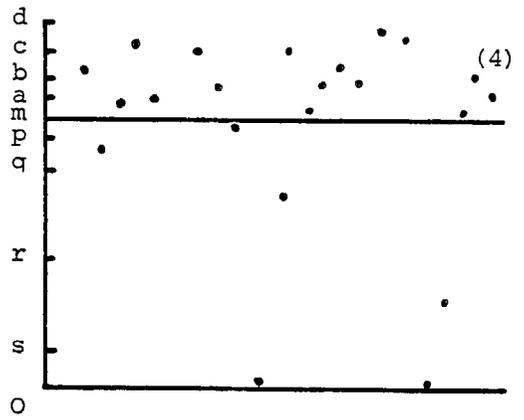
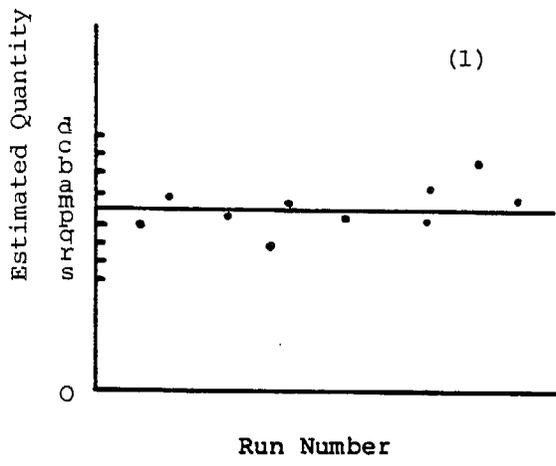


Figure 1. Sample Scatter Plots
(m = sample mean, O = zero)

PROBABILISTIC SKILL TEST (Cont'd)

Q2b. Knowing that the survey oversampled larger boilers and conversely undersampled smaller boilers, please indicate a line that best represents the population mean for all boilers in the district by circling an alphabet:

(a, b, c, d, m, p, q, r, s, O)

Q2c. In the scatter plot you selected, please indicate a smallest interval, in which the true emission factor for all boilers is expected to be found with 95 percent confidence, by circling two alphabets:

(a, b, c, d, m, p, q, r, s, O)

Problem 3. Weekly fuel burning rates for a base-load power plant were recorded by reading the gauge attached to the fuel-oil tank at the plant site. This power plant was shut down periodically for regular maintenance check.

Q3a. Among the six scatter plots shown in Figure 1, please pick the one that seems to best represent the fuel burning data:

(1 2 3 4 5 6)

Q3b. In the scatter plot you selected, please pick a line that appears to represent the median of weekly fuel burning rates at this plant.

(a, b, c, d, m, p, q, r, s, O)

Q3c. Assuming that the operating conditions during the recorded period were representative of this power plant, please indicate a 95 confidence interval for this plant's fuel burning rate by circling two alphabets:

(a, b, c, d, m, p, q, r, s, O)

PROBABILISTIC SKILL TEST (Cont'd)

Problem 4. In order to estimate a distribution of emission factors for all motor vehicles in the state, ARB procured a representative sample of in-use vehicles and tested their exhaust gas emissions according to the Federal Test Procedure.

Q4a. Among the six scatter plots shown in Figure 1, please pick the one that seems to best represent the scatter of exhaust emissions from the procured vehicles:

(1 2 3 4 5 6)

Q4b. In the scatter plot you selected, please pick a line that appears to represent the 50th percentile of exhaust emissions among all procured vehicles:

(a, b, c, d, m, p, q, r, s, O)

Q4c. Although the Smog Check (or I/M) Program detected that nearly 5 percent of vehicles were tampered on their emission control systems, none of the procured vehicles were tampered. With this fact, please indicate an interval in which the true emission factor is expected to be found with 80 percent confidence by circling two alphabets:

(a, b, c, d, m, p, q, r, s, O)

Problem 5. An environmental engineer conducted a literature review to come up with a tentative estimate of the emission factor for off-loading cement at the port facility in the district. He found emission factors for coal, fertilizer and pelletized sulfur, but not for cement itself.

PROBABILISTIC SKILL TEST (Cont'd)

Q5a. Among the scatter plots shown in Figure 1, please pick the one that seems to best represent the emission factors which the engineer found in his literature review:

(1 2 3 4 5 6)

Q5b. After checking a few handbooks for chemical engineers, minerals and other materials, the engineer made his best judgment that cement powders behave most similarly to fertilizer for which he found two reported emission values. Please indicate an emission factor that you would use by circling an alphabet:

(a, b, c, d, m, p, q, r, s, O)

Q5c. Please indicate an interval in which the true emission factor for cement off-loading seems to be found in nine out of ten chances by circling two alphabets:

(a, b, c, d, m, p, q, r, s, O)

AREA SOURCE EMISSIONS

A district engineer was assigned to estimate VOC emissions from dry cleaning operations in the district. Since emissions from large industrial dry cleaners were already estimated in the districts major point source file, only those emissions from commercial and coin-operated dry cleaners were to be estimated.

