United States Environmental Protection Agency

Industrial Environmental Research Laboratory
Research Triangle Park NC 27711

EPA-600/2-78-195
September 1978

Research and Development

EPA

Statistical Analysis of Fugitive Emission Change Due to Refinery Expansion

RESEARCH REPORTING SERIES

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

1. Environmental Health Effects Research
2. Environmental Protection Technology
3. Ecological Research
4. Environmental Monitoring
5. Socioeconomic Environmental Studies
6. Scientific and Technical Assessment Reports (STAR)
7. Interagency Energy-Environment Research and Development
8. "Special" Reports
9. Miscellaneous Reports

This report has been assigned to the ENVIRONMENTAL PROTECTION TECHNOLOGY series. This series describes research performed to develop and demonstrate instrumentation, equipment, and methodology to repair or prevent environmental degradation from point and non-point sources of pollution. This work provides the new or improved technology required for the control and treatment of pollution sources to meet environmental quality standards.

REVIEW NOTICE

This report has been reviewed by the U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policy of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

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

EPA-600/2-78-195
September 1978

Statistical Analysis of Fugitive Emission Change Due to Refinery Expansion

by

Hugh J. Williamson and Lloyd P. Provost

Radian Corporation
P. O. Box 9948
Austin, Texas 78766

Contract No. 68-02-2608
Task No. 27
Program Element No. IAB604C

EPA Project Officer: Irvin A. Jefcoat

Industrial Environmental Research Laboratory
Office of Energy, Minerals, and Industry
Research Triangle Park, NC 27711

Prepared for

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Research and Development
Washington, DC 20460

DISCLAIMER

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

ACKNOWLEDGMENT

Consultations with Dr. Robert G. Wetherold of Radian Corporation have been very helpful in this research.

DCN 78-200-187-27-09

CONTENTS

TABLES

EXECUTIVE SUMMARY

The purpose of this report is to discuss a statistical approach for testing whether a planned petroleum refinery expansion can be made without increasing the fugitive emissions, when this is required by regulations. Emission factors can be used to estimate the emission increase or decrease due to the expansion. Since emission rates are empirically determined, however, they are subject to random sampling errors. Thus, the effect of the expansion on emissions cannot be computed exactly.

For this reason, the problem has been treated statistically. Analytical methods are presented which can be used to compute the mean and standard deviation of the emission change, whether positive or negative, due to the expansion. A method for computing the probability that the emissions will, in fact, not be increased by the expansion is also discussed, along with other related probability calculations.

The methods presented here can be used as an aid in comparing several options for reducing emissions to acceptable levels. This involves simply performing the statistical calculations in parallel for each option to be considered. In a given case, the options might include, for example:

- instigating an improved maintenance program for certain devices (e.g., valves),
- venting certain emissions to a flare, and
- shutting down a particular processing unit.

The methods are designed to handle different control strategies in the existing and in the planned facilities and different strategies for different types of devices, if this is necessary. The following situation, for example, could be handled.

- A new unit, such as a catalytic reformer, is to be added to an existing refinery.
- To reduce emissions in the existing facilities, a maintenance program is instituted for valves, and the API separator is to be covered.
- To reduce emissions in the new facility, the same type of maintenance program used in the old facility will be instituted for valves, and double seals rather than single seals will be used on all pumps. Single seals remain on pumps in the old facility, however.

Additionally an approach is presented for obtaining an estimate of the emission decrease for a particular type of device, such as valves, due to a particular emissions control program. The approach involves a paired-measurement experimental design which eliminates certain sources of extraneous influence and, therefore, increases the precision of the estimates obtained.

Although the discussion and examples in this report are specific to petroleum refineries, the methods developed and outlined here can be generalized to similar situations in other plant expansion problems.

SECTION 1

INTRODUCTION

The purpose of this report is to discuss a statistical approach for testing whether a planned refinery expansion can be carried out without increasing emissions, when this is required by regulations. The problem has been addressed statistically, since emission rates, being determined empirically, are subject to random errors. Thus, the decision as to whether a planned expansion is in compliance with regulations cannot be made with absolute certainty but can be made with a certain level of confidence. Probability calculations can be used to establish the level of confidence with which the judgement can be made in a given case.

The basic approach is first to determine, for each type of device in the refinery which independently contributes to the emissions, the emission rate from the new facility minus the emission reduction (if any) from the old facility due to improved maintenance practices or equipment modifications. This difference is the emissions increment for the device type being considered. The device types would probably include pipeline valves and flanges, pressure relief valves, pumps, compressors, cooling towers, etc. Any system of categorization which is convenient and physically sensible can be used.

Secondly, the emissions increments for all important device types are summed to get the overall increment for the refinery. If this increment is less than or equal to zero, this means that the expansion will "probably" not result in an emissions increase. If the increment is positive, the expansion probably will increase emissions.

These qualitative statements, however, are unsatisfactory due to their vagueness. Thus, an approach for calculating the numerical probability that emissions will not be increased is also presented in this report. The confidence with which a decision can be made regarding compliance with regulations, then, can be assessed.

An effort has been made to present the analysis in a self-contained form, so that an extensive statistical background is not required to follow the development. Thoroughness and generality, however, have not been sacrificed for the sake of simplicity. Some of the situations which could arise in a refinery expansion are complex. The equations necessary to evaluate the emission change due to the expansion in these cases, therefore, are also complex.

The statistical method presented here is illustrated with a number of numerical examples. The emission rates used in the examples are believed to be reasonable but are employed here for illustrative purposes only; they are in no way being presented as emission factors for use other than in this report.

The analysis presented herein also provides certain guidelines for experimental design for determination of emission factors. This is because the information needed to perform a refinery tradeoff study is listed, and the manner in which the information would be used to do the study is presented. These guidelines would be beneficial if existing estimates of emission rates were not appropriate for a particular tradeoff study, and, therefore, new estimates had to be developed.

Radian is currently performing an extensive research study for the Environmental Protection Agency, however, in which statistically valid estimates of the emission rates and their uncertainties will be obtained.

References 3 and 4 were very helpful in choosing a set of numerical test cases.

SECTION 2

AVERAGE EMISSION INCREMENT AND ITS UNCERTAINTY
(Given Type of Device, e.g., Valves)

In this section, the calculation of the estimated emission increment for a given device type is discussed. The standard deviation of this increment, which reflects random errors in the emission factors is also discussed. Four cases are discussed which include most trade-off situations anticipated.

The purpose of this section is not to present a rigorous mathematical treatment but rather to give, in an easily understandable way, the basic equations and the circumstances under which they should be used. The statistical background for this section involves only the basic properties of mean values, variances, and covariances. These properties are discussed briefly in Appendix A. Also see Reference 5.

Case 1- Same Provisions for Reducing Emissions in New and Old Facilities

In this case, it is assumed that control measures such as covering an API separator or improving maintenance programs are applied to all sources of a certain type. Another example is venting emission gases from all sources to a flare.

Define:

E^1 = average emission rate per source without improved maintenance,

S_{E^1} = standard deviation of E^1,

E^{11} = average emission rate per source with improved maintenance,

$S_{E^{11}}$ = standard deviation of E^{11},

N_o = number of sources of type being considered in the old facility, and

N_n = number of sources in the new facility.

Then the emission reduction in the old facility corresponding to this particular type of source is:

$$N_o(E^1 - E^{11})$$

and the emission increment in the new facility is

$$N_nE^{11}$$

Thus, the total emission increment I is

$$I = N_nE^{11} - N_o(E^1 - E^{11})$$
$$= (N_n + N_o)E^{11} - N_oE^1$$

The variance, S_I^2, of I is

$$S_I^2 = (N_n + N_o)^2 \; S_{E^{11}}^2 + N_o^2 S_{E^1}^2$$

It is important to remember that the emission factors E^1 and E^{11} correspond to a single device (e.g., a single valve), while I is the emission rate from all devices of a given type. A similar convention is used in the other cases discussed, although additional variables are introduced.

Example 1 for Case 1

The calculations for Case 1 will now be demonstrated with a numerical example. As is mentioned in the Introduction, the emission factors used in this and other examples are intended for illustrative purposes only. Statistically valid estimates of emission rates and their uncertainties will be calculated through another project currently being performed by Radian.

In this example, the emission increment for valves will be calculated in a hypothetical case in which a catalytic reformer is being added to an existing refinery. It is assumed that the reformer has 850 valves and that an improved maintenance program is introduced which reduces the average emission rate from each valve from 0.040 to 0.008 lb/hr. These and other necessary statistics are presented below.

E^1 = 0.040 lb/hr,

S_{E^1} = 0.006 lb/hr (estimated standard deviation of E^1),

E^{11} = 0.008 lb/hr,

$S_{E^{11}}$ = 0.003 lb/hr (estimated standard deviation of E^{11}),

N_o = 14,000 (number of valves in the refinery before the catalytic reformer was added), and

N_n = 850.

The estimated emission increment can be obtained by direct substitution of these values in the expression for I.

$$I = (N_n + N_o)E^{11} - N_o E^1$$

$$= (850 + 14000)\ 0.008 - 14000(0.040)$$

$$= -441$$

Thus, the addition of the catalytic reformer and the instigation of improved maintenance are estimated to decrease the total emissions from all valves by 441 lb/hr.

If the values of E^1 and E^{11} were exact, then the value of I given above would also be exact. Due to the uncertainty in the empirically determined emission rates E^1 and E^{11}, however, the value of I is also uncertain; its estimated standard deviation, S_I, is calculated as follows.

$$S_I^2 = (N_o + N_n)^2 S_{E^{11}}^2 + N_o^2 S_{E^1}^2$$

$$= (850 + 14000)^2(0.003)^2 + (14000)^2(0.006)^2$$

$$= 9041$$

$$S_I = \sqrt{S_I^2} = 95 \text{ lb/hr}$$

This concludes the calculations which would be necessary for valves alone. Similar calculations would also have to be performed for other device types, such as flanges, pumps and compressors. The values of I and S_I would then be combined to estimate the emission increment for the entire refinery and its uncertainty. This set of calculations is presented in the examples of this section and the next. The final calculations are presented in Example 1 in Section 3.

