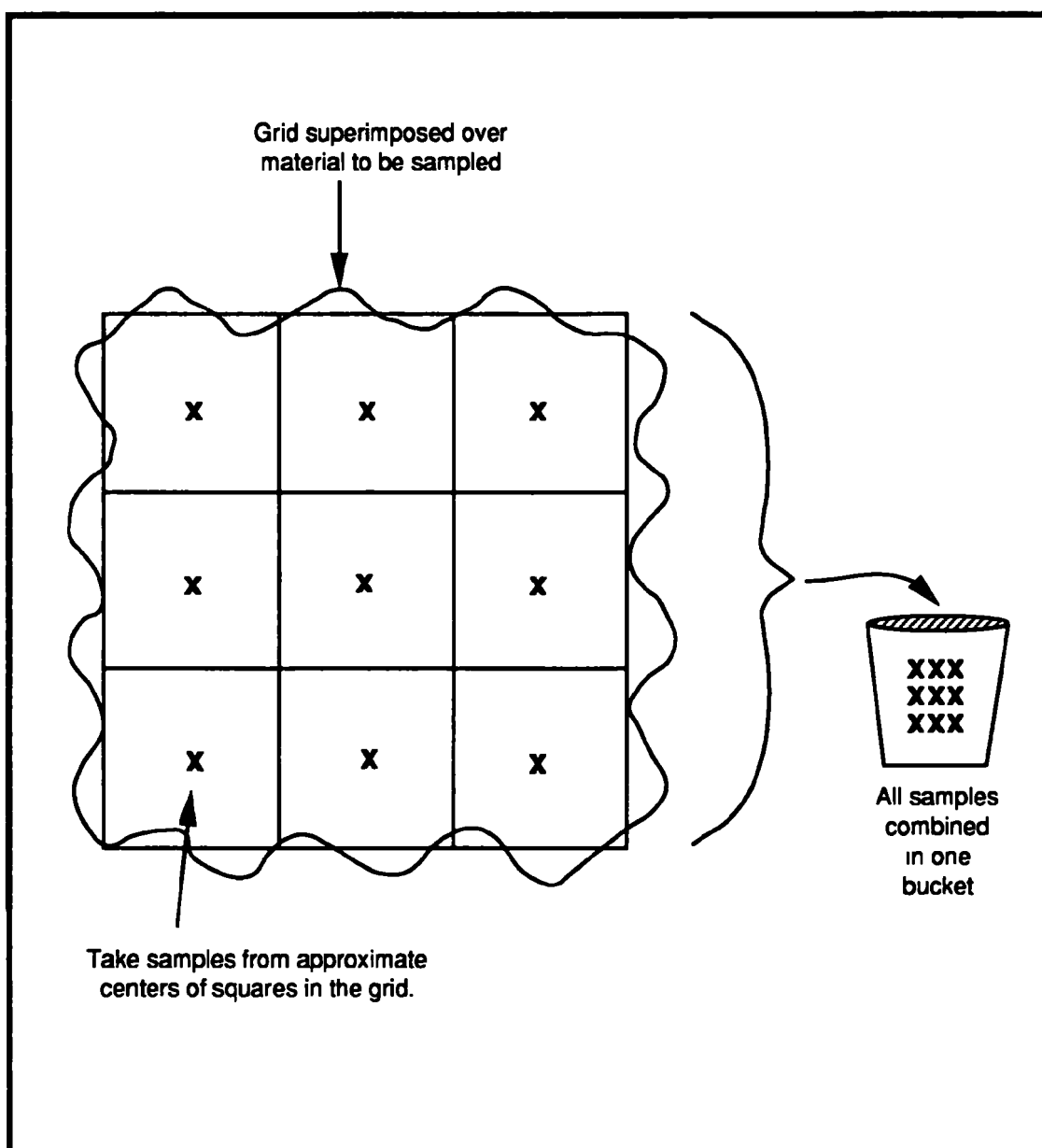

Office of Prevention, Pesticides, and Toxic Substances



SAMPLING GUIDANCE FOR SCRAP METAL SHREDDERS

Technical Background



**SAMPLING GUIDANCE FOR SCRAP
METAL SHREDDERS**

Technical Background

August 1993

**United States Environmental Protection Agency
Office of Prevention, Pesticides
and Toxic Substances
Washington, DC 20460**

DISCLAIMER

This document has been reviewed and approved for publication by the Office of Pollution Prevention and Toxics, U.S. Environmental Protection Agency. The use of trade names or commercial products does not constitute Agency endorsement or recommendation for use.

Authors and Contributors

Westat Project Staff:

James Bethel
Ralph DiGaetano
Mary Pepler

EPA Project Staff:

Edith Sterrett, Project Officer
Susan Dillman, Task Manager

Acknowledgments

The Office of Pollution Prevention and Toxics wishes to thank everyone involved with this project at Westat and at the Environmental Protection Agency, as well as others that have contributed to the document.

In particular, we acknowledge the technical help of John Rogers and William Devlin of Westat in preparing this document. Mary Lou Pieranunzi, Angelia Murphy, Maida Montes and Anna Page also helped to prepare and proof-read the manuscript.

We are grateful to the many reviewers in the Office of Pollution Prevention and Toxics and in other branches of the Environmental Protection Agency for reading the document. In particular, we thank Brad Schultz and Dan Reinhart of the Exposure Evaluation Division.

We owe special thanks to Mitchell D. Erickson of Argonne National Laboratory, David N. Speis of Environmental Testing and Certification Corporation, and Herschel Cutler of the Institute of Scrap Recycling Industries for reviewing this document and contributing many useful comments about it. Their review does not constitute approval.

TABLE OF CONTENTS

TECHNICAL BACKGROUND

1.	Overview	1
2.	Modeling Sampling and Analytical Errors.....	1
3.	Standard Error Calculations	13
3.1	Data Sources.....	13
3.2	Numerical Methods.....	17
4.	Confidence Intervals and Relative Error.....	20
4.1	Numerical Methods.....	20
4.2	Alternative Confidence Levels	25
5.	Hypothesis Testing.....	25
5.1	Monitoring	25
5.2	Clean-up Verification	28
6.	Power Calculations.....	30
6.1	Monitoring	30
6.2	Clean-up Verification	36

LIST OF TABLES

Table		Page
1	Summary of preliminary data	15
2	Relative errors for estimating PCB levels with sample sizes of 2 to 25.....	21
3	t-values for confidence interval and hypothesis tests.....	23
4	Relative errors for estimating PCB levels with sample sizes of 2 to 25 with 90% confidence	26
5	Relative errors for estimating PCB levels with sample sizes of 2 to 25 with 99% confidence	27
6	Cut-off values for monitoring.....	29
7	Cut-off values for clean-up verification.....	31
8	Chance of finding violations in monitoring with a 25 ppm standard.....	32

LIST OF TABLES (Continued)

Table		Page
9	Chance of finding violations in monitoring with a 50 ppm standard.....	33
10	Chance of finding violations in monitoring with a 100 ppm standard	34
11	Chance of finding violations in monitoring with a 50 ppm standard with 90% confidence	37
12	Chance of finding violations in monitoring with a 50 ppm standard with 99% confidence	38
13	Chance of requiring additional clean-up with a 25 ppm standard.....	39
14	Chance of requiring additional clean-up with a 50 ppm standard.....	40
15	Chance of requiring additional clean-up with a 100 ppm standard	41
16	Chance of requiring additional clean-up with a 50 ppm standard with 90% confidence	43
17	Chance of requiring additional clean-up with a 50 ppm standard with 99% confidence	44

SAMPLING GUIDANCE FOR SCRAP METAL SHREDDERS

Technical Background

1. Overview

The purpose of this volume is to provide additional detail on the assumptions, background data, and numerical calculations that support *PCB Sampling Guidance for Scrap Metal Shredders*. The discussion is somewhat technical at times, but in general can be read by anyone with a modest background in mathematical statistics. The primary emphases are on the following areas:

- Assumptions regarding sampling and analytical errors;
- Background data for estimating standard errors;
- The effects of compositing on standard errors;
- Calculations used to estimate relative error in confidence intervals;
- Calculations for finding cut-off values and power in hypothesis testing; and
- Effects of alternative confidence and significance levels on power and relative error.

For the convenience of the reader, tables and definitions are reproduced here, wherever possible, so that it should not be necessary to refer back to the *Field Manual* while reading this volume.

2. Modeling Sampling and Analytical Errors

Variance Components. The sampling model involves three components:

- Sampling over time, which will generally mean the relatively short period of time in which the samples are being collected, such as several hours or, at most, several days. In some monitoring programs, the reference time period might be longer, if certification is determined based on samples collected over several weeks or months.

- Sampling over space, such as taking samples from a fluff pile or from volume of fresh fluff that has been spread out for sampling.
- Potential errors in analysis of PCB content.

The first stage may not be applicable in all cases, but the second and third elements will always be present.

