Developing Improved Emission Factors and Assessment of Uncertainties (or Filling the Holes in Swiss Cheese)

James H. Southerland
Ronald Myers
Emission Inventory Branch (MD-14)
U.S. Environmental Protection Agency
Research Triangle Park, North Carolina 27711

Dennis Wallace
Midwest Research Institute
401 Harrison Oaks Boulevard
Cary, NC 27513

ABSTRACT

This paper reports on the current emission factor development process and examines some of the uncertainties in emission factors and how they affect the estimation of emissions used for various purposes. One use of emission factors is, of course, in area wide emission inventories. Emission factors may be quite useful and adequate for this purpose. However, emission factors are often used under exacting circumstances which may result in their application in a manner whereby precise control limits are developed, permits are granted, fees are levied or emissions traded among sources. This paper builds upon these needs and delineates what is being done to prioritize and improve the emission factors now in existence in the U.S. Environmental Protection Agency's (EPA) data bases and documents, such as AP-42, Compilation of Air Pollutant Emission Factors, which has been the major reference on the topic for about 30 years. Gradually, the "holes" in the data bases are being plugged. This paper raises questions as to the necessary certainty of the data and what the requirements will be to fill all the major gaps.

*This paper is the product of the authors and does not represent approved policy of he United States Environmental Protection Agency.

INTRODUCTION

The concept of emission factors and emission inventories for area wide assessment of emissions of various air pollutants has been in use for 30 years or more. EPA's Compilation of Air Pollutant Emission Factors, AP-42, originally a United States Public Health Service document, was a pioneering vehicle in the extension of this concept.[1] Today, the Fifth Edition of AP-42 has been conceived and is gestating in the halls of the Emission Inventory Branch of EPA. It has grown from a small collection of papers in the mid-sixties, addressing only a hand full of source categories and a very limited number of pollutants, to an approximate length today of nearly 2000 pages, and almost 200 major source categories. Many of these categories were not even conceived of as air pollution sources in those beginning years of its history. In spite of warnings on the limitations of the emission factors, the application of emission factors has changed even more. Uses of emission factors that would not have been conceivable a few years ago are now commonplace. What was originally an attempt to estimate long term emissions in an wide area in

order to develop and evaluate strategies to abate their emissions, has now changed to one where emission estimates are used in exacting photo-chemical and short-term models, precise emission controls are mandated and enforced, emissions are tracked within sensitivity levels of 3 percent per year, fees are charged based on estimates, and emissions are even traded in an economic market of real dollars and cents.

The ever expanding scope of AP-42 emission factor applications in terms of both the sources of interest and the ways in which the emission factors are being used has placed a great demand on EPA to develop more and better emission factors. We are attempting to meet this challenge through a two-pronged iterative process. First, in response to the need for more emission factors, we are revising AP-42 sections with the best data that we can find. At the same time, we are developing tools that we can use to assess the quality and uncertainty of these new emission factors and are applying those tools with the data collected during the emission factor development effort to characterize the uncertainty of these new emission factors. These uncertainty assessments will provide emission factor users with additional information on appropriate uses and limitations of published emission factors. They also feed into the iterative process by helping identify source categories where further improvement in emission factors is needed and in helping identify the most appropriate strategies for collecting data for making these improvements. The data from these future studies can then feed into improving the uncertainty assessments.

EMISSION FACTOR DEVELOPMENT/IMPROVEMENT

Current AP-42 Emission Factor Development

It is also important to proceed in a logical manner to improve these important tools to make them as useful and certain as is technically possible within the time and resource constraints that exist. This goal is made even more important by the Clean Air Act Amendments of 1990. Section 130 of the Amended Act requires EPA to review and revise existing VOC, CO and NO_x emission factors every three years.[2] November 1993 marks the conclusion of the first of these three year periods, after an initial 6 month review and revision effort that was mandated by the Act. The first six month review concluded with Supplement D to AP-42.[3] Although this effort also produced a prioritization of needs for further update,it was instrumental in prompting a realization that an in depth approach to improving individual high priority emission factors could not readily be undertaken without a cover to cover update of the document to reflect information, data and other changes that were determined to be needed. Consequently, a 3 year effort was undertaken to accomplish this update (not limited to the pollutants above but also including PM-10, SO_2, CO_2, Pb and Hazardous and other speciated air pollutants), much of which is represented in the 282 page Supplement E published in October 1992 and the 628 page Supplement F to AP-42 which went to print in August of 1993.[4,5] The remainder of the fruits of this update should be completed in late 1993 or early 1994 and then published in a new Fifth Edition. Thereafter, a more deliberate and prioritized effort to address gaps (holes in the cheese) for the source categories which are of highest concern can be executed and errors and uncertainties can be reduced.