Example 2 for Case 1

In this example, we assume that the refinery's API separator is covered and that this reduces the estimated emission rate from the separator from 6.2 to 0.31 lb/1000 gallons of wastewater. This example is different from the preceding one in that no units of the type being considered are added; thus, N_n is zero. Covering the separator could be one of the steps to reduce emissions from existing facilities, so that additional processing equipment could be added without increasing emissions.

The required inputs are as follows:

$E^1 = 6.2,$

$S_{E1} = 2.9,$

$E^{11} = 0.31,$

$S_{E^{11}} = 0.15,$

$N_o = 1$, and

$N_n = 0.$

The quantities E^1, S_{E^1}, E^{11} and $S_{E^{11}}$ are in pounds per 1000 gallons of wastewater. Then:

$$
\begin{aligned}
I &= (N_n + N_o)E^{11} - N_oE^1 \\
&= (0 + 1)\ 0.31 - (1)(6.2) \\
&= -5.9
\end{aligned}
$$

Thus, covering the separator reduces the estimated hourly emissions by 5.9 lb/1000 gallons of wastewater. The standard deviation of this estimate, which is due almost completely to the uncertainty in E^1, is calculated as follows.

$$
\begin{aligned}
S_I^{\ 2} &= (N_n + N_o)^2\ S_{E^{11}}^{\ 2} + N_o^{\ 2}\ S_{E^1}^{\ 2} \\
&= (0 + 1)(0.15)^2 + (1)^2(2.9)^2 \\
&= 8.4 \\
S_I &= \sqrt{S_I^{\ 2}} = 2.9
\end{aligned}
$$

Further Discussion of Case 1

The case in which sources of a certain type are added in a new facility but reduced in number in the old facility is discussed in this subsection. This case would apply, for example, if a light ends/gas processing unit were to be added to an existing refinery and a reformer feed pretreating unit were to be shut down. This would result in the addition of a number of valves, for example, in the new facility and in the elimination of all valves which were in the reformer feed pretreating unit.

The formulation presented above must be modified slightly to handle this type of situation. Define:

N_c = number of sources of the type being discussed which are eliminated from the old facility.

All other notation is exactly as defined above. The emission reduction in the old facility due to the improvements in hardware or maintenance programs is:

$$(N_o - N_c)(E^1 - E^{11})$$

and the reduction in the old facility due to the elimination of N_c units is

$$N_c E^1$$

The added emissions in the new facility, as before, is:

$$N_n E^{11}$$

Thus, the total emission increment I is:

$$I = N_n E^{11} - (N_o - N_c)(E^1 - E^{11}) - N_c E^1$$
$$= (N_n + N_o - N_c)E^{11} - N_o E^1$$

and the variance $S_I{}^2$ of I is:

$$S_I{}^2 = (N_n + N_o - N_c)^2 \, S_{E^{11}}{}^2 + N_o{}^2 S_{E^1}{}^2$$

Example 3 for Case 1

The calculations for valves for the example mentioned above will now be performed. It will be assumed that putting in a light ends/gas processing unit adds 1300 valves and that eliminating the reformer feed pretreating unit removes 800 valves. These and other needed inputs are listed below.

E^1 = 0.040 lb/hr,

S_{E^1} = 0.006 lb/hr,

E^{11} = 0.008 lb/hr,

$S_{E^{11}}$ = 0.0012 lb/hr,

N_o = 14,000 valves,

N_n = 1300 valves, and

N_c = 800 valves.

As in Example 1 for Case 1, it has been assumed that an improved maintenance program is to be instituted which will reduce the average emissions from valves from 0.040 to 0.008 lb/hr.

$$I = (N_n + N_o - N_c)E^{11} - N_o E^1$$
$$= (1300 + 14000 - 800)\ 0.008 - (14000)(0.040)$$
$$= -444$$

Thus, the maintenance program together with the addition and removal of equipment results in a net decrease of 444 lb/hr in the estimated fugitive emissions from all valves. The standard deviation is computed as follows:

$$S_I{}^2 = (N_n + N_o - N_c)^2 S_{E^{11}}{}^2 + N_o{}^2 S_{E^1}{}^2$$

$$= (1300 + 14000 - 800)^2 (0.0012)^2 + (14000)^2 (0.006)^2$$

$$= 7359$$

$$S_I = 86$$

Example 4 for Case 1

This example is exactly like the preceding one, except that it is assumed that a light gas oil hydrotreating unit is also to be shut down, resulting in the elimination of an additional 800 valves. The only change to the required input quantities, then, is that N_c is increased from 800 to 1600.

Since the value of I obtained when only the reformer feed hydrotreating unit was eliminated is negative, it is not necessary to eliminate another unit to balance emissions from valves. It is possible, however, that a large decrease in emissions in valves would be necessary to balance an increase in emissions from another type of device, such as flanges.

The value of I, then, assuming elimination of the pretreatment and hydrotreating unit is:

$$I = (N_n + N_o - N_c)E^{11} - N_o E^1$$

$$= (1300 + 14000 - 1600)(0.008) - (14000)(0.040)$$

$$= -450 \text{ lb/hr}$$

Eliminating the additional hardware, then, changes the emission increment by only 6 lb/hr.

The standard deviation S_I is computed as follows.

$$S_I{}^2 = (N_n + N_o - N_c)^2 S_{E^{11}}{}^2 + N_o{}^2 S_{E^1}{}^2$$

$$= (1300 + 14000 - 1600)(0.0012)^2 + (14000)^2 (0.006)^2$$

$$= 7326$$

$$S_I = 86$$

This and the preceding example illustrate the use of the analysis methods presented herein as an aid in evaluating the various options for reducing fugitive emissions. The emission reductions achieved by equipment shutdowns and different maintenance programs can be estimated. The final decision, of course, should also include economic and other considerations, as well as reductions in emissions.

The analysis below for Cases 2 and 3 is somewhat more complicated than that for Case 1, and the number of required inputs is greater. Cases 2 and 3 cover situations in which emissions monitoring is performed, and only the high-leaking devices are repaired. In these cases, if the mean and standard deviation of the emissions from a collection of sources which are selectively maintained can be estimated, then the analysis under Case 1 should be used. The mean value would then play the role of E^{11}, and the standard deviation would play the role of $S_{E^{11}}$ in the notation used in Case 1.

The analysis below, however, can be used when the information in this exact form is not known. Although more numerous, the required inputs for Cases 2 and 3 may be more easily obtained than those for Case 1 in some instances.

Case 2 - Emissions Monitoring and Subsequent Corrective Action Performed the Same Way in the Old and New Facilities

An example of this case is the periodic checking of emissions from valves and performing maintenance if required. The possibility exists that sources with different ranges of emission rates would be maintained in a different way; for example, high leakers might be maintained immediately, but medium leakers might be tagged, rechecked periodically, and ultimately maintained only if the emissions exceeded a certain rate. This situation, however, reduces to the case in which units are maintained if they leak more than a certain amount.

Define:

F_1 = average emission rate of maintainable leakers if maintenance is not performed,

F_2 = average emission rate of maintainable leakers if maintenance is performed,

F_3 = average emission rate of other units,

S_{F_i} = standard deviation of F_i, i = 1, 2, 3,

f = proportion of the units of the type being considered which would be high leakers at any given time on the average, if detailed maintenance were not performed,

S_f = standard deviation of f,

N_o = total number of units of the type being considered in the old facility, and

N_n = total number in the new facility.

The emission rate F_2 is intended to apply to the more recently maintained units, while F_3 is intended to apply to the other units. More specifically, the emission rate F_2 applies to the fN_o most recently maintained units in the old facility and to the fN_n most recently maintained units in the new facility.

The emission reduction in the old facility then is:

$$fN_o(F_1 - F_2)$$

and the emission increment in the new facility is:

$$fN_nF_2 + (1-f)N_nF_3$$

The total emission increment is:

$$\begin{aligned} I &= fN_nF_2 + (1-f)N_nF_3 - fN_o(F_1 - F_2) \\ &= [N_nF_2 - N_nF_3 - N_oF_1 + N_oF_2]f + N_nF_3 \\ &= [(N_n + N_o)F_2 - N_nF_3 - N_oF_1]f + N_nF_3 \\ &= Af + N_nF_3 \end{aligned}$$

where A denotes the coefficient of f, which is enclosed in brackets in the expression above.

$$S_I^{\,2} = A^2S_f^{\,2} + f^2S_A^{\,2} + S_f^{\,2}S_A^{\,2} + (N_n)^2S_{F_3}^{\,2} + 2C$$

where $S_A^{\,2}$ is the variance of A, and C is the covariance between Af and N_nF. These two expressions are given below.

$$S_A^{\,2} = (N_n + N_o)^2S_{F_2}^{\,2} + (N_n)^2S_{F_3}^{\,2} + (N_o)^2S_{F_1}^{\,2}$$

$$C = -(N_n)^2fS_{F_3}^{\,2}$$

If it were assumed that the more recently maintained sources had the same mean and variance as the other sources, then the following equations would hold:

$$F_2 = F_3$$

$$S_{F_2}^{\,2} = S_{F_3}^{\,2}$$

The correlation between the estimated emission rates F_2 and F_3 would then be one, and this fact should be taken into account in the analysis. The expressions which should be used for I and S_I^2 are as follows:

$$I = [(N_n + N_o - N_n)F_2 - N_oF_1]f + N_nF_2$$

$$= N_o(F_2 - F_1)f + N_nF_2$$

$$= Bf + N_nF_2$$

where B denotes the coefficient of f.

$$S_I^2 = B^2S_f^2 + f^2 S_B^2 + S_f^2 S_B^2 + (N_n)^2 S_{F_2}^2 + 2C$$

where

$$S_B^2 = (N_o)^2(S_{F_2}^2 + S_{F_1}^2)$$

and

$$C = N_oN_nfS_{F_2}^2$$

<u>Example for Case 2</u>

To illustrate Case 2, we again assume that a catalytic reformer is being added and that an improved maintenance program is being instigated for valves, for which the values of I and S_I will be calculated. In this example, however, the available information is assumed to conform to Case 2 rather than Case 1.