Sampling over time can actually have two different meanings. If a site is visited for 8 hours and samples of fresh output are taken for 10-minute intervals once an hour, then there is some sampling variability due to the fact that some 10-minute intervals were selected and not others. On the other hand, if a site is visited over a longer period of time, say several months, then there is sampling variability because of the days that are selected to visit the site.

Sampling in space requires grid sampling or some other method that insures that samples are distributed over the volume of material that is to be assessed. Whenever grid sampling is involved, the steps suggested in the manual require that all cells have samples taken from them. Since samples are taken from all cells, there is no “between-cell” variance component, but there is still some variability due to the fact that we are selecting one of the many different samples that could be selected from each cell.

Finally, besides this uncertainty about exactly which samples of material might be selected, there is uncertainty about the accuracy of laboratory measurements. (Note that because of this uncertainty, it is recommended in the *Field Manual* that at least 10% of the samples be analyzed in duplicate. See M.D. Erickson, *Analytical Chemistry of PCBs*, 1986, Butterworth, Stoneham, Massachusetts, p. 308.)

All these components are reflected, directly or indirectly, in making estimates of variance calculations for the sampling guidance document. Each component of variance corresponds to one of the possible sources of error discussed in Section 4.1. In the following sections we will describe the sampling and variance models in detail.

Sampling Model. The sampling model discussed in this section is intended to give a formal theoretical model for the grid sampling and other sampling procedures discussed in Chapter 2. It should be noted, however, that the procedures

recommended in Chapter 2 are not based on theoretical considerations but rather their apparently successful use in a study conducted for EPA.*

In its most general form, the sampling model consists of samples collected at m points in time, with n samples being selected at each time period, for a total of mn samples. Usually, when sampling over time, we would have $n = 1$, so that m samples would be collected. On the other hand, when sampling stored fluff, $m = 1$, and n samples would be collected. Furthermore, each of the collected samples would be divided into c subsamples for compositing. Finally, each composite sample would be divided into r replicates for laboratory analysis.

First, let us assume that output is sampled over some period of time, which may consist of several hours or several days. This will be achieved by stopping the shredder and taking samples at, say, m points over this period of time. For example, shredder output may be sampled over an 8-hour period, with samples being taken each half hour. In this case, there would be 16 samples over time, one for each half hour. If grid samples are being taken, then at time i , $1 \leq i \leq 16$, output would be collected and spread into a 3x3 grid. Within the l -th cell of the grid, $1 \leq l \leq 9$, a sample of material would be selected, after which all 9 samples would be combined and mixed. If grid sampling is not used (e.g., if no front loader is available), then grab samples would be taken from the pile of fresh fluff as described in Section 2 of the *Field Manual*.

When stored fluff is being sampled, there is only one time period (i.e., $m = 1$). If the steps described in Section 2 of the *Field Manual* are followed, then 20 samples would be taken, so that we would have $n = 20$.

If desired, the m samples over time, or n samples taken at a single time, may be randomly sorted into groups of c samples each and then composited into q (where $q = m/c$ or $q = n/c$) *composite* samples. Finally, each of the q composite samples is separated into r replicates for laboratory analysis.

* USEPA Fluff Pilot Program, conducted by Westat, Inc., Rockville, MD, and Midwest Research Institute, Kansas City, MO, for the Office of Pollution Prevention and Toxics (see *PCB, Lead, and Cadmium Levels in Shredder Waste Materials: A Pilot Study*. USEPA, Office of Toxic Substances. EPA 560/5-90-008B. 1991).

Thus the sampling model involves four steps: sampling over time, selecting a grid sample (or several grab samples*), compositing, and the formation and sampling of replicates. Three of the steps in this sampling model involve randomization: sampling over time, grid sampling, and sampling of replicates. In sampling over time, it is theoretically possible to eliminate the between-time sampling component by taking samples from every time period. For example, this would be the case if the shredder operation were shut down each half-hour and *all* output were arranged in a grid. In most applications, however, the output from even as brief a period as a half-hour would be too large in volume for such a procedure, so that for each half-hour only a fraction of the material produced will actually be selected and arranged in a grid. In the discussion below, it is assumed that this fraction is small, so that sampling over time is considered to be like sampling from an infinite population of possible times. Similarly, in selecting samples of material from grid cells, it is assumed that there are many possible samples of material, and this stage of sampling is also assumed to be like sampling from an infinite population of such samples. In sampling replicates, one or two of the eight replicates will be selected at random for laboratory analysis.

To facilitate the discussion of expectations and variances, it will be convenient to indicate the first stage of sampling – the selection of times – by S_1 , the second stage – the selection of grid samples – by S_2 , and the third stage – the selection of replicates – by S_3 . That is, S_1 will denote the actual sample of times selected, S_2 will denote the actual grid samples selected, and S_3 will denote the actual replicates selected. Furthermore, the following notations will be useful:

θ = actual average concentration among all material generated over the target time period,

θ_i = actual average concentration among all material at all time i , and

θ_{il} = actual average concentration at time i in cell l .

Other notations will be introduced as necessary in the discussion which follows.

* As we noted earlier, grab samples may be taken instead of grid samples. However, the grab sampling procedures described in Section 2 are relatively similar to grid sampling, in that several samples are taken from a variety of positions in the body of material and then combined into a single sample for analysis.

Variance Model. Let

X_{ijl} = actual concentration in the j -th sample taken from the l -th grid cell at time i .

Since X_{ijl} is the concentration of a sample of material that is selected at random, it is assumed that

$$E(X_{ijl} | S_1) = \theta_{il}.$$

That is, it is assumed that conditional on the time periods selected, the expected value of the concentration of a randomly selected sample of material is the average concentration of all the material. Unfortunately, this assumption requires something of a leap of faith. For example, if the sampling process is not done carefully, then “fines” – small particles of fluff – may be less likely to be selected. If, as is suspected, fines are more likely to contain PCBs, or other toxic materials, then the concentration of toxic materials in a sample might tend to be smaller than the overall average concentration.

The variance of X_{ijl} is defined as

$$V(X_{ijl} | S_1) = E((X_{ijl} - \theta_{il})^2 | S_1) = \sigma_{il}^2.$$

This term represents the variance of picking a single sample at random from cell l , at time i . As noted above, the population of possible samples is finite, but since the size of this population will usually be quite large, the finite population correction will be neglected in our calculations.

When samples from the nine grid cells have been selected, they are combined and mixed. The resulting concentration is denoted as:

$$\tilde{X}_{ij} = \frac{1}{9} \sum_{l=1}^9 X_{ijl} = \text{average concentration among the mixed grid samples.}$$

Notice that

$$\begin{aligned}
E(\tilde{X}_{ij} | S_1) &= E\left(\frac{1}{9} \sum_{l=1}^9 X_{ijl} | S_1\right) \\
&= \frac{1}{9} \sum_{l=1}^9 E(X_{ijl} | S_1) \\
&= \frac{1}{9} \sum_{l=1}^9 \theta_{il} \\
&= \theta_i \\
&= \text{actual average concentration throughout all grid cells at} \\
&\quad \text{time } i.
\end{aligned}$$

(This calculation assumes implicitly that the grids are constructed so that each cell has an equal volume of material.) Furthermore, because the sampling is done independently in each cell,

$$V(\tilde{X}_{ij} | S_1) = \frac{1}{81} \sum_{l=1}^9 V(X_{ijl} | S_1) = \frac{1}{81} \sum_{l=1}^9 \sigma_{il}^2 \equiv \sigma_i^2.$$

This term represents the variance created by testing only selected samples within each cell, rather than testing all material in each cell.

As discussed above, compositing may also be used. We will index the composite samples as ij' , denoting the mean concentration in the ij' -th composite as $\bar{X}_{ij'}$. The mean concentration in the ij' -th composite sample can be written as follows:

$$\bar{X}_{ij'} = \sum_{j=c(j'-1)+1}^{cj'} \tilde{X}_{ij} / c.$$

For example, if $c = 2$ and $n = 8$, then $j = 1, 2, \dots, 8$, and $j' = 1, 2, \dots, 4$. In particular, for $j' = 2$,

$$\bar{X}_{i2} = \sum_{j=2(2-1)+1}^{(2)(2)} \tilde{X}_{ij} / 2 = \sum_{j=3}^4 \tilde{X}_{ij} / 2 = \frac{\tilde{X}_{i3} + \tilde{X}_{i4}}{2}.$$

Notice that this notation includes the case where $c = 1$, i.e., where there is no compositing.

This notation implies that the compositing is done only within time periods, since i indexes time periods and j indexes samples within time periods. However, samples over time can also be composited. The notation above could be adapted by allowing i' to index composite samples over time. However, we recommend that samples over time be composited only when $n = 1$. In this case, it is mathematically equivalent to set $i = 1$ and to let j index the samples over time. Thus we will treat only the case shown above where j' indexes composite samples.