Plans for Continued Emission Factor Improvement

An underlying philosophy of improving emission factor quality is reflected in the current updates being completed on AP-42. However, we expect to focus more in the future on individual source

categories and to place more effort in reducing errors and uncertainties for the "more important" source categories. To this end, we have developed a rating scheme and priorities which will allow us to better identify what categories are most important, and go after them with more fervor. In these efforts, we must pay more attention to statistical approaches and to the documentation of the variables and judgements involved.

First, we have to be more careful in selecting and characterizing the population of data that we use for factor development. We currently use all the data that we can get and afford. We search EPA files, scour state data files, plead with industries and their trade associations for useful data, and develop and run testing programs with our own funds. However, unless we have sufficient data to characterize the whole population (which we never will), we can not be sure just how good our emission factors and estimates really are or in how much source to source variation there may be across the industry. Existing EPA data tend to be limited as to the pollutants evaluated and measured values may be biased due to the focus of the specific sponsoring program office; industry may or may not be willing to "find" and give us data, depending on how they think they will be affected by the results; state agencies often have data but mostly limited to compliance tests, which many times have limited documentation, may generate emission estimates that are biased relative to "typical" conditions and require the state to expend scarce resources to provide; and lastly, it takes considerable time, money and project engineering time to locate the sources which best represent the "typical" facility, conduct tests and develop new data points. On top of that, we ideally need sufficient tests to allow us to determine the variabilities within each of the populations. One informal rule of thumb says we need at least 30 data sets. Our current data base for combustion sources, for criteria pollutants, comes closest to being an adequate data base, with about 50 to 100 tests in some categories. Most other source categories have much more limited data; perhaps 2-3, but almost always less than 10 data sets to work with. These data sets are almost always less than would be needed to do any parametric studies that are warranted in many complex processes to define the important process variables and to estimate the effects of those variables on emission factors.

Our program is focused on selecting source categories which are important for one or more pollutants, either on a national emission basis, or of importance in non-attainment areas as a major source category. To put these efforts into some logical order, we have developed an emission factor needs prioritization scheme. This scheme reflects a National need focus, but also has other attributes. This scheme, for example, also takes into account the tendency of some source categories to be "politically sensitive," we and may tend to try to collect data on them disproportionate to their "numerical" rating. We often are influenced by the "squeaky wheels." For example, welding and abrasive blasting are not typically a major area wide problem, but they are the source of much localized concern, prompting many calls to our engineering staff. In response to these repeated inquiries, the rating scheme allowed us to move these categories higher on the priority list. As a result, over the last two years, we have conducted tests, obtained worker exposure studies and developed draft sections for these two source categories which will eventually be incorporated into AP-42.

So how are we getting the data to fill these holes in the Swiss cheese of emission factors?

EPA files and literature data- usually well documented data, but many times of specialized nature. (eg. Tests may be limited in that they only characterize the emissions which are controlled by the "Best

Available Technology" and as a result may generate emission factors that are biased low relative to the full industry spectrum.

State data- In an effort to overcome the reluctance of states to use their scarce resources to provide us with emission tests located in their files we have instituted a project to "mine" their project files using our resources. We have collected over 2000 screened source tests for various source categories in the last year and a half. These data are probably limited/biased in many cases due to the purpose and conditions at time of test, but mining of these data is expected to produce many improvements and statistical insights into true values and variabilities.

Trade Association data- In increasing numbers, trade associations are undertaking test programs and proposing to work with EPA to develop better data bases and emission estimation tools, so that their members can have a better basis for dealing with air agencies, so that they will have an opportunity to be a part of work which may likely influence permit fees, and to provide better information for permit applications. The challenge of the EPA emission factor group is to utilize and steer these resources/efforts so that truly representative data are generated, and that results are real and fair to the industry's interest and to the environmental and regulatory considerations.