Again, it is assumed that 850 valves are added. Further we assume that 12% of the valves would leak above the allowed rate if maintenance were not performed. These and other required inputs are as follows:

$F_1 = 0.267,$

$S_{F_1} = 0.035,$

$F_2 = 0.000,$

$S_{F_2} = 0.000,$

$F_3 = 0.0091,$

$S_{F_3} = 0.0045,$

$f = 0.120,$

$S_f = 0.010,$

$N_o = 14000,$ and

$N_n = 850.$

Then,

$$I = Af + N_nF_3$$

where

$$A = [(N_n + N_o)F_2 - N_nF_3 - N_oF_1]$$

$$= [(14850)(0) - (850)(0.0091) - (14000)(0.267)]$$

$$= -3745.74$$

Thus,

$$I = (-3745.74)(0.12) + (850)(0.0091)$$

$$I = -441$$

The standard deviation of I can then be calculated as follows.

$$S_I^{\ 2} = A^2S_f^{\ 2} + f^2S_A^{\ 2} + S_f^{\ 2}S_A^{\ 2} + N_n^{\ 2}S_{F_3}^{\ 2} + 2C$$

where

$$S_A^{\ 2} = (N_n + N_o)^2S_{F_2}^{\ 2} + N_n^{\ 2}S_{F_3}^{\ 2} + N_o^{\ 2}S_{F_1}^{\ 2}$$

$$= (14850)^2(0) + (850)^2(0.0045)^2 + (14000)^2(0.035)^2$$

$$= 240{,}115$$

$$C = -N_n^{\ 2}fS_{F_3}^{\ 2} = -(850)^2(0.120)(0.0045)^2$$

$$= -1.76$$

$$S_I^{\ 2} = (-3745.74)^2(0.010)^2 + (0.120)^2(240{,}115) + (0.010)^2(240{,}115) + (850)^2(0.0045)^2 + 2(-1.76)$$

$$= 4896$$

$$S_I = \sqrt{S_I^{\ 2}} = 70$$

This example was patterned after Example 1 for Case 1. The emission rates used in the two cases are consistent, and the resulting values of I are the same, -441. The values of S_I are different, however, due to the very different formulations and types of information used in calculating S_I in the two cases. The covariance structures are different in Cases 1 and 2, for example, and this affects S_I.

Case 3 - Same Hardware Provisions for Reducing Emissions in New and Old Facilities, Along With Emissions Monitoring and Subsequent Corrective Action Performed in Both Facilities

This case is, in a sense, a composite of Cases 1 and 2. The analysis presented here would apply in the case of pumps, for example, if the following conditions held:

- in the old facilities, single mechanical pump seals were replaced by double seals, resulting in a reduction in the average emissions,
- double mechanical pump seals were also used in the new facility, and
- in both facilities, emissions monitoring and subsequent corrective action were performed the same way.

The following terms will be needed:

E = average emission rate for hardware modifications, without detailed maintenance, and

S_E = standard deviation of E.

The F_i and f, defined below, apply to sources with hardware modifications:

F_1 = average emission rate of maintainable leakers if maintenance is not performed,

F_2 = average emission rate of maintainable leakers if maintenance is performed,

F_3 = average emission rate of other units,

f = fraction of sources of the type being considered which would be high-leakers at any given time on the average if detailed maintenance were not performed,

S_f = standard deviation of f,

N_o = total number of sources of the type being considered in the old facility, and

N_n = total number in the new facility.

The emission reduction in the old facility is:

$$N_oE - [fN_oF_2 + (1 - f)N_oF_3]$$

The emission increment in the new facility is:

$$fN_nF_2 + (1 - f)N_nF_3$$

The emission increment for the entire refinery for the type of device being considered is:

$$I = fN_nF_2 + (1 - f)N_nF_3 - [N_oE - fN_oF_2 - (1 - f)N_oF_3]$$

$$= f[(N_n + N_o)F_2 - (N_o + N_n)F_3] + (N_n + N_o)F_3 - N_oE$$

$$= Af + BF_3 - N_oE$$

where

$$A = (N_n + N_o)F_2 - (N_o + N_n)F_3 = (N_n + N_o)(F_2 - F_3)$$

and

$$B = (N_n + N_o)$$

The variance of I, then, is given by the following:

$$S_I{}^2 = A^2S_f{}^2 + f^2S_A{}^2 + S_f{}^2S_A{}^2 + B^2S_{F_3}{}^2 + 2C + N_o{}^2S_E{}^2$$

where $S_A{}^2$ is the variance of A, and C is the covariance of Af and BF_3. These two expressions are given below.

$$S_A{}^2 = (N_n + N_o)^2S_{F_2}{}^2 + (N_o + N_n)^2S_{F_3}{}^2$$

$$= (N_n + N_o)^2(S_{F_2}{}^2 + S_{F_3}{}^2)$$

$$C = -B^2fS_{F_3}{}^2$$

<u>Example for Case 3</u>

In this example, the emission increment for pump seals will be examined assuming that a catalytic reformer is being added, that single seals are being replaced by double seals on existing pumps, and that all pumps on the new equipment will have double seals. In addition to the hardware modifications, it is assumed that pump seals which are found to leak excessively (say, over one lb/hr) are replaced. The required inputs are listed below:

$E = 0.308$ lb/hr,

$S_E = 0.080$ lb/hr,

$F_1 = 1.500$ lb/hr,

$F_2 = 0.010$ lb/hr,

$F_3 = 0.015$ lb/hr,

$S_{F_1} = 0.400$ lb/hr,

$S_{F_2} = 0.005$ lb/hr,

$S_{F_3} = 0.006$ lb/hr,

$f = 0.130$,

$S_f = 0.020$,

$N_o = 264$, and

$N_n = 17$.

$$I = Af + BF_3 - N_oE$$

where

$$A = (N_n + N_o)(F_2 - F_3)$$

$$= 281\ (0.010 - 0.015) = -1.41$$

$$B = N_n + N_o = 281$$

Thus,

$$I = (-1.41)(0.130) + 281(0.015) - 264(0.308)$$

$$= -77$$

Thus, adding the catalytic reformer and replacing single seals with double seals results in a net decrease of 77 lb/hr in the estimated fugitive emissions from all pumps. The standard deviation of I can be calculated as follows:

$$S_I{}^2 = A^2S_f{}^2 + f^2S_A{}^2 + S_f{}^2S_A{}^2 + B^2S_{F_3}{}^2 + 2C + N_o{}^2S_E{}^2$$

where,

$$S_A^2 = (N_n + N_o)^2(S_{F_2}^{\ 2} + S_{F_3}^{\ 2})$$

$$= (281)^2[(0.005)^2 + (0.006)^2]$$

$$= 4.82$$

$$C = -BfS_{F_3}^{\ 2} = 281(0.130)(0.006)^2 = -0.001$$

Then,

$$S_I^2 = (-1.41)^2(0.020)^2 + (0.130)^2(4.82)^2$$
$$+(0.020)^2(4.82)^2 + (281)^2(0.006)^2 + 2(-0.001)$$
$$+ (264)^2\ (0.080)^2$$

$$= 449$$

$$S_I = \sqrt{S_I^{\ 2}} = 21$$

Case 4 - Entirely Different Emission Factors Apply for the New and Old Facilities

This case would apply if, for example, emissions from a particular type of device were vented to a flare in the new facility, but it was infeasible to introduce such a system in the old facility. The results for this case follow from the analyses for Cases 1, 2 and 3, with appropriate choices for the parameters.

Suppose, for example, that emissions monitoring with corrective action as required is to be performed in the old facility. Then the emissions increment I_{old} and its standard deviation $S_I{}^2{}_{old}$ are obtained by using the analysis for Case 2, with

$N_n = 0.$

Suppose further that hardware capabilities have been introduced to reduce fugitive emissions in the new facility. The emissions increment I_{new} and its standard deviation $S_I{}^2{}_{new}$ can be obtained from the analysis of Case 1, with

$N_o = 0.$

The total emissions increment corresponding to this particular type of device is:

$$I = I_{old} + I_{new}$$

and

$$S_I^2 = S_I^2{}_{old} + S_I^2{}_{new}$$

Example 1 for Case 4

In this example, it is once again assumed that a catalytic reformer is to be added to an existing refinery. The emission statistics for compressors are computed below.

Four compressors are assumed to be required for the new reformer, and these are all vented to a flare. No provisions for reducing emissions from existing compressors, which are not presently vented to a flare, are planned. Thus, different emission factors entirely apply in the old and new facilities, and this situation is handled under Case 4.

This is a very special case in which the emission increment and its standard deviation are both zero, as is shown below.

Old Facility--

It is clear that nothing has been done to change the compressor emissions in the existing facility. Thus,

$I_{old} = 0$, and

$S_{I_{old}} = 0$.

New Facility--

Since emissions are vented to a flare, the fugitive emissions are essentially reduced to zero. Thus,

$I_{new} = 0$, and

$S_{I_{new}} = 0$.

Both Facilities--

The emission increment for the entire facility and its uncertainty are readily seen to be zero:

$$I = I_{old} + I_{new} = 0$$

$$S_I^2 = S_I^2{}_{old} + S_I^2{}_{new} = 0$$

Example 2 for Case 4

In this case, we assume that a light ends/gas processing unit is to be added, resulting in the addition of 15 pumps. The single seals on the existing pumps are replaced by double seals. Hardware provisions, including venting to a flare, again essentially reduce the emissions from the new pumps to zero. The required calculations for pumps, then, are as follows.