The final stage of sampling involves forming and selecting replicates. Denote the concentration in the k -th replicate as $\bar{X}_{ij'k}$. Assuming that the selection of replicates is random,

$$E(\bar{X}_{ij'k} | S_1, S_2) = \bar{X}_{ij'}.$$

(Notice that this expectation is conditional on both the selection of time periods and the samples within grid cells.) The notation

$$\gamma_{1ij'}^2 = E\left((\bar{X}_{ij'k} - \bar{X}_{ij'})^2 | S_1, S_2\right)$$

will be convenient for indicating the variance between replicates.

Measurement Error. When samples of material are submitted to laboratory analysis, some measurement error is inevitable. Thus it is reasonable to let

$$\begin{aligned} Y_{ij'k} &= \text{measured concentration in the } k\text{-th replicate of the } ij'\text{-th composite sample} \\ &= (\text{Actual Concentration}) + (\text{Measurement Error}) \\ &= \bar{X}_{ij'k} + \varepsilon_{ij'k} \end{aligned}$$

where

$$E(\varepsilon_{ij'k}) = 0$$

$$V(\varepsilon_{ij'k}) = \gamma_2^2$$

and where $\bar{X}_{ij'k}$ and $\varepsilon_{ij'k}$ are assumed to be independent.

Total Variance. As discussed earlier, there are three stages of sampling, consisting of m time periods, n/c composite samples, and r replicates. To estimate the overall concentration, θ , the average measured concentration among all samples is calculated. The overall average measured concentration can be written as

$$\bar{\bar{Y}} = \frac{1}{m} \frac{c}{n} \frac{1}{r} \sum_i \sum_{j'} \sum_k Y_{ij'k}.$$

Notice that

$$\begin{aligned} (1) \quad E(\bar{\bar{Y}}) &= E\left(E\left(\bar{\bar{Y}} \mid S_1, S_2, S_3\right)\right) \\ &= E\left(E\left(\frac{1}{m} \frac{c}{n} \frac{1}{r} \sum_i \sum_{j'} \sum_k Y_{ij'k} \mid S_1, S_2, S_3\right)\right) \\ &= E\left(\frac{1}{m} \frac{c}{n} \frac{1}{r} \sum_i \sum_{j'} \sum_k E\left(Y_{ij'k} \mid S_1, S_2, S_3\right)\right) \\ &= E\left(\frac{1}{m} \frac{c}{n} \frac{1}{r} \sum_i \sum_{j'} \sum_k E\left(\bar{X}_{ij'k} + \varepsilon_{ij'k} \mid S_1, S_2, S_3\right)\right). \end{aligned}$$

Conditional on S_1, S_2 , and S_3 – that is, conditional on the time periods, grid samples and replicates that were selected – $\bar{X}_{ij'k}$ is constant and $\varepsilon_{ij'k}$ has an expected value of zero, independent of S_1, S_2 , and S_3 . Thus

$$E\left(\bar{X}_{ij'k} + \varepsilon_{ij'k} \mid S_1, S_2, S_3\right) = \bar{X}_{ij'k} + E\left(\varepsilon_{ij'k}\right) = \bar{X}_{ij'k}.$$

It follows that

$$E(\bar{\bar{Y}}) = E\left(\frac{1}{m} \frac{c}{n} \frac{1}{r} \sum_i \sum_{j'} \sum_k \bar{X}_{ij'k}\right)$$

$$\begin{aligned}
&= E\left(E\left(\frac{1}{m} \frac{c}{n} \frac{1}{r} \sum_i \sum_{j'} \sum_k \bar{X}_{ij'k} \mid S_1, S_2\right)\right) \\
&= E\left(\frac{1}{m} \frac{c}{n} \sum_i \sum_{j'} E\left(\frac{1}{r} \sum_k \bar{X}_{ij'k} \mid S_1, S_2\right)\right).
\end{aligned}$$

Now, conditional on S_1 and S_2 , as discussed earlier,

$$E\left(\frac{1}{r} \sum_k \bar{X}_{ij'k} \mid S_1, S_2\right) = \frac{1}{r} \sum_k E\left(\bar{X}_{ij'k} \mid S_1, S_2\right) = \frac{1}{r} \sum_k \bar{X}_{ij'} = \bar{X}_{ij'}$$

so that

$$\begin{aligned}
E(\bar{\bar{Y}}) &= E\left(\frac{1}{m} \frac{c}{n} \sum_i \sum_{j'} \bar{X}_{ij'}\right) \\
&= E\left(\frac{1}{m} \frac{c}{n} \sum_i \sum_{j'} \sum_{j=c(j'-1)+1}^{cj'} \tilde{X}_{ij'/c}\right)
\end{aligned}$$

Notice however, that

$$\sum_{j'} \sum_{j=c(j'-1)+1}^{cj'} \tilde{X}_{ij'/c} = \sum_j \tilde{X}_{ij}$$

so that

$$\begin{aligned}
E(\bar{\bar{Y}}) &= E\left(\frac{1}{m} \frac{1}{n} \sum_i \sum_j \tilde{X}_{ij}\right) \\
&= E\left(E\left(\frac{1}{m} \frac{1}{n} \sum_i \sum_j \tilde{X}_{ij} \mid S_1\right)\right) \\
&= E\left(\frac{1}{m} \frac{1}{n} \sum_i \sum_j E\left(\tilde{X}_{ij} \mid S_1\right)\right) \\
&= E\left(\frac{1}{m} \sum_i \theta_i\right) \\
&= \theta.
\end{aligned}$$

(The validity of the last step depends on the assumption that the time periods are randomly selected.) This shows that $\bar{\bar{Y}}$ is an unbiased estimator of θ , and thus the notation

$$\hat{\theta} = \bar{\bar{Y}}$$

will be used in the sections which follow.

Using the conditional expectation formula for variances,

$$\begin{aligned} V(\hat{\theta}) &= V(\bar{\bar{Y}}) \\ &= E\left(V(\bar{\bar{Y}} \mid S_1, S_2, S_3)\right) + V\left(E(\bar{\bar{Y}} \mid S_1, S_2, S_3)\right) \\ &= E\left(\left(\frac{1}{m} \frac{c}{n}\right)^2 \sum_i \sum_j V\left(\frac{1}{r} \sum_k (\bar{X}_{ij \cdot k} + \varepsilon_{ij \cdot k}) \mid S_1, S_2, S_3\right)\right) \\ &\quad + V\left(\frac{1}{m} \frac{c}{n} \sum_i \sum_j E\left(\frac{1}{r} \sum_k (\bar{X}_{ij \cdot k} + \varepsilon_{ij \cdot k}) \mid S_1, S_2, S_3\right)\right). \end{aligned}$$

Conditional on S_1, S_2 , and S_3 , $\bar{X}_{ij \cdot k}$ is constant, while $\varepsilon_{ij \cdot k}$ has mean 0 and variance γ_2^2 , neither of which depends on S_1, S_2 , or S_3 . Thus

$$\begin{aligned} V\left(\frac{1}{r} \sum_k (\bar{X}_{ij \cdot k} + \varepsilon_{ij \cdot k}) \mid S_1, S_2, S_3\right) &= \frac{1}{r^2} \sum_k V\left(\bar{X}_{ij \cdot k} + \varepsilon_{ij \cdot k} \mid S_1, S_2, S_3\right) \\ &= \frac{\gamma_2^2}{r} \end{aligned}$$

and

$$\begin{aligned} E\left(\frac{1}{r} \sum_k (\bar{X}_{ij \cdot k} + \varepsilon_{ij \cdot k}) \mid S_1, S_2, S_3\right) &= \frac{1}{r} \sum_k E\left(\bar{X}_{ij \cdot k} + \varepsilon_{ij \cdot k} \mid S_1, S_2, S_3\right) \\ &= \frac{1}{r} \sum_k \bar{X}_{ij \cdot k}. \end{aligned}$$

Combining these expressions, we have

$$(2) \quad V(\hat{\theta}) = E\left(\left(\frac{1}{m} \frac{c}{n}\right)^2 \sum_i \sum_j \frac{\gamma_2^2}{r}\right) + V\left(\frac{1}{m} \frac{c}{n} \sum_i \sum_j \frac{1}{r} \sum_k \bar{X}_{ij \cdot k}\right)$$

$$\begin{aligned}
&= \frac{1}{m} \frac{c}{n} \frac{1}{r} \gamma_2^2 + E \left(\left(\frac{1}{m} \frac{c}{n} \right)^2 \sum_i \sum_{j'} V \left(\frac{1}{r} \sum_k \bar{X}_{ij'k} \mid S_1, S_2 \right) \right) \\
&\quad + V \left(\frac{1}{m} \frac{c}{n} \sum_i \sum_{j'} E \left(\frac{1}{r} \sum_k \bar{X}_{ij'k} \mid S_1, S_2 \right) \right) \\
&= \frac{1}{m} \frac{c}{n} \frac{1}{r} \gamma_2^2 + E \left(\left(\frac{1}{m} \frac{c}{n} \right)^2 \sum_i \sum_{j'} \frac{\gamma_{1ij'}^2}{r} \right) + V \left(\frac{1}{m} \frac{c}{n} \sum_i \sum_{j'} \bar{X}_{ij'} \right) \\
&= \frac{1}{m} \frac{c}{n} \frac{1}{r} \gamma_2^2 + E \left(\left(\frac{1}{m} \frac{c}{n} \right)^2 \sum_i \sum_{j'} \frac{\gamma_{1ij'}^2}{r} \right) + V(\bar{\bar{X}})
\end{aligned}$$

where

$$\bar{\bar{X}} = \frac{1}{m} \frac{c}{n} \sum_i \sum_{j'} \bar{X}_{ij'}$$

Similarly,

$$\begin{aligned}
V(\bar{\bar{X}}) &= E(V(\bar{\bar{X}} \mid S_1)) + V(E(\bar{\bar{X}} \mid S_1)) \\
&= E \left(\left(\frac{1}{m} \right)^2 \sum_i V \left(\frac{c}{n} \sum_{j'} \bar{X}_{ij'} \mid S_1 \right) \right) + V \left(\frac{1}{m} \sum_i E \left(\frac{c}{n} \sum_{j'} \bar{X}_{ij'} \mid S_1 \right) \right)
\end{aligned}$$