EPA Test Program- For years, resources available were totally inadequate to undertake a testing program for emission factor development. In FY '92 this drought of funds seemed to be over, but returned in the budget cutting era of FY '93. Funds are still available for a few well selected tests and piggyback opportunities on tests of others. Careful selection of sources, test methods and general objectives of this program are paramount to lending insights into and elimination of the uncertainties in emission factors.

Once we have obtained all available data from the above sources, we follow a standardized model to develop consistent and defensible rated emission factors.[6] In the model, 1) the source is characterized using engineering judgement with respect to variations in process characteristics, potential pollutants and emission controls; 2) the test data are rated in comparison to established EPA reference or recommended test methods; 3) the data to be used are are segmented and aggregated based upon the assumed source characteristics; 4) final selection of test data follows a general guideline; 5) average emission factors are developed and 5) quality ratings are assigned based upon subjective evaluations using an establised methodology. Though this model will not eliminate uncertainty and error in the resulting factors, it does help insure that they are reduced and can be reproduced and more easily and efficiently improved in the future. Let's look at this process in more detail.

EMISSION FACTOR UNCERTAINTY CHARACTERIZATION

Sources of Uncertainty in Emission Factors

As the use of AP-42 emission factors has expanded in such areas as permit development and air quality modeling, concerns for the quality, or uncertainty of these emission factors have also increased. Characterizing emission factor concepts within a statistical framework is often a useful tool for uncertainty assessment. Simply stated, an emission factor is an estimate of some true population parameter (such as emissions/activity level). When an emission factor is used to develop an annual emission inventory for a large area, the population parameter of interest is the mean unit emission rate for

the population of sources within the inventory area. When an emission factor is used to estimate VOC emissions from a single source for an 8-hour ozone model, the parameter being estimated is the mean unit emission rate for that source for that 8-hour period. In the above examples a "unit emission rate" is the emission rate per unit of process operation.

Because our knowledge of the behavior of the "population" that generates the emissions is crude and incomplete, there is uncertainty associated with calculated emission factors. This uncertainty is related to the error with which the emission factor estimates this true parameter of interest, and the error is function of both the bias of the calculated emission factor relative to the population parameter and the imprecision of the estimate. The bias component, the uncertainty of an emission factor, is a direct function of the underlying parameter that the emission factor is estimating. The imprecision component of the uncertainty is a function of the variability of the members of the population to which the emission factor is applied. Furthermore, both bias and imprecision are related to the data base used to develop the emission factor, the uses of the emission factor and the procedures used to develop the emission factor.

Information provided by emission factor users indicates that the uses are quite varied on both temporal and spatial scales. Relative to spatial scale, emission factor applications fall into three general categories. For the first category, individual facility or source specific emission estimates are used in the analyses, but decisions based on the results are for emissions from multiple facilities aggregated in a specified area. Examples of applications that fall into this category are state emission inventories, NAAQS equivalency determinations, and model inputs for photochemical grid modeling or PSD demonstrations. The pollutants of greatest concern for this category appear to be VOC, NO_x, and PM-10. For the second category, emissions estimates are developed for specific sources or facilities, and decisions are made on a facility basis. Permit limits, permit fee determinations, and innovative regulatory strategy development (including marketable permits) are the primary examples for this category, and essentially all pollutants are of concern. The third category includes those policy applications that are not particularly concerned with plant specific emissions, but with emission levels from broad source categories. Examples include lead strategy development, air toxic control programs, and RACT equivalency studies. Overlaying these areas is the temporal scale to which the emission factor is applied, which may range from an hour or less to annual.

In spite of the diversity of uses of emission factors, and their many inherent flaws, source category average emission factors are often the "best available" means to estimate emissions. If an emission factor is used to estimate emissions for a relatively large geographic and long-term scale, the uncertainty of that estimate is partially a function of the bias and imprecision of the emission factor. However, if the emission factor is used to estimate facility-specific emissions on a shorter time scale, the uncertainty is also a function of the inherent facility-to-facility and temporal variability of the "true emission factors" for specific facilities and time periods as well as the bias and imprecision of the estimator itself. Hence, sources of uncertainty can be divided into two major categories--those related to the bias and imprecision of emission factors as estimators of true source category emissions and those related to the inherent facility to facility and temporal variability of the "true emissions."