New Facility--

As in the preceding example,

$I_{new} = 0$, and

$S_{I_{new}} = 0.$

Old Facility--

The calculations for the old facility can be made by using the Case 1 analysis and setting N_n equal to zero. The required inputs, then, are as follows:

$E^1 = 0.308,$

$S_{E^1} = 0.080,$

$E^{11} = 0.060,$

$S_{E^{11}} = 0.015,$

$N_o = 264$, and

$N_n = 0.$

Then,

$$\begin{aligned} I_{old} &= (N_n + N_o)E^{11} - N_oE^1 \\ &= (264)(0.060) - (264)(0.308) \\ &= -65 \end{aligned}$$

and,

$$\begin{aligned} S_I{}^2{}_{old} &= (N_o + N_n)^2 S_{E^{11}}{}^2 + N_o{}^2 S_{E^1}{}^2 \\ &= (264)^2(0.080)^2 + (264)^2(0.015)^2 \\ &= 462 \end{aligned}$$

$$S_{I_{old}} = \sqrt{S_I{}^2}_{old} = 21$$

Both Facilities--

The total emission increment for valves is:

$$I = I_{old} + I_{new} = -65$$

and its variance is

$$S_I{}^2 = S_I{}^2{}_{old} + S_I{}^2{}_{new} = 462$$

Thus, $S_I = 21$

Method for Reducing Error in Emissions Increments

A method designed to reduce the uncertainty in the emissions estimates is presented in Appendix B. The resulting analysis required to calculate I and S_I is more complex than in the cases presented above. The circumstances under which this method can be used advantageously and the analysis itself are discussed in detail in Appendix B.

The method does not lend itself to brief summarization without loss of accuracy. It should be considered, however, if the conditions discussed below hold.

Define:

x = measured emission rate for a particular source of the type being considered before a planned program for reducing emissions is put into effect, and

y = measured emission rate for the same source after the program is put into effect.

Then, if x and y are correlated (see Reference 5), the method discussed in Appendix B can be used to reduce the variance in the estimated value of I.

Other Cases

While every effort has been made to cover the most likely scenarios, it is almost inevitable that situations will arise which do not clearly fall under any of the cases discussed above. In these cases, the analysis necessary to estimate the emission increment and its uncertainty for the device being considered should be worked out. The analysis under "Further Discussion of Case 1," earlier in this section, is an example of how the basic equations can be extended to handle other cases.

It is probable, however, that the cases presented here will cover most, if not all, device types in a given refinery. The approach discussed in Section 3 for combining the I's and S_I's for all device types in the refinery, moreover, can be used in any case.

Systematic Errors

It is correctly pointed out in Reference 4 that, if present, biases in measurements could affect the estimated emission rates. If a known bias exists in measuring emissions, then the data will not be representative of the true emissions and therefore must be corrected before the analysis is performed. A quality-assurance program using specified calibration techniques and utilizing known standards for verification will be essential in determining the magnitude of the bias and the resulting correction factors needed. If quality-assurance procedures are not incorporated, then unknown systematic errors are best handled as random errors, as is discussed below.

If an emission factor were estimated from measurements at several refineries, and if each data collection involved an independent equipment setup and calibration, then one would expect that the data set would include a random collection of systematic or bias errors (if, in fact, biases were present). These errors would then be averaged, along with the other random variations, in obtaining the final estimated emission rate. If present, bias errors would increase the standard deviation of the emission rate.

Under these conditions, the analysis should be carried out exactly as described above; one should assume that the standard deviations of the emission rates reflect all measurement errors.

Suppose, on the other hand, that data were only available from a single equipment setup and calibration. If the variance from setup to setup for the particular type of measuring equipment used were known from experience, this variance should be added to the variance of the mean calculated directly from the data. The results, then, would represent the uncertainty due both to measurement-to-measurement variations and the random error due to setup and calibration.

If in collecting emissions data an unknown bias error existed which could not be measured or accounted for, then the resulting estimates would be erroneous. The standard deviations calculated would underestimate the actual uncertainty of the emission rates. There is always the potential for this type of occurrence in any measurement project. Proper calibration and quality-assurance standards will minimize this problem. If no attempt is made to measure these biases, then statistical procedures are of no help in deriving estimates of emissions.

SECTION 3

MEAN EMISSION INCREMENT AND ITS UNCERTAINTY
(Entire Refinery)

Now, suppose, that for each device type, the mean (or expected) emission increment, I, and standard deviation, S_I, have been computed. The emission increment I_T for the entire refinery is:

$$I_T = \Sigma I$$

and the variance of I_T is

$$S_{I_T}^{\,2} = \Sigma S_I^{\,2}$$

In both cases, the summation is over all device types.

Now, I_T is a linear combination of the various emission factors. Assuming that the most important emission factors were estimated from a large number of observations, these factors can reasonably be assumed approximately normally distributed. This statement does not imply that the original emissions data are normally distributed.

The quantity I_T then is a linear sum whose largest terms are approximately normal; thus, I_T is also approximately normal. The inclusion of the multiplicative random factor f, however, weakens the argument for normality somewhat. This objection would be insignificant, however, if f were estimated from a large sample and S_f^2 were small.

If I_T is assumed to be normal, and if the emission factors which contribute most to the standard deviation estimate S_{I_T} are calculated from samples of size 30 or greater, then the following variable Z can be assumed to have the standard normal distribution (i.e., Z is normally distributed and has mean zero and variance one)

$$Z = \frac{I_T - \mu}{S_{I_T}}$$

where μ is the true and unknown mean of the emission increment.

As is mentioned in the Introduction of this report, Radian is currently performing an extensive research study in which statistically valid estimates of emission rates and their uncertainties will be obtained. Due to the magnitude of the data collection being performed in this project, it is felt that the emission factors will be calculated from sample sizes large enough to satisfy the requirements for Z to have approximately the standard normal distribution. The standard normality assumption, moreover, greatly simplifies the probability calculations.

A few comments about cases in which this assumption is not valid, however, are in order. If I_T is approximately normal, but some of the important variance estimates are calculated from samples of size less than 30, then Z as defined is not normally distributed. Since $S_{I_T}^{\ 2}$ is the sum of a set of estimates of <u>different</u> variances, moreover, Z does not have exactly the t - distribution, either. It is possible that the distribution of Z could be approximated by a t - distribution, however. A similar application of the t - distribution is discussed in Section 4.14 of Reference 7. Cases in which I_T is not approximately normally distributed should probably be treated individually. Further analytical work regarding the distribution of Z will be beneficial if such cases arise in the future.

Assuming that Z has approximately the standard normal distribution, then, well-known statistical methods can be used to make probabilistic statements about the value of μ, as is discussed below. Of particular interest is the probability that μ is less than or equal to zero; that is, that the plant expansion can be made with no increase in emissions.

The probability that μ is less than some value a can be obtained as follows:

$$P(\mu<a) =$$

$$P(-a<-\mu) =$$

$$P\left(\frac{I_T - a}{S_{I_T}} < \frac{I_T - \mu}{S_{I_T}}\right) =$$

$$P\left(\frac{I_T - a}{S_{I_T}} < Z\right)$$

The last step follows from the definition of Z given above. The expression on the left of the inequality sign can be computed, given that I_T and S_{I_T} have already been calculated and the value of a has been chosen. The probability can then be looked up in normal probability tables, such as Table 2 of Appendix D, Reference 5.

As discussed above, it has been tacitly assumed that the sample sizes used to compute the emissions factors are large enough, say larger than 30, to justify using the Z-statistic instead of the t-statistic. The use of the Z-statistic would be valid as long as the emissions factors for the greatest contributors to emissions, such as valves, were computed from large samples.

It should be noted that the estimated emission increment I_T for the entire refinery can be calculated even if only the emission factors, but not their variances, are available. The variances are required only for the calculation of S_{I_T} and for the probability calculations.

The analysis method presented here can be used as an aid to assess several options for reducing the emissions to acceptable levels. The following two numerical examples illustrate this process. In Example 1, the calculations are performed assuming certain emission control procedures will be used. In Example 2, then, the control procedures are altered in several ways, and the calculations are updated accordingly. The final selection of control strategies, of course, should also include economic and other considerations, as well as the emission calculations.

Example 1

The calculations will now be worked out for a hypothetical example. It is assumed that a catalytic reformer is being added to an existing refinery and that various provisions to reduce emissions are being taken. The provisions and the values of I and S_I are listed in Table 3-1. The value of I and S_I shown for valves are calculated in Example 1 for Case 1; for the API separator in Example 2, Case 1; for pumps in the example for Case 3; and for compressors in Example 1 for Case 4.

The value of I and S_I for the API separator are -6.0 and 3.0, respectively. Both values are in pounds of emissions per thousand gallons of wastewater and must be converted to lb/hr. It is assumed that 300,000 gallons of wastewater are processed per hour; therefore, the conversion can be made by multiplying I and S_I by 300. The resulting values are:

$$I = (-6)(300) \cong -1800 \text{ lb/hr}$$

$$S_I = 3(300) \cong 900 \text{ lb/hr}$$

The values of I and S_I for flanges, drains, and relief valves must also be calculated. Since no change is made in the old facility for these three sources, the values of I_{old} and $S_{I_{old}}$ are both zero (see the discussion of Example 1 for Case 4). The statistics regarding the emission increment in the new facility, however, must be computed. The required inputs are shown in Table 3-2.

The new-facility calculations will be done as a Case 1 problem, with $N_O = 0$. Then in the case of flanges:

TABLE 3-1. DATA FOR EXAMPLE 1: HYPOTHETICAL TEST CASE IN WHICH A CATALYTIC REFORMER IS ADDED TO AN EXISTING REFINERY

Device	Steps to Reduce Emissions	I	S_I
Valves	Maintain high leakers	-441	95
API Separator	Cover	-6*	3*
Pumps	Replace single seals with double seals, then maintain high leakers	-77	21
Compressors	Vent emissions from new compressors to a flare	0	0
Flanges	None	2	1
Process Drains	None	2	2
Relief Valves	None	2	1

*In pounds per thousand gallons of wastewater. All other values of I and S_I are pounds per hour.