Since the mean concentration of the ij' -th composite sample is

$$\bar{X}_{ij'} = \sum_{j=c(j'-1)+1}^{cj'} \tilde{X}_{ij}/c,$$

we have

$$\frac{c}{n} \sum_{j'} \bar{X}_{ij'} = \frac{c}{n} \sum_{j'} \sum_{j=c(j'-1)+1}^{cj'} \tilde{X}_{ij}/c = \frac{1}{n} \sum_j \tilde{X}_{ij}.$$

Thus

$$\begin{aligned}
V(\bar{\bar{X}}) &= E \left(\left(\frac{1}{m} \right)^2 \sum_i V \left(\frac{1}{n} \sum_j \tilde{X}_{ij} \mid S_1 \right) \right) + V \left(\frac{1}{m} \sum_i E \left(\frac{1}{n} \sum_j \tilde{X}_{ij} \mid S_1 \right) \right) \\
&= E \left(\left(\frac{1}{m} \right)^2 \sum_i \sigma_i^2/n \right) + V \left(\frac{1}{m} \sum_i \theta_i \right).
\end{aligned}$$

Setting

$$\tau^2 = E(\theta_i - \bar{\theta})^2$$

(where $\bar{\theta} = \frac{1}{m} \sum_i \theta_i$) and

$$\sigma^2 = E(\sigma_i^2)$$

yields

$$V(\bar{X}) = \frac{\sigma^2}{mn} + \frac{\tau^2}{m}.$$

The first term represents the variability due to sampling over space, while the second represents the variability due to sampling over time. Combining this expression with (2) yields

$$V(\hat{\theta}) = \frac{\tau^2}{m} + \frac{\sigma^2}{mn} + E\left(\left(\frac{1}{m} \frac{c}{n}\right)^2 \sum_i \sum_{j'} \frac{\gamma_{1ij'}^2}{r}\right) + \frac{1}{m} \frac{c}{n} \frac{1}{r} \gamma_2^2.$$

Assuming that the variance between replicates, γ_{1ij}^2 , is relatively constant over time and between composite samples, i.e., $\gamma_{1ij'}^2 = \gamma_1^2$, leads to the simplification

$$\begin{aligned} (3) \quad V(\hat{\theta}) &= \frac{\tau^2}{m} + \frac{\sigma^2}{mn} + E\left(\left(\frac{1}{m} \frac{c}{n}\right)^2 \sum_i \sum_{j'} \frac{\gamma_1^2}{r}\right) + \frac{1}{m} \frac{c}{n} \frac{1}{r} \gamma_2^2 \\ &= \frac{\tau^2}{m} + \frac{\sigma^2}{mn} + \frac{1}{m} \frac{c}{n} \frac{1}{r} \gamma_1^2 + \frac{1}{m} \frac{c}{n} \frac{1}{r} \gamma_2^2 \\ &= \frac{\tau^2}{m} + \frac{\sigma^2}{mn} + \frac{c}{n} \frac{\gamma^2}{mr} \end{aligned}$$

where

$$\gamma^2 = \gamma_1^2 + \gamma_2^2.$$

This expression for the variance of the overall average sample concentration includes terms for the variability over time (τ^2/m), the variability between samples within

grid cells (σ^2/mn), and the variability due to measurement errors (γ^2c/nmr). In most cases, the between-time component (τ^2) is probably larger than variability of samples within grid cells (σ^2). One exception would be in sampling stored fluff, where in most cases there would be no between-time component (i.e., $m = 1$) and formula (3) would reduce to

$$V(\hat{\theta}) = \frac{\sigma^2}{n} + \frac{c}{n} \frac{\gamma^2}{r}.$$

On the other hand, when sampling from fresh output streams, it will generally be the case that $n = 1$, in which case formula (3) reduces to

$$V(\hat{\theta}) = \frac{\tau^2 + \sigma^2}{m} + \frac{\gamma^2}{mr}.$$

In this case, however, there may be compositing over time, where the m samples are grouped at random into q ($= m/c$) samples for analysis. Finally, it would generally be expected that $r = 1$, except when unusual concentrations are found. These considerations would lead to the variance formula

$$(4) \quad V(\hat{\theta}) = \frac{\tau^2 + \sigma^2}{m} + \frac{\gamma^2}{q}.$$

In fact, as we will see in Section B.3.1, our preliminary estimates of sampling variability for stored fluff are somewhat lower than those for fresh fluff, so that assuming the model in (4) is probably conservative. Moreover, based on the available data, we are not able to obtain distinct estimates for τ^2 and σ^2 which are sufficiently accurate. For these two reasons, we will assume the model in (4) for standard error calculations.

3. Standard Error Calculations

3.1 Data Sources

Data for estimating the components of the variance model in (4) were available from two sources. First, a pilot study for a proposed national survey of shredder sites gathered samples of fluff and other materials at seven shredder sites distributed across

the United States.* These samples were collected under controlled, monitored circumstances, with adequate quality control and documentation on the methods used in collecting the data. In addition, other data on fluff were available from several states where samples of fluff had been collected and analyzed. Unfortunately, for the latter data, no information is available on the sampling methods or on the types of materials sampled. Both sources of data are summarized in Table 1, which gives the averages, standard deviations, and coefficients of variation (CV's) found in various types of material.

The EPA data show exploratory research on PCB levels in fluff from shredded automobiles and white goods, as well as stored fluff, soil, and other materials. The data presented include estimated standard deviations for analytical error and sampling error, and the estimated overall average PCB concentrations. For example, for automobiles, the estimated standard deviation for analytical error is 9.28, indicating that

$$\hat{\gamma} = 9.28$$

might be used as an estimate of the standard deviation for measurement error in analyzing fluff from automobiles. In the EPA data, the standard deviations for analytical error range are 9.28 to 19.27, while the standard deviations for sampling error range from 23.90 to 155.37. In the State data, no standard deviations are available for analytical error, and the standard deviations for sampling error, which range from 22.32 to 345.55, are not identified by the type of material.

In the EPA study, procedures similar to those described in the *Sampling Guidance* were used to collect multiple samples of stored fluff, soil, and fresh fluff. In sampling fresh fluff, however, only one sample was taken for each time period, so that the sampling error term includes both τ^2 and σ^2 , as indicated in formula (4) shown in Section 2. Thus it is not possible to make separate estimates of the parameters τ^2 and σ^2 . Another source of confounding is that the estimates of CV's in the "Sampling Error" column *also* contain measurement error. Thus the estimates for standard errors and CV's have a slight upward bias, although this bias amounts to only a few percentage points at most.

* USEPA Fluff Pilot Program, conducted by Westat, Inc., Rockville, MD, and Midwest Research Institute, Kansas City, MO, for the Office of Pollution Prevention and Toxics (see *PCB, Lead, and Cadmium Levels in Shredder Waste Materials: A Pilot Study*. USEPA, Office of Toxic Substances. EPA 560/5-90-008B. 1991).

Table 1: Summary of preliminary data

Data from EPA Pilot Study

<u>Type of Material</u>	<u>Overall PCB Level</u>	<u>Analytical Error*</u>		<u>Sampling Error</u>		<u>Number of Samples</u>
		<u>Standard Deviation</u>	<u>CV</u>	<u>Standard Deviation</u>	<u>CV</u>	
Automobile Fluff	24.64	9.28	0.30	23.90	0.97	37
White Goods Fluff	57.16	19.27	0.34	71.45	1.25	21
Other Fresh Fluff	176.56	–	–	155.37	0.88	9
Stored Fluff	77.13	–	–	47.82	0.62	8
Soil	44.14	–	–	41.05	0.93	8
Overall	55.69	17.82	0.32	57.36	1.03	83

Data from State Samples

<u>State</u>	<u>Overall PCB Level</u>	<u>Analytical Error*</u>		<u>Sampling Error</u>		<u>Number of Samples</u>
		<u>Standard Deviation</u>	<u>CV</u>	<u>Standard Deviation</u>	<u>CV</u>	
Massachusetts	33.85	–	–	29.45	0.87	80
Rhode Island	21.52	–	–	25.18	1.17	80
Maryland	228.84	–	–	345.55	1.51	12
California	58.73	–	–	22.32	0.38	13
Indiana	87.13	–	–	29.62	0.34	8
Arizona	190.50	–	–	72.39	0.38	12

*Estimates for analytical error were based on repeated analysis of eight samples of automobile fluff and four samples of white goods fluff.