Sources of uncertainty associated with the second category are primarily a function of process and control device characteristics that vary between facilities and at the same facility over time. Examples include use of different process equipment; process size or process rate differences that affect per unit

emitting potential; process temperature differences; variations in raw materials and fuels; differences in air pollution control device operating characteristics such as scrubber pressure drop, ESP specific collection area, and carbon absorber regeneration frequency; and facility housekeeping, operation and maintenance practices. The uncertainties associated with these sources cannot be reduced by improved emission factor development methodologies; at best, the sources can identified (for specific source categories), and the degree of uncertainty can be estimated. On the other hand, sources of uncertainty associated with the first category are related to the emission factor development process and to the data used to develop the emission factor estimates. The degree of uncertainty associated with these sources can be reduced by identifying these sources and improving the emission factor development process.

Four features of the AP-42 emission factor development process are noteworthy for this discussion. First, as a part of the emission factor development process, the source category is characterized and process variations are identified. This component provides one mechanism for characterizing some of the inherent process variability discussed in the paragraph above. Second, the input emission test data for the analysis are not based on a random sample from the population. Rather, they are a convenience sample from a variety of sources including EPA files, state agency files, and industry submittals. As such, they may represent a biased sample from the population. Third, the data are subjected to a careful evaluation and a rating process is exercised whereby inferior quality data are discarded prior to emission factor development. This selection process can affect emission factor uncertainty in several ways. If some component of the source category is selectively eliminated in the process, the emission factor will thus be biased. For example, if particulate matter results from tests are discarded because they include both condensible and filterable fractions that cannot be separated, and if raw material characteristics from sources represented by that data subset result in different emission factors for those sources, these emission factors are likely to be biased. The selection process can also have counterbalancing effects on emission factor precision. By discarding lower quality data (presumably data that have a higher degree of uncertainty or imprecision) the precision of the estimates may be improved. On the other hand, eliminating a part of the data set reduces the sample size, which in turn increases the imprecision (standard error) of the estimate of the mean emission factor. Finally, the decision about how to segment or aggregate processes within a source category may have mixed effects on uncertainty. If different process types within a category (e.g., wet process and dry process portland cement kilns) have "true emission factors" that are different, using an average calculated emission factor produces a biased estimate for both process types. However, if two emission factors are developed, less data are available to generate each factor, so the estimates are less precise. Hence, segmentation or aggregation of an industry typically involves a trade-off between the bias and precision components of uncertainty. More explicit characterization of each of these sources of uncertainty for a particular emission source category during the development process can lead to reduced uncertainty in the final estimates.

Approaches for Quantifying Uncertainty

Interviews with emission factor users, as a part of an ongoing study of emission factor uncertainty, suggest that uncertainty characterization should address four primary issues:[7]

- Identify potential sources of variability at least qualitatively and to the degree possible, incorporate that variability quantitatively in the average emission factor;

- Characterize potential biases of average published emission factor relative to "population average emission factor";
- Estimate the imprecision (or estimator variability) of the published emission factor as an estimator of the population average emission factor; and
- Characterize the between source variability in source-specific emission factors relative to the population average and the within-facility and overall temporal variability in emission factors.

These factors can be addressed through a combination of engineering and statistical analyses. First, an engineering assessment of each industrial process is conducted as a part of the background study for each AP-42 section. This assessment, coupled with preliminary analyses of available emission data, provides identification of emission sources within an industry segment, characterization of the pollutants and controls for those emission sources, an assessment of the process parameters that are likely to affect emissions, and an evaluation of the relative merits of expressing the emission factors in estimating equation or in a ratio format (i.e., in terms of average mass of emissions per unit of process activity). Further, this preliminary analysis provides some information as to whether different process alternatives for the same operation should be handled separately or aggregated for emission factor purposes. This engineering analysis addresses the first of the four uncertainty issues identified above.