TABLE 3-2. DATA REQUIRED TO CALCULATE EMISSION-INCREMENT STATISTICS FOR NEW FACILITY FOR FLANGES, PROCESS DRAINS AND RELIEF VALVES

Quantity	Flanges	Process Drains	Relief Valves
E^{11}	0.00076	0.034	0.50
$S_{E^{11}}$	0.00050	0.025	0.20
N_o	0	0	0
N_n	2800	60	4

NOTE: The values of E^1 and S_{E^1} are not needed, since $N_o = 0$. E^{11} and $S_{E^{11}}$ are the emission rate and its standard deviation in lb/hr/unit with the programmed maintenance (none in these cases). All variable names are as defined under Case 1.

$$I = N_n E^{11} - N_o(E^1 - E^{11})$$

$$= N_n E^{11}$$

$$= (2800)(0.00076)$$

$$= 2.13$$

and

$$S_I{}^2 = (N_n + N_o)^2 \, S_{E11}{}^2 + N_o{}^2 S_{E1}{}^2$$

$$= N_n{}^2 S_{E11}{}^2$$

$$= (2800)^2 (0.0005)^2$$

$$= 1.96$$

$$S_I = \sqrt{S_I{}^2} = 1.40$$

The calculations for drains and relief valves are very similar. The values of I and S_I for all source types are given in Table 3-1, as is mentioned above.

The mean and standard deviation of the emission increment will now be calculated.

$$I_T = \Sigma I = -441 - 1800 - 77 + 0 + 2 + 2 + 2$$

$$= -2312$$

$$S_{I_T}{}^2 = (95)^2 + (900)^2 + (21)^2 + (0)^2 + (1)^2 + (2)^2 + (1)^2$$

$$= 819472$$

$$S_{I_T} = 905$$

Now we are interested in computing the probability that the true emission increment is less than zero:

$$P(\mu<0) =$$

$$P(0<-\mu) =$$

$$P\left(\frac{I_T}{S_{I_T}} < \frac{I_T - \mu}{S_{I_T}}\right) =$$

$$P\left(\frac{I_T}{S_{I_T}} < Z\right) =$$

$$P\left(\frac{-2312}{905} < Z\right) =$$

$$P(-2.55 < Z) = 0.9946$$

The probability 0.9946 was obtained from a normal probability table. The probability that the emissions are not increased, then, is extremely high.

This example requires one further comment. Any indirect effect of maintenance or hardware changes should be taken into account in establishing the emission rate estimtes to be used. Suppose, for example, that an equipment expansion (such as adding a catalytic reformer) resulted in a significant increase in fugitive emissions from either the API separator or the cooling towers due to an increased volume of water processed. Then the emission increase should be reflected in the calculations.

Example 2

As a further illustration, we will now alter the preceding example somewhat. Suppose the maintenance program for valves is not planned; that high leaking pumps are maintained, but that single seals are used in the new and old facility; and that the API separator is not covered.

The values of:

I = 34 lb/hr, and

S_I = 68 lb/hr

for valves are obtained exactly as were the values for flanges in the prececing example; in this case

$E^{11} = 0.040,$

$S_E{}^{11} = 0.080,$ and

$N_n = 850.$

The calculation of I and S_I for pumps falls under Case 1, with the following required inputs:

$E^1 = 0.308$ lb/hr,

$S_{E^1} = 0.080$ lb/hr,

$E^{11} = 0.060$ lb/hr,

$S_{E^{11}} = 0.015$ lb/hr,

$N_o = 264,$ and

$N_n = 17.$

Thus,

$$I = (N_n + N_o)E^{11} - N_oE^1 = -64$$

$$S_I{}^2 = (N_n + N_o)^2 S_{E^{11}}{}^2 + N_o{}^2 S_{E^1}{}^2 = 464$$

$$S_I = 22$$

Since no maintenance or hardware modifications are to be made regarding the API separator, the values of I and S_I for it become zero.

The estimated emission increment for the entire refinery is:

$$I_T = \Sigma I = 34 + 0 - 64 + 0 + 2 + 2 + 2$$

$$= -24$$

and

$$S_{I_T}{}^2 = \Sigma S_I{}^2 = (68)^2 + (0)^2 + (22)^2 + (0)^2 + (1)^2 + (2)^2 + (1)^2$$

$$= 5114$$

$$S_{I_T} = 72$$

The situation here is much less clear cut than it was in the preceding example. The estimated emission increment is negative, which means that the planned expansion probably would not increase the fugitive emissions. The probability that the true mean μ is actually negative can be calculated as follows:

$$P(\mu<0) =$$

$$P(0<-\mu) =$$

$$P\left(\frac{I_T}{S_{I_T}} < \frac{I_T - \mu}{S_{I_T}}\right) =$$

$$P\left(\frac{-24}{72} < Z\right) =$$

$$P(-0.33 < Z) = 0.63 \text{ (from the tabulated normal distribution)}$$

If it is desired, the probability that the true mean is less than a specified positive value can be computed by essentially the same method. The probability that μ is less than 50 lb/hr, for example, is:

$$P(\mu<50) =$$

$$P(-50<-\mu) =$$

$$P\left(\frac{I_T - 50}{S_{I_T}} < \frac{I_T - \mu}{S_{I_T}}\right) =$$

$$P(-1.03 < Z) = 0.85$$

Thus, there is only a 15% chance that the true emission increment is greater than 50 lb/hr.

SECTION 4

COMBINING DIFFERENT ESTIMATES OF THE SAME EMISSION RATE

In this section, the situation is discussed in which independent estimates of the same emission factor are available from:

(1) data collected at the refinery being investigated and

(2) other refineries (from which general emission factors have been computed).

The estimate from Source (1) may be more accurate in that random differences among refineries do not contribute to the error. Source (2), on the other hand, would be less subject to random measurement errors and random device-to-device differences if it were based on a larger data set, which would usually be the case.

The objective, then, is to determine whether: (a) the estimate from Source (1) only should be used, or (b) the estimates from the two sources should be combined. In the second case, the question is how to combine the estimates most effectively. Alternative (a) or (b) is selected by performing a statistical hypothesis test. Hypothesis tests are discussed in References 3 and 5 and in many other statistics works.

The analysis to be discussed in this section applies equally to any of the emission factors required for any of the cases discussed above.

Define the following variables:

E_1 = emission factor estimated from data collected at the refinery in question,

S_{E_1} = standard deviation of E_1,

E_2 = emission factor estimated from other refineries, and

S_{E_2} = standard deviation of E_2.

Note that S_{E_1} reflects only device-to-device variations and random measurement errors, while S_{E_2} reflects these variations <u>and</u> whatever refinery-to-refinery variations may exist.

Now unless a measurement malfunction invalidates E_1, it can be used to obtain the final estimate. Both E_1 and E_2 should be used if they can be considered independent estimates of the same quantity. If, on the other hand, the particular refinery being studied deviates significantly from the average, so that E_2 is actually a biased estimate of the emission factor of interest, then only E_1 should be used.

To decide which course to take, we will test whether E_1 and E_2 are equal within random variation. The statistical hypothesis test should be made using the Z - statistic, which is defined below, if E_1 and E_2 were both computed from reasonably large samples, say with size over 30. Otherwise, the t - statistic should be used. The t - statistic is discussed in many introductory statistics textbooks, such as References 5 and 7.

As is seen in the numerical examples below, if one sample size is much larger than the other (e.g., by a factor of 16) then very little is gained by including the estimate calculated from the smaller sample size. If both sample sizes are very small, it is unlikely that an adequate emission increment will be obtained even by using both data sets. For these reasons, the discussion below is centered around the Z - statistic; application of the t - statistic, in accordance with References 5 and 7, however, is very similar.

Define:

$$Z = \frac{E_2 - E_1}{\sqrt{S_{E_1}^{\,2} + S_{E_2}^{\,2}}}$$

If E_1 and E_2 are normally distributed, then Z is normally distributed with variance one. If E_1 and E_2 estimate the same emission factor, then Z has mean zero. Standard statistical methods then can be used to perform the statistical hypothesis test, as is discussed below.

If $|Z| > Z_{crit}$, where Z_{crit} is a value chosen from a normal variable table, then the difference between E_1 and E_2 is too great to be explained by random errors. In this case, E_1 should be used as the estimate.

If the test is to be performed at the 0.05 confidence level, for example, then $Z_{crit} = 1.96$. The confidence level is the probability of concluding that there is a true difference between E_1 and E_2 when there is not.

If $|Z| \leq Z_{crit}$, then the difference between E_1 and E_2 can reasonably be explained in terms of random sampling errors alone. In this case, E_1 and E_2 should both be used to derive the emission estimate.

The minimum variance estimate of the emission factor, then, is as follows:

$$E = a_1E_1 + a_2E_2$$

where

$$a_1 = \frac{1/S_{E_1}^{\ 2}}{\frac{1}{S_{E_1}^{\ 2}} + \frac{1}{S_{E_2}^{\ 2}}}$$

and

$$a_2 = \frac{1/S_{E_2}^{\ 2}}{\frac{1}{S_{E_1}^{\ 2}} + \frac{1}{S_{E_2}^{\ 2}}}$$

That is, E_1 and E_2 are weighted by the inverses of their variances. The variance of the estimate E, then, is:

$$S^2 = a_1^{\ 2} S_{E_1}^{\ 2} + a_2^{\ 2} S_{E_2}^{\ 2}$$

Since:

$0 < a_1 < 1$ and $0 < a_2 < 1$ when both variances are greater than zero, the variance S^2 is less than either $S_{E_1}^{\ 2}$ or $S_{E_2}^{\ 2}$. Thus, the uncertainty has been reduced in using both E_1 and E_2 to compute the final emission factor.