Notice that the larger the estimated overall PCB level, the higher the standard deviation. This phenomenon is frequently observed in sampling and it is often taken into account by using CV rather than the standard deviation as a measure of variability. For a random variable x , the coefficient of variation is generally defined as

$$CV = \frac{\sqrt{V(x)}}{E(x)}.$$

(See, for example, Snedecor and Cochran, *Statistical Methods*, Iowa State University Press, Sixth Edition, 1967, p. 62.) In this case, we will be interested in the coefficient of variation for analytical error, which will be defined as

$$CV_{\gamma} = \frac{\gamma}{\theta}.$$

The data in Table 1 show estimated CV's for measurement error for automobile and white goods as .30 and .34, respectively. As an overall figure, we might estimate the CV for measurement error as

$$\hat{CV}_{\gamma} = \frac{\hat{\gamma}}{\hat{\theta}} = \frac{17.82}{55.69} \approx .32.$$

This estimate for CV_{γ} is shown for "Overall" in the "Analytical Error CV" column of Table 1.

CV's can also be estimated for sampling error. In this case we are interested in

$$CV_{\tau\sigma} = \frac{\sqrt{\tau^2 + \sigma^2}}{\theta}$$

since this gives the standard error for sampling relative to the overall PCB concentration. Table 1 gives estimates $\hat{CV}_{\tau\sigma}$ for sampling various materials, as well as for the state data. These values range from .62 to 1.25, with an overall CV of 1.03.

The estimates of variances that were used for calculating sample sizes and power were based on the numbers in Table 1. In all cases, estimates were rounded off. Because of the tendency toward over-estimation discussed above, estimates for CV's were rounded down somewhat. In all cases, estimates are based on the numbers shown in the first and second rows in Table 1. The estimate for CV_γ was obtained by averaging the CV's for automobile and white goods; the estimates for overall PCB levels and for CV's for sampling error were based directly on the respective entries in Table 1. With that in mind, the following estimates were assumed for use in standard error calculations:

- \hat{CV}_γ (CV for analytical error): 32%
- $\hat{CV}_{\tau 1}$ (CV for sampling error) for automobiles: 100%
- Overall PCB level for automobiles: 25 ppm
- $\hat{CV}_{\tau 2}$ (CV for sampling error) for white goods: 100%
- Overall PCB level for white goods: 57 ppm.

In the next section we will discuss in more detail how these numerical values were used.

3.2 Numerical Methods

Mixture Distributions. A *mixture distribution* occurs when two different quantities are combined at random. For example, in the output at shredder sites, fluff from shredding automobiles and white goods is combined. One model for this process can be formulated as follows. Let U , Z_1 , and Z_2 be independent random variables, where U takes the values 1 and 0, with mean and variance

$$E(U) = \pi \quad V(U) = \pi(1 - \pi),$$

and where

$$E(Z_i) = \mu_i \quad V(Z_i) = \eta_i^2.$$

If we define W as

$$W = UZ_1 + (1 - U)Z_2$$

then the distribution of W is a *mixture* of the distributions of Z_1 and Z_2 . The mean and variance of W are given by

$$(5) \quad E(W) = \pi\mu_1 + (1 - \pi)\mu_2$$

and

$$\begin{aligned} (6) \quad V(W) &= V(UZ_1 + (1 - U)Z_2) \\ &= E\left(V(UZ_1 + (1 - U)Z_2 | U)\right) + V\left(E(UZ_1 + (1 - U)Z_2 | U)\right) \\ &= E\left(U^2\eta_1^2 + (1 - U)^2\eta_2^2\right) + V\left(U\mu_1 + (1 - U)\mu_2\right) \\ &= \pi\eta_1^2 + (1 - \pi)\eta_2^2 + V\left(\mu_2 + U(\mu_1 - \mu_2)\right) \\ &= \pi\eta_1^2 + (1 - \pi)\eta_2^2 + \pi(1 - \pi)(\mu_1 - \mu_2)^2. \end{aligned}$$

This model can be applied to samples of shredder output by letting

$$U = \begin{cases} 1 & \text{if the sample was generated by shredding white goods} \\ 0 & \text{if the sample was generated by shredding automobiles} \end{cases}$$

and letting Z_1 and Z_2 represent the distributions of PCB levels in samples of white goods and automobile fluff, respectively. Thus, in this model, π is the percent of white goods that is mixed with automobiles, while μ_1 and μ_2 are the average PCB concentrations in white goods and automobiles, respectively. Similarly, η_1^2 and η_2^2 represent the variability of PCB levels in white goods and automobiles.

Combining the notation in Section 2 with formula (5), the overall concentration of PCBs would be denoted as

$$(7) \quad \theta = \pi\mu_1 + (1 - \pi)\mu_2.$$

To obtain a formula for $V(\hat{\theta})$, remember that, from (4),

$$V(\hat{\theta}) = \frac{\tau^2 + \sigma^2}{m} + \frac{\gamma^2}{q}.$$

As discussed earlier, the “ $\tau^2 + \sigma^2$ ” term represents sampling variability, while the γ^2 term represents measurement error. Since, in formula (6), $V(W)$ represents sampling error only, we have

$$\tau^2 + \sigma^2 = \pi\eta_1^2 + (1-\pi)\eta_2^2 + \pi(1-\pi)(\mu_1 - \mu_2)^2$$

which yields

$$(8) \quad V(\hat{\theta}) = \frac{\pi\eta_1^2 + (1-\pi)\eta_2^2 + \pi(1-\pi)(\mu_1 - \mu_2)^2}{m} + \frac{\gamma^2}{q}.$$

Recall from our discussion of analytical error that γ^2 appears to be larger when the overall PCB level is larger. On the other hand, the coefficient of variation, CV_γ appears to be relatively constant. Notice that

$$CV_\gamma = \frac{\gamma}{\theta}$$

implies

$$(9) \quad \gamma^2 = CV_\gamma^2 \theta^2 = CV_\gamma^2 (\pi\mu_1 + (1-\pi)\mu_2)^2.$$

This gives us an alternative formula for variability due to measurement error, one which reflects the change in γ^2 relative to the change in the overall concentration of PCBs.

Combining formulas (4), (8) and (9), we have the following expression:

$$(10) \quad V(\hat{\theta}) = \frac{\pi\eta_1^2 + (1-\pi)\eta_2^2 + \pi(1-\pi)(\mu_1 - \mu_2)^2}{m} + \frac{CV_\gamma^2 (\pi\mu_1 + (1-\pi)\mu_2)^2}{q}.$$

Formula (10) was used for estimating variances for calculating relative errors, while a simpler variation of it was used for power calculations. Since the actual parameter values are not known, estimates were derived based on the data in Table 1. For example, the CV for analytical error and the overall PCB concentrations were estimated to be

$$\hat{CV}_\gamma = 32\%$$

$$\hat{\mu}_1 = 57 \text{ ppm}$$

$$\hat{\mu}_2 = 25 \text{ ppm}$$

although μ_1 and μ_2 were sometimes allowed to vary (in power calculations, for example). The values for η_1^2 and η_2^2 were derived from the CV's shown on page 15:

$$\hat{\eta}_1 = \hat{CV}_{\tau\sigma 1} \hat{\mu}_1 = (1.00)(57) = 57$$

$$\hat{\eta}_2 = \hat{CV}_{\tau\sigma 2} \hat{\mu}_2 = (1.00)(25) = 25.$$

The value for π was either taken to be 10% or allowed to vary between 0 to 100%.

4. Confidence Intervals and Relative Error

4.1 Numerical Methods

In the *Field Manual*, Section 4.2 discusses confidence intervals and relative error; Table 2 in the *Field Manual* shows the relative errors for various sample sizes and compositing strategies. Here, Table 2 shows the relative errors for various mixtures of white goods and automobiles. Note that Table 2 in the *Field Manual* corresponds to the "25% White Goods" column. In this section, we will discuss how the relative errors were calculated.

Table 2: Relative errors for estimating PCB levels with sample sizes of 2 to 25

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Relative error*				
			Percent white goods				
			0%	10%	25%	50%	85%
2	2	1	944%	1038%	1084%	1066%	984%
4	4		167%	184%	192%	189%	174%
9	9		81%	89%	93%	91%	84%
16	16		56%	62%	64%	63%	58%
25	25		43%	48%	50%	49%	45%
4	2	2	698%	762%	793%	781%	725%
8	4		123%	135%	140%	138%	128%
18	9		60%	65%	68%	67%	62%
32	16		41%	45%	47%	46%	43%
50	25		32%	35%	36%	36%	33%
8	2	4	534%	576%	597%	588%	551%
16	4		94%	102%	106%	104%	98%
36	9		46%	49%	51%	50%	47%
64	16		32%	34%	35%	35%	33%
100	25		24%	26%	27%	27%	25%
16	2	8	429%	455%	468%	463%	440%
32	4		76%	81%	83%	82%	78%
72	9		37%	39%	40%	40%	38%
128	16		25%	27%	28%	27%	26%
200	25		20%	21%	21%	21%	20%

*A relative error of 50% means that with 95% certainty, the estimated average concentration will be within 50% of the actual average concentration.