The remaining three issues can be addressed through a variety of statistical techniques. A very general statistical model that incorporates these different uncertainty concerns, which is based on a modification of the model suggested by Benkovitz, et. al., is:

$$EF_{ij} = EF_{SC} + EF_{Sub} + \theta_i + \epsilon_j \qquad \text{(Eq. 1)}$$

where

EF_{ij} = Source-specific emission factor generated from an emission test at facility i at some point in time j

EF_{SC} = True emission factor for a particular source category

EF_{Sub} = True deviation to the emission factor for a particular subcategory of the source category

θ_i = Deviation in the average emission factor for source i from the subcategory "population" average

ϵ_j = Deviation (or error) of a particular emission factor measurement (from a specific test or run) from the average emission factor for a specific facility

This statistical model is broad enough to cover most emission factor scenarios in that it can have the following properties:

-- Both EF_{SC} and EF_{Sub} can be expressed as a simple ratio measure or as linear or nonlinear functions of operating variables or environmental parameters denoted by $X_1, X_2, \ldots, X_n$.

-- θ_i and ϵ_j are random variables with particular probability distributions. (Typically, the mean of these distributions is assumed to be 0. Also, the distributions are often assumed to be normal, although other continuous distributions can be used)

-- ϵ_j can be decomposed further into more explicitly defined error terms related to such factors as temporal variability and measurement error as appropriate for a particular emission factor.

The objectives of the statistical analyses are to characterize possible biases in the estimation of EF_{SC} and EF_{Sub}, to characterize the precision of these two estimates, and to characterize the distributions (particularly the variances) of θ_i and ϵ_j. Because the true population parameters are never known, the bias of the estimates cannot be determined explicitly. However, as noted in the previous section, the primary sources of bias are associated with the data collection and selection processes. An indication of the potential magnitude can be obtained by subset analyses of the data based on data source and data quality. Measures of the uncertainty associated with the imprecision of the population estimators and of the variability of θ_i and ϵ_j can be obtained through a variety of statistical techniques. Some generic statistical tools that can be used for these analyses include "Linear Models" for mean estimation, mean confidence intervals and bias characterization; "Variance Component Analysis" for partitioning variability; "Resampling Techniques" for distributional characterization and "Time Series Models" for population temporal distributions from CEM's. The appropriate tools for a particular emission factor are selected based on the characteristics of the source and of the available data.

Preliminary Analysis Results

As a part of the ongoing study of uncertainties in AP-42 emission factors, prototype uncertainty analyses are being conducted for five source categories that represent a range of types of source, pollutant, and emission estimating method. This section presents some preliminary results from those analyses for two sources, portland cement kilns and fixed-roof (FR) storage tanks. These two sources were selected because they represent a range of pollutants (PM, SO_2, and NO_x for portland cement plants and VOC for FR tanks) and because they represent two types of estimating procedures. The emission factors for portland cement plants are traditional ratio emission factors, while the emission factors for FR tanks are in the form of estimating equations. The paragraphs below briefly summarize the analysis methods and some of the preliminary results for these two sources.[8,9]

A substantial quantity of emission data is available for PM, SO_2, and NO_x emissions from four types of portland cement kilns (wet process, dry process, preheater, and preheater/precalciner) and three levels of air pollution control (uncontrolled, fabric filter [FF] controlled, and electrostatic precipitator

TABLE 1
Characteristics of the Portland Cement Kiln Emission Data

Kiln Type	APCD	No. of Facilities			No. of Test Runs		
		NO_x	SO_2	PM	NO_x	SO_2	PM
Wet	None	1	2	3	8	9	13
Wet	FF	1	1	1	2	4	3
Wet	ESP	7	10	13	70	63	85
Dry	None	0	0	0	0	0	0
Dry	FF	5	3	4	93	11	10
Dry	ESP	1	2	4	2	14	24
Preheater	None	0	0	1	0	0	4
Preheater	FF	5	6	9	45	48	55
Preheater	ESP	1	0	1	5	0	4
Precalciner	None	0	0	0	0	0	0
Precalciner	FF	3	2	4	27	24	27
Precalciner	ESP	1	2	1	3	6	3

[ESP] controlled). Table 1 shows the number of facilities tested and the number of test runs for each of these pollutants as a function of kiln type and air pollution control device (APCD). The numbers in the table represent only A, B, and C rated test data. The uncertainty analyses for these data were conducted in three stages. First, exploratory descriptive analyses and analysis of variance modeling was conducted to assess the most appropriate level of kiln/APCD aggregation using all available data. Next, possible biases associated with data collection and selection methods were examined. The analyses considered both data quality biases and data source biases. Finally, linear models techniques and variance components procedures were used to assess uncertainties associated with the final emission factor estimates.