To reiterate, it has been assumed that the sample sizes used to compute E_1 and E_2 are large enough (at least 30) to justify using the Z- rather than the t-statistic. At least this sample size would be required in most cases to obtain acceptable accuracy, anyway.

Example

Suppose the analysis described above is to be used to test the emission factor for valves and that the following data are available:

E_1 = 0.0400 lb/hr,

S_{E_1} = 0.0080 lb/hr,

E_2 = 0.0600 lb/hr, and

S_{E_2} = 0.0320 lb/hr.

The value of S_{E_2} is four times as large as the value of S_{E_1}. This is about what one would expect if E_1 were determined from a sample size 16 times as large as that used to determine E_2.

Then:

$$Z = \frac{0.060 - 0.040}{\sqrt{(0.008)^2 + (0.032)^2}} = 0.61$$

Since:

$$|Z| = 0.61 < 1.96$$

the difference between E_1 and E_2 can reasonably be explained in terms of random errors. It is important to notice that this conclusion is reached despite the rather large difference between E_1 and E_2. This is due to the relatively large uncertainty in E_2.

The two emission rate estimates then could be combined as follows:

$$a_1 = \frac{1/(0.008)^2}{\frac{1}{(0.008)^2} + \frac{1}{(0.032)^2}} = 0.94$$

Similarly:

$$a_2 = 0.06$$

NOTE: $a_1 + a_2$ must be one.

Then the updated emission rate is:

$$E = 0.94(0.040) + (0.06)(0.060)$$

$$E = 0.041$$

and

$$S = \sqrt{(0.94)^2(0.008)^2 + (0.06)^2(0.032)^2}$$

$$= 0.0078$$

The emission rate estimate then has been changed by only 0.001, from 0.040 to 0.041, and the standard deviation of the emission rate has also been changed by a very small account. Thus, since E_2 was estimated from a small sample size, the improvement in the emission increment is very small.

Suppose, then, that E_2 had been estimated from a sample one-fourth as large as that used to estimate E_1, and that:

$S_{E_2} = 0.016.$

Note that

$$S_{E_2} = \frac{S_{E_1}}{\sqrt{4}}$$

Then:

$a_1 = 0.80,$

$a_2 = 0.20,$ and

$$S = \sqrt{(0.80)^2(0.008)^2 + (0.20)^2(0.016)^2} = 0.0072$$

In this case, the uncertainty in the emission estimate is reduced by 10% from 0.0080 to 0.0072, by pooling the data.

One might ask, how large a sample of data would be required for the estimation of E_2 for the data pooling to result in a significant improvement in the emission rate estimate. The type of calculations demonstrated here can be used to address this question; the fact that the standard deviation of a sample mean value decreases as the square root of the sample size increases was used in choosing the values of S_{E_2}; e.g., increasing the sample size by a factor of four would be expected to reduce the uncertainty in the mean by a factor of two.

If the sample sizes used to estimate E_1 and E_2 were the same, and if

$$S_{E_1} = S_{E_2} = 0.0080,$$

then,

$$a_1 = a_2 = 0.5$$

and

$$S_E = \sqrt{(0.5)^2(0.008)^2 + (0.5)^2(0.008)^2} = 0.0057$$

In this case, the uncertainty in the emission estimate is reduced from 0.0080 to 0.0057, or by a factor of 0.707.

SECTION 5

SUMMARY

This report presents a statistical approach for testing whether a planned refinery expansion can be made without increasing the fugitive emissions, when this is required by regulations. Emission factors can be used to estimate the emission increase or decrease due to the expansion. Being empirically determined, however, the emission factors are subject to random errors. Thus, the effect of the expansion on emissions cannot be computed exactly.

For this reason, the problem has been treated statistically. Anaytical methods are presented which can be used to estimate the emission increment as a function of:

- the facilities which are to be added,
- existing facilities which will be shut down (if any), and
- hardware changes or improved maintenance programs designed to reduce emissions in the old facilities and to limit the new emissions in the planned facilities.

The emission increment, I, can be thought of as:

$$I = A - D$$

where

A = emissions from the new facility in, say, lb/hr, and

D = emission <u>decrease</u> in the existing facility due to hardware changes or improved maintenance practices.

If the value of I is negative, the expansion "probably" can be made without increasing emissions. If I is positive, emissions "probably" will be increased. As is indicated above, however, I is affected by the random errors in the emission factors. Thus, the standard deviation of I is also computed, and the probability that I is less than zero is obtained; this is the probability that the expansion as planned will not increase fugitive emissions.

Other probabilities, e.g. that I is less than 50 lb/hr, can also be computed, if this is desired.

The methods presented here can be used as an aid in comparing several options for reducing emissions to acceptable levels. This involves simply performing the statistical calculations in parallel for each option to be considered. The final decision, of course, should also involve economic and other considerations.

The methods are designed to handle different control strategies in the existing and in the planned facilities and different strategies for different types of device, if this is necessary. The following situation, for example, could be handled.

- A catalytic reformer is to be added to an existing refinery.

- To reduce emissions in the existing facilities, an improved maintenance program is instituted for valves, and the API separator is to be covered.

- To limit the added emissions in the new facility, the same type of maintenance program used in the old facility will be instituted for valves, and double seals will be used rather than single seals for all pumps. Single seals remain on all pumps in the old facility, however.

REFERENCES

1. Aitchison, John, "On the Distribution of a Positive Randon Variable Having a Discrete Probability Mass at the Origin", Journal of the American Statistical Association, 50, 9 (1955), pp. 901-908.

2. Finney, D. J., "On the Distribution of a Variate Whose Logarithm is Normally Distributed", Journal of the Royal Statistical Society Series B, 7, pp. 155-161.

3. Jefcoat, I. A., Leigh Short, R. G. Wetherold, "Fugitive Emission Control Strategy for Petroleum Refineries", Paper Presented at the Refinery Emissions Symposium, Jekyll Island, Georgia, April 26-28, 1978.

4. Jones, Harold R., Pollution Control in the Petroleum Industry, Noyes Data Corporation, Park Ridge, New Jersey, 1973.

5. Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes, Introduction to the Theory of Statistics, McGraw-Hill Book Company, New York, 1963.

6. Serth, R. W. and T. W. Hughes, "Error Analysis for Plant Expansion Problem", Monsanto Research Corporation, Dayton, Ohio, March 18, 1977.

7. Snedecor, George W. and William G. Cochran, Statistical Methods, The Iowa State University Press, Ames, Iowa, 1967.

APPENDIX A

PROPERTIES OF THE MEAN AND VARIANCE

This Appendix includes a brief discussion of the properties of the mean, variance, and covariance. These statistical measures have been used extensively in the study reported herein. Also, Reference 5 includes an excellent discussion of these statistical measures. Some properties of the covariance that are not readily available in statistical texts are derived in this appendix.

Mean

The mean or arithmetic average of a random variable X can be estimated from a set of N values, X_i, i = 1 to N, as follows:

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$$

If μ_X is the mean of X and K is a constant, then the mean of KX is $K\mu_X$. If Y has mean μ_Y, then the mean of X + Y is $\mu_X + \mu_Y$. From these properties the mean of any linear combination of random variables can be obtained. The mean of:

$$3 + 4X + 2Y$$

for example is:

$$3 + 4\mu_X + 2\mu_Y$$

If X and Y are independently distributed, that is, if the value of one is not influenced by the value of the other, then the mean of XY is $\mu_X\mu_Y$.

Variance

The variance σ^2 is a measure of the amount of scatter or dispersion a quantity has; the more widely it varies, the greater the variance is. Thus, a quantity with a large variance is considered to have a large uncertainty.

The variance $\sigma_X{}^2$ of a variable X can be estimated from a set of N values as follows:

$$S^2 = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N - 1}$$

The standard deviation, σ, is the positive square root of the variance.

If X and Y are independently distributed, the variance of X + Y is:

$$\sigma_X{}^2 + \sigma_Y{}^2$$

where σ_Y^2 is the variance of Y. The variance of K_X is

$$K^2\sigma_X^2$$

where, as before, K is a constant. The variance of any constant is zero. From these properties, the variance of a linear combination of independent variables can be derived. The variance of:

$$3 + 4X + 2Y$$

for example is

$$16\sigma_X^2 + 4\sigma_Y^2$$

If X and Y are independent, the variance of XY is:

$$\mu_X^2\sigma_Y^2 + \mu_Y^2\sigma_X^2 + \sigma_X^2\sigma_Y^2$$

If X and Y are not independent, the variance of X + Y is

$$\sigma_X^2 + \sigma_Y^2 + 2\,\mathrm{cov}(X,Y)$$

where cov(X,Y), the covariance between X and Y, can be estimated from a set of N values of X and Y as follows:

$$\frac{\sum_{i=1}^{N} (X_i - \overline{X})(Y_i - \overline{Y})}{N - 1}$$

The more closely X and Y are (linearily) related, the larger the covariance is in magnitude. If Y is a perfect linear function of X with positive slope, then the covariance equals its maximum possible value, $\sigma_X\sigma_Y$. When X and Y are independent, the covariance is zero.

The variance of X + XY where X and Y are independent is

$$\sigma_X^2 + \sigma_{XY}^2 + 2\ \mathrm{cov}(X, XY) =$$

$$\sigma_X^2 + \mu_Y^2\sigma_X^2 + \mu_X^2\sigma_Y^2 + \sigma_X^2\sigma_Y^2 + 2\ \mathrm{cov}(X, XY).$$

The covariance between X and XY is

$$\mu_Y\sigma_X^2$$

This relationship, which is not ordinarily given in textbooks, is proved below.

The same expression for the variance of X + XY can be obtained by writing X + XY as X(1 + Y) and using the formula given above for the variance of a product of two independent random variables.

Derivation of the Covariance Between X and XY

Suppose X and Y are independent random variables with respective means μ_X and μ_Y and variances $\sigma_X{}^2$ and $\sigma_Y{}^2$. The covariance between X and XY, denoted cov(X, XY), is derived below.