The formula for a confidence interval for θ is given by

$$(11) \quad \hat{\theta} \pm t_{\alpha/2} \sqrt{V(\hat{\theta})}$$

where $t_{\alpha/2}$ is the percentile corresponding to $1-\alpha/2$ of a Student's t -distribution with $(q-1)$ degrees of freedom. Since sample sizes will be small in most cases, the t -distribution is more appropriate than the standard normal. Unfortunately, the distribution of PCB levels in samples of fluff is highly skewed and, as a result, the normal distribution gives only a rough approximation to the actual coverage probabilities of confidence intervals. Based on simulations done with data from the EPA Pilot Study, 95% confidence intervals may be accurate as little as 75% of the time with sample sizes of 15 to 25. In nearly all cases, when confidence intervals are incorrect they fall *below* the actual PCB level.

In any case, the relative error is given by

$$(12) \quad \text{Relative Error} = \frac{t_{\alpha/2} \sqrt{V(\hat{\theta})}}{\theta}.$$

Combining formulas (7), (10), and (12), we have

$$(13) \quad \text{Relative Error} =$$

$$\frac{t_{\alpha/2}}{\pi\mu_1 + (1-\pi)\mu_2} \sqrt{\frac{\pi\eta_1^2 + (1-\pi)\eta_2^2 + \pi(1-\pi)(\mu_1 - \mu_2)^2}{m} + \frac{(\text{CV}_\gamma(\pi\mu_1 + (1-\pi)\mu_2))^2}{q}}.$$

As an example calculation, suppose that $\pi = .10$, $c = 2$, $m = 8$, and $\alpha = .05$ (i.e., the mixture of white goods is 10%, there are 8 total samples consisting of 4 composites of 2 subsamples each, and the confidence level is 95%). Notice from Table 3, $t_{\alpha/2} = 3.18$. For the other parameters, we will use the estimates discussed in Section 3.2:

$$\hat{\text{CV}}_\gamma = 32\%$$

$$\hat{\mu}_1 = 57$$

$$\hat{\mu}_2 = 25$$

Table 3: t-values for confidence intervals and hypothesis tests

Number of composite samples $q = m/c$	t-values					
	Confidence intervals			Hypothesis tests		
	90%	95%	99%	90%	95%	99%
2	6.31	12.71	63.66	3.08	6.31	31.82
3	2.92	4.30	9.93	1.89	2.92	6.97
4	2.35	3.18	5.84	1.64	2.35	4.54
5	2.13	2.78	4.60	1.53	2.13	3.75
6	2.02	2.57	4.03	1.48	2.02	3.37
7	1.94	2.45	3.71	1.44	1.94	3.14
8	1.89	2.36	3.50	1.42	1.89	3.00
9	1.86	2.31	3.35	1.40	1.86	2.90
10	1.83	2.26	3.25	1.38	1.83	2.82
11	1.81	2.23	3.17	1.37	1.81	2.76
12	1.80	2.20	3.11	1.36	1.80	2.72
13	1.78	2.18	3.05	1.36	1.78	2.68
14	1.77	2.16	3.01	1.35	1.77	2.65
15	1.76	2.15	2.98	1.35	1.76	2.63
16	1.75	2.13	2.95	1.34	1.75	2.60
17	1.75	2.12	2.92	1.34	1.75	2.58
18	1.74	2.11	2.90	1.33	1.74	2.57
19	1.73	2.10	2.88	1.33	1.73	2.55
20	1.73	2.09	2.86	1.33	1.73	2.54
21	1.73	2.09	2.85	1.33	1.73	2.53
22	1.72	2.08	2.83	1.32	1.72	2.52
23	1.72	2.07	2.82	1.32	1.72	2.51
24	1.71	2.07	2.81	1.32	1.71	2.50
25	1.71	2.06	2.80	1.32	1.71	2.49
30	1.70	2.05	2.76	1.31	1.70	2.46
50	1.68	2.01	2.68	1.30	1.68	2.41
75	1.67	1.99	2.64	1.29	1.67	2.38
100	1.66	1.98	2.63	1.29	1.66	2.37
>100	1.65	1.96	2.58	1.28	1.65	2.33

$$\hat{\eta}_1 = 57 \quad \hat{\eta}_2 = 25.$$

Substituting these estimates for their respective parameter values in formula (13) gives a relative error of

$$\begin{aligned} \text{Relative Error} &= \frac{3.18}{28.2} \sqrt{\frac{324.9 + 562.5 + 92.16}{8} + \frac{81.43}{4}} \\ &= 135\% \end{aligned}$$

as shown in Table 2.

Notice that the Relative Error depends on the parameters α , m , and on c as follows:

- As α increases, $t_{\alpha/2}$ decreases, making the Relative Error smaller;
- As m increases, $V(\hat{\theta})$ decreases, making the Relative Error smaller;
- As c increases, $V(\hat{\theta})$ increases, making the Relative Error larger.

In the latter cases, the sample size is altered and the sampling and/or measurement errors actually change. When α changes, however, the sampling and measurement errors remain the same, but the possibility of error in our conclusions changes. For example, suppose that the level of confidence is changed from 95% to 90% – i.e., α changes from .05 to .10. This means that we are willing to accept a 10% chance of being wrong rather than a 5% chance. In exchange for this sacrifice, both $t_{\alpha/2}$ and the resulting confidence interval will be smaller, so that we can make a closer estimate the actual level of PCBs. Thus the Relative Error, as we have defined it in formula (12), is smaller, although the chance that this error occurs has increased.

4.2 Alternative Confidence Levels

Tables 4 and 5 show what the relative errors would be under the alternative confidence levels of 90% and 99%. When the confidence level is decreased from 95% to 90%, the decrease in relative error is quite striking for very small sample sizes (e.g., $2 \leq q \leq 4$), but relatively modest for larger sample sizes (e.g., $16 \leq q \leq 25$). Similarly, increasing the level of confidence from 95% to 99% sharply increases the relative errors when q is small, but the increase is much smaller when the sample sizes are larger.

5. Hypothesis Testing

5.1 Monitoring

To find cut-off values in monitoring, we employ the standard statistical procedures for testing the hypothesis

$$H_0: \mu \leq \mu_0$$

against the alternative

$$H_a: \mu > \mu_0,$$

where μ_0 is taken to be 25, 50 or 100. That is, we compare the standardized statistic

$$t^* = \frac{\bar{x} - \mu_0}{s/\sqrt{m}}$$

with $t_{.05}$, the 95-th percentile of the t -distribution with $m-1$ degrees of freedom, rejecting H_0 if $t^* > t_{.05}$. For simplicity, it is suggested that the user implement this procedure by comparing \bar{x} with a cut-off value given by

$$(14) \quad \text{Cut-off Value} = \mu_0 + t_{.05} \frac{s}{\sqrt{q}}.$$

Table 4: Relative errors for estimating PCB levels with sample sizes of 2 to 25 and 90% confidence

Total samples collected <i>m</i>	Number of composites analyzed <i>m/c</i>	Subsamples in each composite <i>c</i>	Relative error*				
			Percent white goods				
			0%	10%	25%	50%	85%
2	2	1	469%	516%	539%	530%	489%
4	4		124%	136%	142%	140%	129%
9	9		65%	72%	75%	74%	68%
16	16		46%	51%	53%	52%	48%
25	25		36%	40%	41%	41%	37%
4	2	2	347%	378%	394%	388%	360%
8	4		91%	100%	104%	102%	95%
18	9		48%	53%	55%	54%	50%
32	16		34%	37%	39%	38%	35%
50	25		27%	29%	30%	30%	28%
8	2	4	265%	286%	296%	292%	274%
16	4		70%	75%	78%	77%	72%
36	9		37%	40%	41%	41%	38%
64	16		26%	28%	29%	29%	27%
100	25		20%	22%	23%	22%	21%
16	2	8	213%	226%	233%	230%	218%
32	4		56%	60%	61%	61%	58%
72	9		30%	31%	32%	32%	30%
128	16		21%	22%	23%	23%	21%
200	25		16%	17%	18%	18%	17%

*A relative error of 50% means that with 95% certainty, the estimated average concentration will be within 50% of the actual average concentration.