He first estimated the emission using the AP-42 emission factors, namely 1.3 lb /yr /capita for commercial and 0.4 lb/yr/capita for coin-operated. The emission factor rating was B in the scale of A through E, with A being the best (see Appendix A). In this estimate, he picked 100 telephone numbers from the telephone directories in the district and conducted a mini-questionnaire survey on residents' usage of commercial and coin-operated dry cleaners. Results of the survey were:

Commercial only	35
Commercial and coin-operated	10
Coin-operated only	25
No answer	30
TOTAL	100

Based on the survey results, he assumed that 60% of the resident population were for commercial dry cleaners and 40% for coin-operated. The engineer calculated the emission using the county population of 827,000 as follows:

$$E = (1.3 \times 0.60 + 0.4 \times 0.40) \times 827,000/2,000 \\ = 387 \text{ tons per year}$$

Then, he found a recent ARB publication (Rogozen 1985) in which the following empirical equation was given:

$$E = \exp [4.362 + 0.00808 (P - 193.85)] \quad (1)$$

where E is annual emission in tons and P is the county's urban population in thousands. By substituting the county's urban population of 627,000 into this equation, he computed a total amount of VOC emissions from both commercial and coin-operated dry cleaners in the district.

AREA SOURCE EMISSIONS (Cont'd)

$$E = 2,596 \text{ tons per year}$$

The engineer was surprised to see the huge discrepancy between the two estimation methods. Later, he found that Equation (1) was applicable only up to 400,000 urban residents in a county. From the ARB publication, he prepared Table 1 and Figure 2. In the table, per capita dry cleaning emissions are computed for each of the eight counties in the San Joaquin Valley from which Equation (1) was derived. By taking an arithmetic average of these per capita emissions, he derived a new emission factor of 1.2 lb/yr/capita. Using the new emission factor, the engineer obtained another emission estimate as:

$$\begin{aligned} E &= 1.2 \times 627,000/2,000 \\ &= 376 \text{ tons per year} \end{aligned}$$

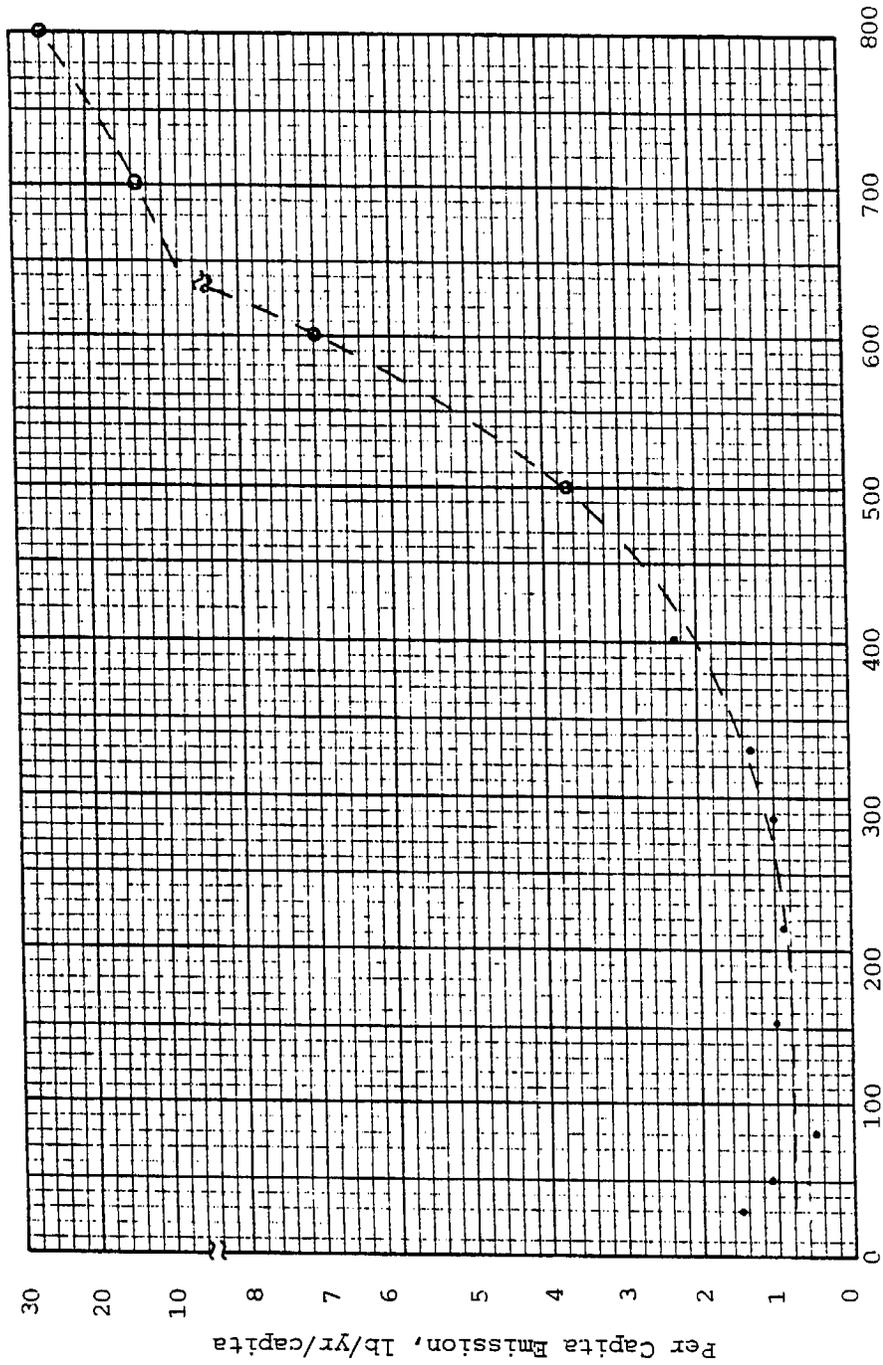
Standard Method (SM)