The following formal definition of covariance will be needed:

$$\text{cov}(\alpha,\beta) = E[(\alpha - \mu_\alpha)(\beta - \mu_\beta)]$$

where α and β are random variables with respective means μ_α and μ_β, and "E" denotes the "expected value" or "average value." Thus, the covariance is the average value of $(\alpha - \mu_\alpha)(\beta - \mu_\beta)$.

By the definition of covariance, then,

$$\begin{aligned}\text{cov}(X,\ XY) &= E[(X - \mu_X)(XY - \mu_X\mu_Y)] \\ &= E[X^2Y - X\mu_X\mu_Y - \mu_X XY + \mu_X{}^2\mu_Y] \\ &= E[X^2Y] - \mu_X{}^2\mu_Y - \mu_X{}^2\mu_Y + \mu_X{}^2\mu_Y \\ &= \mu_Y E[x^2] - \mu_X{}^2\mu_Y\end{aligned}$$

Now, it can easily be shown that:

$$\sigma_X{}^2 = E[x^2] - \mu_X{}^2$$

from the formal definition of the variance

$$\sigma_X{}^2 = E[(X-\mu_X)^2]$$

so

$$E[x^2] = \sigma_X{}^2 + \mu_X{}^2$$

Thus,

$$\begin{aligned}\text{cov}(X,\ XY) &= \mu_Y(\sigma_X{}^2 + \mu_X{}^2) - \mu_X{}^2\mu_Y \\ &= \mu_Y\sigma_X{}^2\end{aligned}$$

Derivation of the Covariance Between XY and Z

Suppose X, Y, and Z are random variables with respective means μ_X, μ_Y and μ_Z, and X is independent of Y and Z. Then:

$$\begin{aligned}
\text{cov}(XY, Z) &= E[(XY - \mu_X\mu_Y)(Z - \mu_Z)] \\
&= E[XYZ - XY\mu_Z - \mu_X\mu_Y Z + \mu_X\mu_Y\mu_Z] \\
&= \mu_X E[YZ] - \mu_X\mu_Y\mu_Z - \mu_X\mu_Y\mu_Z + \mu_X\mu_Y\mu_Z \\
&= \mu_X\mu_Y\mu_Z + \mu_X\text{cov}(Y, Z) - \mu_X\mu_Y\mu_Z \\
&= \mu_X\text{cov}(Y, Z)
\end{aligned}$$

Derivation of the Covariance Between $\overline{X}$ and $\overline{Y}$

An addition property of covariance is needed for use in Appendix B. Suppose the covariance cov(X, Y) between two random variables X and Y is known, and the covariance between the sample means $\overline{X}$ and $\overline{Y}$ is needed. The quantities $\overline{X}$ and $\overline{Y}$ are the means of samples of size N of values of X and Y, respectively.

Then, if μ_X and μ_Y denote the means of X and Y, respectively.

$$\text{cov}(\overline{X}, \overline{Y}) = \text{cov}\left(\frac{\sum_{i=1}^{N} X_i}{N}, \frac{\sum_{i=1}^{N} Y_i}{N}\right)$$

$$= E\left[\left(\frac{\sum_{i=1}^{N} X_i}{N} - \mu_X\right)\left(\frac{\sum_{i=1}^{N} Y_i}{N} - \mu_Y\right)\right]$$

$$= \frac{1}{N^2} E\left[\left(\sum_{i=1}^{N} (X_i - \mu_X)\right)\left(\sum_{i=1}^{N} (Y_i - \mu_Y)\right)\right]$$

$$= \frac{1}{N^2} \sum_{i=1}^{N} E[(X_i - \mu_X)(Y_i - \mu_Y)] + \frac{1}{N^2} \sum_{i \neq j} E[(X_i - \mu_X)(Y_j - \mu_Y)]$$

$$= \frac{1}{N^2}[N\,\text{cov}(X, Y)] + \frac{1}{N^2}(0)$$

$$= \frac{\text{cov}(X, Y)}{N}$$

Estimation of Mean and Variance for Skewed Distributions

The expressions given above for estimating the mean and variance are unbiased estimators, that is, their expected values equal the population values they estimate. This is true regardless of the underlying probability distribution.

If the distribution is highly skewed (asymmetric), however, these estimators are not efficient; other estimators exist which have smaller uncertainties. The lognormal distribution is an example of such a distribution.

Reference 2 presents a discussion of the efficient estimation of the mean and variance of a lognormally distributed random variable.

APPENDIX B

A PAIRED MEASUREMENT SCHEME FOR REDUCING RANDOM ERRORS

In this appendix, a method designed to reduce the variance of I is described. This is beneficial in that if I is known more accurately, then a more confident decision can be made as to whether a refinery can be expanded without increasing emissions. The extent of the reduction in uncertainty, or whether any reduction is achieved, depends on factors discussed below.

The basic approach involves making emissions measurements before and after a program to reduce emissions is instituted. The program could include hardware modification or improved maintenance practices. Then define

X_i = measured emission rate from the i^{th} tested unit before the program is started

Y_i = measured rate after the program is started for the <u>same</u> <u>unit</u>, and

$r_i = X_i - Y_i$, the decrease in emissions for the i^{th} unit due to the program.

The average emission reduction, then, is

$$R = \frac{\sum_{i=1}^{N} r_i}{N}$$

where N is the number of units tested and the variance σ_R^2 is estimated by

$$S_R^2 = \frac{\left(\frac{\sum_{i=1}^{N} (r_i - R)}{(N-1)}\right)}{N}$$

In the expression above, the quantity in parentheses is the variance of the individual values of r, and this must be divided by N to obtain the variance of the mean of N values of r. (More complicated estimates exist for the mean and variance which are more efficient when the distribution is highly skewed, as is discussed in Appendix A.)

Now, it is clear that the average emission reduction can also be expressed

$$R = \frac{\sum_{i=1}^{N} r_i}{N} = \frac{\sum_{i=1}^{N} X_i - Y_i}{N}$$

$$= \frac{\sum_{i=1}^{N} X_i}{N} - \frac{\sum_{i=1}^{N} Y_i}{N}$$

$$= \overline{X} - \overline{Y}$$

where $\overline{X}$ and $\overline{Y}$ are the means of the X's and the Y's, respectively.

The final expression for R indicates that the variance of R can also be written

$$\sigma_{\bar{X}}^2 + \sigma_{\bar{Y}}^2 - 2 \text{ cov } (\bar{X},\bar{Y})$$

where cov $(\bar{X},\bar{Y})$ is the covariance between $\bar{X}$ and $\bar{Y}$. The following brief discussion of covariance relates importantly to the physical problem:

The covariance, which is defined formally and discussed in Appendix A, is a measure of the extent to which two variables are linearly related. If there is no relationship, then the covariance is zero. If one variable tends to increase as the other increases, then the covariance is positive.

Thus, if the emission rates before and after the program is started are not related, then the variance of R becomes

$$\sigma_{\bar{X}}^2 + \sigma_{\bar{Y}}^2$$

and the effort to ensure that the same set of sources are measured before and after achieves nothing; any set of sources could have been tested before and any other set after, and the result would have been the same (excluding differences in random errors).

If X_i and Y_i are linearly related, however, cov $(\bar{X},\bar{Y})$ is positive, and using a paired measuring scheme reduces the uncertainty in R; the amount of the variance reduction is 2 cov $(\bar{X},\bar{Y})$. Whether the paired approach should be used in a given case depends on the extent to which X and Y are related and the expense and inconvenience of using a paired scheme for estimating R.

Additionally, estimated emission rates may be available which were not obtained by using the paired approach. If a paired experiment were performed, then either (1) the existing factors would not be used or (2) a much more complicated analysis approach would be employed to combine emission rates estimated by different methods. The second option is not considered to be desirable, since it is not at all certain that a significant reduction in uncertainty would be achieved over using the simpler, unpaired approach.

In the analysis which follows, the covariances between related emission rates (such as $\bar{X}$ and $\bar{Y}$) and between emission rates and emission reductions (such as X and R) are needed to compute the uncertainty of the estimated emission increment. The equations needed to calculate these covariances are given below.

The covariance between X and Y is estimated as follows

$$\widehat{cov}(X,Y) = \frac{\sum_{i=1}^{N} (X_i-\bar{X})(Y_i-\bar{Y})}{N-1}$$

The covariance between $\bar{X}$ and $\bar{Y}$, then, is

$$cov(\bar{X},\bar{Y}) = \frac{cov(X,Y)}{N}$$

This relationship is proved in Appendix A. The covariance between X and r is estimated by

$$\widehat{cov}(X,r) = \frac{\sum_{i=1}^{N} (X_i-\bar{X})(r_i-R)}{N-1}$$

and the estimated covariance between $\bar{X}$ and R is

$$\widehat{cov}(X,R) = \frac{\widehat{cov}(X,r)}{N}$$

In Section 2, the calculation of the mean and variance of the emissions increment for a given type of device is discussed. Several scenarios representing different types of effort to reduce emissions are analyzed. The same scenarios are discussed below, except that the emissions reductions are assumed to have been estimated by using the paired-measurement scheme.

Case 1^1 - Same Provisions for Reducing Emissions in the New and Old Refineries

Define:

R = average reduction in emissions per source due to maintenance,

S_R = standard deviation of R,

E^1 = average emission rate per source without maintenance,

$S_E{}^1$ = standard deviation of E^1,

E^{11} = average emission rate per source with maintenance,

$S_E{}^{11}$ = standard deviation of E^{11},

N_o = number of sources of the type being considered in the old facility, and

N_n = number of sources in the new facility.

This case corresponds to Case 1 discussed in Section 2. The covariance between R and E^{11} will also be needed, as is discussed below.