Table 5: Relative errors for estimating PCB levels with sample sizes of 2 to 25 and 99% confidence

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Relative error*				
			Percent white goods				
			0%	10%	25%	50%	85%
2	2	1	4726%	5199%	5430%	5339%	4927%
4	4		307%	337%	352%	346%	320%
9	9		117%	129%	135%	133%	122%
16	16		77%	85%	89%	87%	81%
25	25		59%	65%	67%	66%	61%
4	2	2	3494%	3815%	3972%	3910%	3630%
8	4		227%	248%	258%	254%	236%
18	9		87%	95%	99%	97%	90%
32	16		57%	62%	65%	64%	59%
50	25		43%	47%	49%	49%	45%
8	2	4	2672%	2883%	2988%	2947%	2761%
16	4		173%	187%	194%	191%	179%
36	9		66%	72%	74%	73%	69%
64	16		44%	47%	49%	48%	45%
100	25		33%	36%	37%	37%	34%
16	2	8	2146%	2279%	2345%	2319%	2202%
32	4		139%	148%	152%	150%	143%
72	9		53%	57%	58%	58%	55%
128	16		35%	37%	38%	38%	36%
200	25		27%	28%	29%	29%	27%

*A relative error of 50% means that with 95% certainty, the estimated average concentration will be within 50% of the actual average concentration.

Alternatively, the cut-off values can be approximated by using the values in Table A-1 (reproduced here as Table 6). The values in this table were calculated using formula (14). For example, for $q = 9$, $s = 150$ and $\mu_0 = 50$, we have

$$\text{Cut-off Value} = 50 + 1.86 \frac{150}{\sqrt{9}} = 143.0$$

as shown in Table 6.

5.2 Clean-Up Verification

To find cut-off values for clean-up verification, we reverse the direction of the hypothesis test for monitoring. Here we want to test

$$H_0: \mu \geq \mu_0$$

against the alternative

$$H_a: \mu < \mu_0.$$

We reverse the direction in clean-up verification because it is known that the site, or output at the site, has been found to be contaminated. Thus, we assume that the site is contaminated until the data demonstrate otherwise.

In this case, the cut-off value is calculated as

$$(15) \quad \text{Cut-off Value} = \mu_0 - t_{.05} \frac{s}{\sqrt{m}}$$

and H_0 is rejected if \bar{x} is smaller than the cut-off value. Notice that when s is large, it may occur that

$$\mu_0 - t_{.05} \frac{s}{\sqrt{q}} < 0$$

in which case it will be impossible to reject the null hypothesis and further clean-up will be required.

Table 6: Cut-off values for monitoring*

Safety Standard	Standard Deviation	Number of Composite Samples Analyzed (q=m/c)				
		2	4	9	16	25
25	20	114.2	48.5	37.4	33.8	31.8
	35	181.2	66.1	46.7	40.3	37.0
	50	248.1	83.8	56.0	46.9	42.1
	75	359.6	113.1	71.5	57.8	50.7
	100	471.2	142.5	87.0	68.8	59.2
	150	694.3	201.3	118.0	90.6	76.3
	250	1,140.5	318.8	180.0	134.4	110.5
50	20	139.2	73.5	62.4	58.8	56.8
	35	206.2	91.1	71.7	65.3	62.0
	50	273.1	108.8	81.0	71.9	67.1
	75	384.6	138.1	96.5	82.8	75.7
	100	496.2	167.5	112.0	93.8	84.2
	150	719.3	226.3	143.0	115.6	101.3
	250	1,165.5	343.8	205.0	159.4	135.5
100	20	189.2	123.5	112.4	108.8	106.8
	35	256.2	141.1	121.7	115.3	112.0
	50	323.1	158.8	131.0	121.9	117.1
	75	434.6	188.1	146.5	132.8	125.7
	100	546.2	217.5	162.0	143.8	134.2
	150	769.3	276.3	193.0	165.6	151.3
	250	1,215.5	393.8	255.0	209.4	185.5

*If the average of the analyzed samples is larger than the cut-off value in the table, then conclude that the shredder output violates the given standard. Otherwise, assume that the output meets the standard. The chance of incorrectly finding a violation is 5%.

in which case it will be impossible to reject the null hypothesis and further clean-up will be required.

The cut-off values can be approximated by using the values in Table A-5 (reproduced here as Table 7). The values in this table were calculated using formula (15). For example, for $q = 4$, $s = 10$, and $\mu_0 = 50$, we have

$$\text{Cut-off Value} = 50 - 2.35 \frac{10}{\sqrt{4}} = 38.3$$

as shown in Table 7.

6. Power Calculations

6.1 Monitoring

Power Calculations. In hypothesis tests for monitoring, the power of the test against a specific alternative, say $\mu = \mu_a$, is given by

$$\begin{aligned} (16) \quad \text{Power} &= P(\bar{x} > \text{Cut-off Value} \mid \mu = \mu_a) \\ &= P\left(\bar{x} > \mu_0 + t_{.05} \frac{s}{\sqrt{q}} \mid \mu = \mu_a\right) \\ &= P\left(\frac{\bar{x} - \mu_a}{s/\sqrt{q}} > \frac{\mu_0 - \mu_a}{s/\sqrt{q}} + t_{.05}\right) \end{aligned}$$

To calculate estimates for power shown in Tables A-2 through A-4 (reproduced here as Tables 8 - 10), the probability in formula (16) was approximated by using the standard normal distribution. That is, we used the approximation

$$\text{Power} \approx 1 - \Phi\left(\frac{\mu_0 - \mu_a}{s/\sqrt{q}} + t_{.05}\right)$$

where Φ is the cumulative distribution function of the standard normal distribution.

Table 7: Cut-off values for clean-up verification

Safety Standard	Standard deviation	Number of composite samples analyzed (q=m/c)				
		2	4	9	16	25
25	10	—	13.3	18.8	20.6	21.6
	15	—	7.4	15.7	18.4	19.9
	20	—	1.5	12.6	16.3	18.2
	25	—	—	9.5	14.1	16.5
	35	—	—	3.3	9.7	13.0
	50	—	—	—	3.1	7.9
	65	—	—	—	—	2.8
50	10	5.4	38.3	43.8	45.6	46.6
	20	—	26.5	37.6	41.3	43.2
	30	—	14.8	31.4	36.9	39.7
	50	—	—	19.0	28.1	32.9
	60	—	—	12.8	23.8	29.5
	75	—	—	3.5	17.2	24.4
	125	—	—	—	—	7.3
100	15	33.1	82.4	90.7	93.4	94.9
	25	—	70.6	84.5	89.1	91.5
	50	—	41.3	69.0	78.1	82.9
	75	—	11.9	53.5	67.2	74.4
	100	—	—	38.0	56.3	65.8
	150	—	—	7.0	34.4	48.7
	250	—	—	—	—	14.5

*A dash (—) indicates that the standard deviation is too large to establish that the site is clean.

Table 8: Chance of finding violations in monitoring with a 25 ppm standard

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Chance of detecting violation*				
			Actual PCB concentration				
			30	35	40	50	60
2	2	1	0.00	0.00	0.00	0.00	0.00
4	4		0.02	0.04	0.05	0.08	0.11
9	9		0.08	0.15	0.22	0.33	0.42
16	16		0.13	0.25	0.37	0.56	0.68
25	25		0.18	0.36	0.53	0.75	0.86
4	2	2	0.00	0.00	0.00	0.00	0.00
8	4		0.03	0.05	0.08	0.14	0.20
18	9		0.11	0.22	0.34	0.53	0.65
32	16		0.19	0.39	0.57	0.79	0.89
50	25		0.26	0.55	0.76	0.93	0.98
8	2	4	0.00	0.00	0.00	0.00	0.00
16	4		0.04	0.08	0.14	0.25	0.35
36	9		0.15	0.34	0.51	0.75	0.86
64	16		0.26	0.57	0.78	0.95	0.99
100	25		0.38	0.76	0.93	0.99	1.00
16	2	8	0.00	0.00	0.00	0.00	0.00
32	4		0.05	0.12	0.22	0.40	0.54
72	9		0.21	0.48	0.69	0.90	0.96
128	16		0.36	0.74	0.92	0.99	1.00
200	25		0.51	0.90	0.99	1.00	1.00

*Power calculations assume a 5% chance of incorrectly finding a violation.

Table 9: Chance of finding violations in monitoring with a 50 ppm standard

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Chance of detecting violation*				
			Actual PCB concentration				
			60	70	85	100	125
2	2	1	0.00	0.00	0.00	0.00	0.00
4	4		0.02	0.04	0.06	0.08	0.11
9	9		0.08	0.15	0.25	0.33	0.44
16	16		0.13	0.25	0.43	0.56	0.70
25	25		0.18	0.36	0.60	0.75	0.87
4	2	2	0.00	0.00	0.00	0.00	0.00
8	4		0.03	0.05	0.10	0.14	0.21
18	9		0.11	0.22	0.39	0.53	0.68
32	16		0.19	0.39	0.64	0.79	0.91
50	25		0.26	0.55	0.83	0.93	0.98
8	2	4	0.00	0.00	0.00	0.00	0.00
16	4		0.04	0.08	0.17	0.25	0.37
36	9		0.15	0.34	0.59	0.75	0.88
64	16		0.26	0.57	0.85	0.95	0.99
100	25		0.38	0.76	0.96	0.99	1.00
16	2	8	0.00	0.00	0.00	0.00	0.00
32	4		0.05	0.12	0.27	0.40	0.56
72	9		0.21	0.48	0.77	0.90	0.97
128	16		0.36	0.74	0.96	0.99	1.00
200	25		0.51	0.90	1.00	1.00	1.00

*Power calculations assume a 5% chance of incorrectly finding a violation.