Preliminary analyses of the available data indicated that different aggregation patterns were appropriate for the three pollutants. For NO_x, no meaningful differences were noted between controlled and uncontrolled emissions. Some differences were noted between kiln types, but the differences were relatively small and not totally consistent with the pattern expected based on engineering principles. Also, a more precise estimate was generated by aggregating across kiln types, so a single emission factor was developed for all kilns. For SO_2, the emission factors were averaged across kiln types, but separate factors were developed for uncontrolled, FF controlled and ESP controlled kilns. For PM, separate factors were developed for uncontrolled and FF or ESP controlled kilns averaged across kiln types. Using these aggregation patterns, replicate analyses were conducted using only A and B rated data. Then additional replicate analyses were conducted by sequentially deleting data generated from EPA regulatory development tests, EPA or industry sponsored research tests, compliance tests, and other types of tests. Although these replicate analyses did result in some changes in the emission factor estimates, the changes were generally small (30 percent or less) and they exhibited no consistent pattern of magnitude or direction. Taken together, these analyses suggested no pattern of bias in the C-rated data or in the data

developed from different sources. Consequently, data from all sources with A, B, and C ratings were used to develop the final emission factors.

TABLE 2
Summary of Preliminary Portland Cement Kiln Results

Pollutant	Scenario	Emission Factor (lb/ton clinker)		Uncertainty Estimate[b]			
				Between Facility		Within Facility	
		Mean	Confidence Interval[a]	Variance	Std. Dev.	Variance	Std. Dev.
NO_x	All	5.6	(5.3,5.9)	13.4	3.7	1.9	1.4
SO_2	No Control	13.3	(11.0,15.6)	88.7	9.4	72.0	8.5
SO_2	FF	2.8	(1.9,3.7)	19.3	4.4	1.5	1.2
SO_2	ESP	7.2	(6.3,8.1)	50.4	7.1	18.0	4.2
PM	No Control	217	(210,224)	10,500	102	3020	55
PM	FF/ESP	0.48	(<0,2.9)	0.24	0.49	0.16	0.40

[a] 95% confidence interval based on normal approximation for mean estimate.
[b] Obtained via variance components analysis with restricted maximum likelihood estimates

Table 2 presents the final emission factor estimates and 95 percent confidence intervals for estimates as estimators of the true mean emission factor. Recall that the general form of the emission factor model for an individual source is:

$$EF_{ij} = EF_{SC} + EF_{Sub} + \theta_i + \epsilon_j$$

where for this example EF_{SC} and EF_{Sub} are combined into a single emission factor. The characteristics of θ_i and ϵ_j were examined via variance components analyses based on facility-to-facility differences in emissions and run-to-run differences within facilities. Evaluation of these data indicated that they tend to be right skewed and more heavily tailed than normally distributed data. Consequently, confidence intervals based on the normal distribution for facility-to-facility differences and temporal differences within facilities could not be developed. However, the variance (and standard deviation) estimates for the between facility and within facility distributions presented in Table 2 indicate that these sources of uncertainty are much greater than the uncertainty in the estimate of the mean. These data will be analyzed further using bootstrap techniques to develop confidence intervals for between and within facility uncertainty in the emission factors.

Emission estimates for FR storage tanks are generated via a series of estimating equations that are described in Section 12 of AP-42 and that can be implemented via the TANKS program that is available from EPA. These estimating equations, which are based on an analytical derivation from first principles rather than on empirical equations generated from test data, generate separate estimates for standing storage losses and working losses. The only data available for evaluating the uncertainty associated with these estimating equations were generated from tests conducted by EPA, the American Petroleum

Institute, and the Western Oil and Gas Association on standing storage losses from 30 operating storage tanks.

Preliminary analyses of the data suggested that separate analyses based on type of liquid stored (chemical product, crude oil, and fuel oil) were warranted. Also, the measure that appeared to best characterize the uncertainty associated with the estimating equation was the ratio of calculated to measured emissions. Table 3 presents the results of the uncertainty analyses based upon this measure. The table includes information on the range of measured emissions from the test programs that formed the basis for the analysis, the vapor pressure range that was tested for each product type, the estimate of the average ratio of calculated to measured emissions associated with the estimating equations, and 95 percent confidence intervals for the ratio as measures of the uncertainty with which the equations estimate the overall mean emissions and with which they estimate emissions for an individual facility. The data indicate that the estimating equations appear to be biased for both chemicals and fuel oil, although the bias is in opposite directions. The results also show a high level of uncertainty, particularly if the equations are used to estimate emissions for a particular tank. However, these results should be interpreted cautiously given the relatively small number of tests for each product class.