Then the emission reduction in the old facility corresponding to this particular type of source is

$$N_o R$$

and the added emissions in the new facility is

$$N_n E^{11}$$

Thus, the total emission increment I is

$$I = N_n E^{11} - N_o R$$

The variance S_I^2 of I is

$$S_I^2 = N_n^2 S_{E^{11}}^2 + N_o^2 S_R^2 - 2N_n N_o \hat{\text{cov}}(R,E^{11})$$

It will now be shown that the analysis here is consistent with that presented in Section 2. To do this, we will replace R by $E^1 - E^{11}$ in the expression for I, to obtain:

$$N_n E^{11} - N_o R =$$

$$N_n E^{11} - N_o(E^1 - E^{11}) =$$

$$(N_n + N_o) E^{11} - N_o E^1$$

This is exactly the expression for I given in Case 1, Section 2.

If E^1 and E^{11} are assumed to be independent, the variances given here and in Section 2 can also be shown to be consistent. Under this assumption,

$$S_R^2 = S_{E^1}^2 + S_{E^{11}}^2$$

and

$$\hat{\text{cov}}(R,E^{11}) = - S_{E^{11}}^2$$

Thus, S_I^2 becomes

$$N_n^2 \sigma_{E^{11}}^2 + N_o^2 (S_{E^1}^2 + S_{E^{11}}^2) + 2N_n N_o S_{E^{11}}^2 =$$

$$(N_n^2 + 2N_n N_o + N_o^2) S_{E^{11}}^2 + N_o^2 S_{E^1}^2 =$$

$$(N_n + N_o)^2 S_{E^{11}}^2 + N_o^2 S_{E^1}^2$$

and this is exactly the expression given for S_I^2 in Section 2.

<u>Example for Case 1[1]</u>

As in Example 1 for Case 1, suppose a catalytic reformer is being added to an existing refinery and that this adds 850 valves. An improved maintenance program is introduced which reduces the average emission rate from a given valve from 0.040 to 0.008 pounds per hour. These and other necessary statistics are presented below.

$R = 0.040 - 0.008 = 0.032$ lb/hr,

$S_{\hat{R}} = 0.004$ lb/hr,

$E^1 = 0.040$ lb/hr,

$S_{E^1} = 0.006$ lb/hr,

$E^{11} = 0.008$ lb/hr,

$S_{E^{11}} = 0.003$ lb/hr,

$N_o = 14000$,

$N_n = 850$, and

$\hat{cov}(R,E^{11}) = 3.6 \times 10^{-6}$ (lb/hr)^2

$\hat{cov}(R,E^{11})$ and S_R would have the values given above if the emission factors E^1 and E^{11} had a correlation of 0.7.

The value of I is the same as in Example 1 for Case 1:

I = -441 lb/hr

but S_I^2 is reduced:

$$S_I^2 = N_n^2 S_{E^{11}}^2 + N_o^2 S_R^2 - 2 N_n N_o \hat{cov}(R,E^{11})$$

$$= (850)^2(0.003)^2 + (14000)^2 (0.004)^2 - (850)(14000)(3.6 \times 10^{-6})$$

$$= 3100$$

$$S_I = 56 \text{ lb/hr.}$$

Thus, a reduction from S_I = 95 pounds per hour, which is the result given in Example 1 for Case 1^1, to S_I = 56 pounds per hour is achieved by using the pairing scheme.

If the emission factors E^1 and E^{11} had a correlation of 0.3, instead of 0.7 as in the case above, S_R and $\hat{cov}(R,E^{11})$ would have the following values:

$$S_R = 0.006$$

$$\hat{cov}(R,E^{11}) = -3.6 \times 10^{-6}$$

The resulting value of S_I is 84 pounds per hour. In this case a smaller reduction (84 versus 95) in S_I is achieved by using the pairing scheme.

As is discussed in Section 2, the analysis of Case 1 (Case 1^1 here) can be used to handle the situations covered under Cases 2 and 3 (Cases 2^1 and 3^1 here). Cases 2^1 and 3^1 treat scenarios in which screening for maintenance is performed, and high-leaking devices are repaired. In these cases, if the mean (E^{11}) and standard deviation ($S_{E^{11}}$) of the emissions from a collection of units which are selectively maintained are known, then the analysis discussed under Case 1^1 should be used. The analysis below, however, can be used when the information in this exact form is not known. Although more numerous, the required inputs for Cases 2^1 and 3^1 may be more easily obtained than those for Case 1^1 in some instances.

<u>Case 2^1 - Emissions Monitoring and Subsequent Corrective Action Performed the Same in the Old and New Facilities</u>

Define

R_1 = reduction in emissions per unit due to maintenance,

$S^2_{R_1}$ = variance of R_1,

F_2 = average emission rate of maintainable leakers if maintenance is performed,

F_3 = average emission rate of other units,

S_{F_i} = standard deviation of F_i, $i=2, 3$,

f = proportion of the units of the type being considered which would be high-leakers at any given time on the average, if detailed maintenance were not performed,

S_f = standard deviation of f,

N_o = total number of units of the type being considered in the old facility, and

N_n = total number in the new facility.

This case corresponds to Case 2 in Section 2.

The emission reduction in the old faclity, then, is

$$f\ N_o\ R$$

and the emission increment in the new facility is

$$f\ N_n F_2 + (1-f)\ N_n F_3$$

The total emission increment is

$$\begin{aligned} I &= f\ N_n F_2 + (1-f)\ N_n F_3 - f\ N_o R \\ &= [\,N_n F_2 - N_n F_3 - N_o R\,]\ f + N_n F_3 \\ &= Af + N_n F_3 \end{aligned}$$

where A denotes the coefficient of f, which is enclosed in brackets in the expression above. The variance S_I^2 of I is given by

$$S_I^2 = A^2 S_f^2 + f^2 S_A^2 + S_A^2 S_f^2 + N_n^2 S_{F_3}^2$$

$$+ 2f \left[N_n^2 \hat{\text{cov}}(F_2, F_3) - N_n^2 S_{F_3}^2 - N_o N_n \hat{\text{cov}}(R, F_3) \right]$$

where

$$S_A^2 = N_n^2 S_{F_2}^2 + N_n^2 S_{F_3}^2 + N_o^2 S_R^2$$

<u>Case 3[1] - Same Hardware Provisions for Reducing Emissions in New and Old Facilities, Along with Emissions Monitoring and Subsequent Corrective Action Performed in Both Facilities</u>

Define

R = average reduction in emissions per source due to the hardware modification,

S_R^2 = variance of R,

R_1 = average reduction in emissions per source due to maintenance,

$S_{R_1}^2$ = variance of R , and

F_2, F_3, S_{F_2}, S_{F_3}, f, S_f, N_o,

and N_n are as defined in the preceding section.

This case corresponds to Case 3 in Section 2.

The emission reduction in the old facility is

$$N_o R + f\, N_o\, R_1$$

and the emission increment in the new facility is

$$f\, N_n F_2 + (1-f)\, N_n F_3$$

The emission increment for the entire refinery for the type of device being considered is

$$I = f\,N_n F_2 + (1-f)\,N_n F_3 - N_o R - f\,N_o R_1$$

$$= f\,[\,N_n F_2 - N_n F_3 - N_o R_1\,] + N_n F_3 - N_o R$$

$$= Af + N_n F_3 - N_o R$$

where A denotes the coefficient of f which appears in brackets above.

$$S_I^{\,2} = A^2 S_f^{\,2} + f^2 S_A^{\,2} + S_A^{\,2} S_f^{\,2} + N_n^{\,2} S_{F_3}^{\,2} + N_o^{\,2} S_R^{\,2}$$

$$+ 2f\,[\,N_n^{\,2}\,\hat{\mathrm{cov}}\,(F_2,F_3) - N_n^{\,2} S_{F_3}^{\,2} - N_o N_n\,\hat{\mathrm{cov}}\,(R_1,F_3)$$

$$- N_o N_n\,\hat{\mathrm{cov}}\,(R,F_2) + N_o N_n\,\hat{\mathrm{cov}}\,(R,F_3) + N_o^{\,2}\,\hat{\mathrm{cov}}\,(R_1,R)\,]$$

$$+ 2N_o N_n\,\hat{\mathrm{cov}}\,(R,F_3)$$

Case 4' - Different Emission Factors Entirely Apply for the New and Old Facilities

The discussion presented under Case 4 in Section 2 applies directly in this case whether the paired-measurement approach is used or not.

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
EPA-600/2-78-195		78-200-187-27-09

4. TITLE AND SUBTITLE Statistical Analysis of Fugitive Emission Change Due to Refinery Expansion	5. REPORT DATE September 1978 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Hugh J. Williamson and Lloyd P. Provost	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Radian Corporation P.O. Box 9948 Austin, Texas 78766	10. PROGRAM ELEMENT NO. 1AB604C 11. CONTRACT/GRANT NO. 68-02-2608, Task 27
12. SPONSORING AGENCY NAME AND ADDRESS EPA, Office of Research and Development Industrial Environmental Research Laboratory Research Triangle Park, NC 27711	13. TYPE OF REPORT AND PERIOD COVERED Task Final; 2-7/78 14. SPONSORING AGENCY CODE EPA/600/13

15. SUPPLEMENTARY NOTES IERL-RTP project officer is Irvin A. Jefcoat, Mail Drop 62, 919/541-2547.

16. ABSTRACT The report discusses a statistical approach for determining if a planned petroleum refinery expansion can be carried out without increasing fugitive emissions. The random uncertainty of the empirically determined emission factors is taken into account during the determination. The method presented is designed to handle different control strategies in the existing and planned facilities and different strategies for different types of devices (e.g., pumps and valves) if necessary. It is also possible to evaluate different options for reducing emissions to acceptable levels.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS		b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group	
Pollution	Random Error	Pollution Control	13B	12A
Processing	Measurement	Stationary Sources	13H	
Leakage	Probability Theory	Fugitive Emissions	14B	
Refineries			13I	
Petroleum Refining				
Petroleum Industry			05C	
Expansion				

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
Unlimited	Unclassified	60
	20. SECURITY CLASS *(This page)*	22. PRICE
	Unclassified	

EPA Form 2220-1 (9-73)