Q1. The reported emissions from area source dry cleaners are $m=387$ tons per year. After carefully considering all factors that may affect the emission value, please pick an interval in which the true emission level is expected to lie with greatest certainty among the following intervals:

1. 1.0m to 1.6m
2. 0.9m to 1.5m
3. 0.8m to 1.4m
4. 0.7m to 1.3m
5. 0.6m to 1.2m
6. 0.5m to 1.1m
7. 0.4m to 1.0m

Table 1. Dry Cleaning Emissions in the San Joaquin Valley
(after Rogozen 1985)

County	1979 VOC Emissions (ton/yr)	1980 Urban Population (1,000)	Per Capita Emission (lb/yr/capita)
Fresno	468	403.1	2.3
Kern	212	330.5	1.3
Kings	28	48.9	1.1
Madera	22	30.1	1.5
Merced	22	83.8	0.5
San Joaquin	150	286.0	1.0
Stanislaus	97	215.2	0.9
Tulare	75	153.2	1.0
Mean			1.20
Standard Deviation			0.53



1980 Urban Population in 1000

Figure 2. Per capita Dry Cleaning Emissions in the San Joaquin Valley

AREA SOURCE EMISSIONS (Cont'd)

Q2. Within the interval you selected above, how confident are you to find the true emission level?

1. with 99.9 percent (1 out of 1000)
2. with 99.0 percent (1 out of 100)
3. with 95 percent (1 out of 20)
4. with 90 percent (1 out of 10)
5. with 80 percent (1 out of 5)
6. with 50 percent (1 out of 2)
7. with 20 percent (less than 50/50)

Q3. Imagine a pin-wheel that is a disk with a colored sector that can be varied from 0 to 360 degrees. Suppose you set the colored sector equal to the confidence level you selected above (e.g., 90 % $0.90 \times 360 = 324$ degrees). Then, is the chance that a spinner would land on the colored portion the same as the confidence probability that you attached to in Q2?

(YES NO)

If NO is answered, then go back to Q2 and correct your confidence level that leads to YES in Q3. Your correct choice is:

(1 2 3 4 5 6 7)

Scatter Plot Method (SP)

Q4. The reported emissions from area source dry cleaners are $m=387$ tons per year. After carefully considering all factors that may affect the emission value, please imagine a scatter of estimates of the emission and pick one of the six scatter plots shown in Figure 1 that best matches the scatter of estimates you mentally imagined:

(1 2 3 4 5 6)

AREA SOURCE EMISSIONS (Cont'd)

- Q5. In the scatter plot you selected, please indicate an interval in which you expect to find the true emission level in one out of ten chances (i.e., with 90 percent confidence) by circling two alphabets:

(a, b, c, d, m, p, q, r, s, O)

- Q6. Please make your best estimate for the upper boundary of the 90 percent confidence interval you selected. My estimated upper confidence level is:

() times the original estimate, m.

- Q7. Please make your best estimate for the lower boundary of the 90 percent confidence interval you selected. My estimated lower confidence level is:

() times the original estimate, m.

APPENDIX A

AP-42 EMISSION FACTOR RATINGS

To help users understand the reliability and accuracy of AP-42 emission factors, each Table (and sometimes individual factors within a Table) is given a rating (A through E, with A being the best) which reflects the quality and the amount of data on which the factors are based. In general, factors based on many observations or on more widely accepted test procedures are assigned higher rankings. For instance, an emission factor based on ten or more source tests on different plants would likely get an A rating, if all tests were conducted using a single valid reference measurement method or equivalent techniques. Conversely, a factor based on a single observation of questionable quality, or one extrapolated from another factor for a similar process, would probably be labeled D or E. Several subjective schemes have been used in the past to assign these ratings, depending upon data availability, source characteristics, etc. Because these ratings are subjective and take no account of the inherent scatter among the data used to calculate factors, they should be used only as approximations, to infer error bounds or confidence intervals about each emission factor. At most, a rating should be considered an indicator of the accuracy and precision of a given factor used to estimate emissions from a large number of sources. This indicator will largely reflect the professional judgement of the authors and reviewers of AP-42 Sections concerning the reliability of any estimates derived with these factors.