Conclusions

TABLE 3
Summary of Fixed Roof Tank Uncertainty Results

Product Stored	Crude Oil	Fuel Oil	Chemicals
Number of Tests	8	10	12
Measured Emission Estimate Range (lb/day)	80.3-578	0.008-0.050	3.41-45.3
Vapor Pressure Range (psia)	0.19-4.1	0.0050-0.012	0.21-1.9
Mean Calculated/Measured Ratio	1.3	0.88[a]	2.1[a]
95% Confidence Interval (Mean Ratio)[b]	(0.33,2.2)	(0.79,0.97)	(1.3,2.6)
95% Confidence Interval (Between Facility Ratio)[c]	(0.091,7.9)	(0.66,1.2)	(0.60,2.3)

[a] Significantly different from 1 at the 0.05 level

[b] Obtained by exponentiating the 95% confidence interval for the logarithm of the ratio. This confidence interval was based on a t-distribution with degrees of freedom 1 less than the number of facilities tested.

[c] Obtained from exponentiating the 2.5 and 97.5 percentiles of the distribution based on a normal approximation.

Emission factors, by their very existence as an attempt to provide a simplistic model of often extremely complex activities, are prone to many errors and uncertainties. Though we must often use them for purposes for which they are technically inappropriate, they are frequently the only way to make needed emission estimates and proceed with mandated and desired environmentally enhancing programs. Consequently, EPA has undertaken a renewed prioritization of activities needed to update and improve

emission factors in a manner that will have the most payoff in terms of decreases in uncertainty of individual source and inventory-wide emission estimates. Some of these activities will involve the collection of new test data through programs specifically designed to meet narrowly defined emission factor objectives. Other of these needs may be filled through collection and utilization of data from testing done by States and others for other reasons and the adaptation of these data to fill a void for which they were not originally collected. Vast improvements have been made and more are being made. Much remains to be done, but we must keep in mind that even at best, the uncertainties will not disappear and we will continue to need further improvements the more that we are called upon to provide factors to estimate emissions for permits or emission fees, or even to quantify emissions to a precision that will facilitate emissions being traded on the open market in terms of very precise dollars and cents.

1. *Compilation of Air Pollutant Emission Factors, Volume I: Stationary Point and Area Sources*, AP-42, 4th Edition, U. S. Environmental Protection Agency, Research Triangle Park, NC, September 1985.

2. *The Clean Air Act*, §130, 42 U.S.C. 7401-7626

3. *Compilation of Air Pollutant Emission Factors, Supplement D to Volume I: Stationary Point and Area Sources*, AP-42, U. S. Environmental Protection Agency, Research Triangle Park, NC, September, 1991.

4. *Compilation of Air Pollutant Emission Factors, Supplement E to Volume I: Stationary Point and Area Sources*, AP-42, U. S. Environmental Protection Agency, Research Triangle Park, NC, October, 1992.

5. *Compilation of Air Pollutant Emission Factors, Supplement F to Volume I: Stationary Point and Area Sources*, AP-42, U. S. Environmental Protection Agency, Research Triangle Park, NC, July, 1993.

6. J. Southerland, *et al.*, *Technical Procedures for Developing AP-42 Emission Factors and Preparing AP-42 Sections*, Office Of Air Quality Planning And Standards, U. S. Environmental Protection Agency, Research Triangle Park, NC, September 1993.

7. Written communication from Dennis Wallace, Midwest Research Institute, Cary, NC, to Ronald E. Myers, U. S. Environmental Protection Agency, Research Triangle Park, NC, June 23, 1993.

8. *Draft Emission Factor Documentation for AP-42 Section 8.6, Portland Cement Manufacturing*, Office Of Air Quality Planning And Standards, U. S. Environmental Protection Agency, Research Triangle Park, NC, July 30, 1993.

9. *Emission Factor Documentation for AP-42 Chapter 12, Storage Of Organic Liquids*, Office Of Air Quality Planning And Standards, U. S. Environmental Protection Agency, Research Triangle Park, NC, October 1992.