Table 10: Chance of finding violations in monitoring with a 100 ppm standard

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Chance of detecting violation*				
			Actual PCB concentration				
			125	150	175	200	250
2	2	1	0.00	0.00	0.00	0.00	0.00
4	4		0.02	0.04	0.06	0.08	0.11
9	9		0.10	0.18	0.26	0.33	0.44
16	16		0.16	0.31	0.45	0.56	0.70
25	25		0.22	0.45	0.63	0.75	0.87
4	2	2	0.00	0.00	0.00	0.00	0.00
8	4		0.03	0.07	0.11	0.14	0.21
18	9		0.14	0.28	0.42	0.53	0.68
32	16		0.24	0.49	0.68	0.79	0.91
50	25		0.34	0.67	0.85	0.93	0.98
8	2	4	0.00	0.00	0.00	0.00	0.00
16	4		0.05	0.11	0.18	0.25	0.37
36	9		0.20	0.43	0.62	0.75	0.88
64	16		0.34	0.69	0.87	0.95	0.99
100	25		0.49	0.86	0.97	0.99	1.00
16	2	8	0.00	0.00	0.00	0.00	0.00
32	4		0.06	0.17	0.29	0.40	0.56
72	9		0.27	0.59	0.80	0.90	0.97
128	16		0.47	0.85	0.97	0.99	1.00
200	25		0.65	0.96	1.00	1.00	1.00

*Power calculations assume a 5% chance of incorrectly finding a violation.

To calculate the values in Table 8, we allowed the alternative PCB level, μ_a , to vary between 30 and 250, depending on the standard used, and the number of composited samples, $q (= m/c)$, to vary between 2 and 25. Finding an appropriate approximation for s was somewhat complicated. First, notice that s^2 is computed using q composite samples, and that, in general, s^2/q should be an unbiased estimate of $V(\hat{\theta})$. Thus, from formula (4),

$$\frac{1}{q}E(s^2) = V(\hat{\theta}) = \frac{\tau^2 + \sigma^2}{m} + \frac{\gamma^2}{q} = \frac{1}{q}\left(\frac{\tau^2 + \sigma^2}{c} + \gamma^2\right)$$

suggesting that

$$E(s^2) = \frac{\tau^2 + \sigma^2}{c} + \gamma^2 = \frac{\text{Sampling Variability}}{c} + (\text{Measurement Error}).$$

Since sampling error tends to increase with the level of the PCB concentration, we estimated sampling variability as

$$\tau^2 + \sigma^2 \approx (\hat{CV}_{\tau\sigma i} \hat{\mu}_a)^2 = (1.0 \hat{\mu}_a)^2 = \hat{\mu}_a^2$$

(recalling that $\hat{CV}_{\tau\sigma i}$ was determined empirically to be 1.0 for both white goods and automobiles). Similarly, to estimate measurement error, we used

$$\gamma^2 \approx (\hat{CV}_{\gamma} \hat{\mu}_a)^2 = (.32 \hat{\mu}_a)^2.$$

Combining these terms,

$$s^2 \approx \frac{\hat{\mu}_a^2}{c} + (.32 \hat{\mu}_a)^2 = \hat{\mu}_a^2 \left(\frac{1}{c} + .32^2 \right)$$

or

$$(17) \quad s \approx \sqrt{\hat{\mu}_a^2 \left(\frac{1}{c} + .32^2 \right)}.$$

For example, if $c = 2$, $m = 18$, $\mu_0 = 50$, and $\mu_a = 70$, then

$$\frac{50 - \mu_a}{s/\sqrt{q}} + t_{.05} = \frac{50 - 70}{54.3/\sqrt{9}} + 1.86 = .76.$$

Consulting a table of the standard normal distribution shows that

$$1 - \Phi(.76) = .22$$

as shown in Table 9.

Alternative Significance Levels. Tables 11 and 12 show the effects of alternative significance levels, 10% and 1%, respectively, on the power of hypothesis test for monitoring when the standard is 50 ppm. The effects are analogous to those seen when changing the level of confidence earlier in Section 4. Specifically, increasing the significance level from 5% to 10% results in improvements in power, while lowering the significance level from 5% to 1% results in sharp decreases in power.

6.2 Clean-Up Verification

Power Calculations. The power calculations for clean-up verification are similar to those for monitoring. Recall, as discussed in Section 5, that the direction of the hypothesis is reversed for clean-up verification. In addition, the question of power was phrased in reverse in Tables A-6 through A-8 (reproduced here as Tables 13 through 15), since one would want the chance of additional clean-up to be low if the site was in fact clean. Thus, against a specific alternative, $\mu = \mu_a$, the chance of additional clean-up was calculated as

$$(18) \quad \text{Chance of additional clean-up} = P(\bar{x} > \text{Cut-off Value} \mid \mu = \mu_a)$$

$$= P\left(\frac{\bar{x} - \mu_a}{s/\sqrt{q}} > \frac{\mu_0 - \mu_a}{s/\sqrt{q}} - t_{.05}\right)$$

$$= 1 - \Phi\left(\frac{\mu_0 - \mu_a}{s/\sqrt{q}} - t_{.05}\right).$$

Table 11: Chance of finding violations in monitoring with a 50 ppm standard with 90% confidence

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Chance of detecting violation*				
			Actual PCB concentration				
			60	70	85	100	125
2	2	1	0.00	0.00	0.01	0.01	0.01
4	4		0.09	0.14	0.20	0.25	0.31
9	9		0.18	0.28	0.41	0.51	0.62
16	16		0.24	0.40	0.59	0.71	0.83
25	25		0.30	0.52	0.74	0.86	0.94
4	2	2	0.00	0.01	0.01	0.02	0.02
8	4		0.11	0.18	0.28	0.36	0.46
18	9		0.23	0.38	0.58	0.70	0.82
32	16		0.31	0.55	0.78	0.89	0.96
50	25		0.40	0.70	0.91	0.97	0.99
8	2	4	0.00	0.01	0.02	0.03	0.05
16	4		0.14	0.25	0.40	0.52	0.65
36	9		0.29	0.52	0.75	0.87	0.95
64	16		0.41	0.72	0.92	0.98	1.00
100	25		0.53	0.86	0.98	1.00	1.00
16	2	8	0.00	0.01	0.03	0.06	0.10
32	4		0.17	0.33	0.54	0.68	0.81
72	9		0.36	0.66	0.88	0.96	0.99
128	16		0.52	0.85	0.98	1.00	1.00
200	25		0.67	0.95	1.00	1.00	1.00

*Power calculations assume a 10% chance of incorrectly finding a violation.

Table 12: Chance of finding violations in monitoring with a 50 ppm standard with 99% confidence

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Chance of detecting violation*				
			Actual PCB concentration				
			60	70	85	100	125
2	2	1	0.00	0.00	0.00	0.00	0.00
4	4		0.00	0.00	0.00	0.00	0.00
9	9		0.01	0.02	0.04	0.07	0.12
16	16		0.02	0.07	0.15	0.24	0.38
25	25		0.04	0.13	0.30	0.46	0.64
4	2	2	0.00	0.00	0.00	0.00	0.00
8	4		0.00	0.00	0.00	0.00	0.00
18	9		0.01	0.04	0.10	0.17	0.28
32	16		0.04	0.13	0.32	0.49	0.69
50	25		0.08	0.26	0.56	0.77	0.92
8	2	4	0.00	0.00	0.00	0.00	0.00
16	4		0.00	0.00	0.00	0.00	0.01
36	9		0.02	0.07	0.21	0.36	0.55
64	16		0.07	0.25	0.57	0.78	0.93
100	25		0.14	0.47	0.84	0.96	0.99
16	2	8	0.00	0.00	0.00	0.00	0.00
32	4		0.00	0.00	0.00	0.01	0.02
72	9		0.03	0.14	0.38	0.60	0.81
128	16		0.11	0.42	0.80	0.94	0.99
200	25		0.23	0.69	0.97	1.00	1.00

*Power calculations assume a 1% chance of incorrectly finding a violation.

Table 13: Chance of requiring additional clean-up with a 25 ppm standard

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Chance of requiring more clean-up*				
			Actual PCB concentration				
			1	5	10	15	20
2	2	1	—	0.82	1.00	1.00	1.00
4	4		—	—	0.31	0.86	0.97
9	9		—	—	0.01	0.48	0.87
16	16		—	—	—	0.22	0.79
25	25		—	—	—	0.07	0.70
4	2	2	—	0.16	1.00	1.00	1.00
8	4		—	—	0.07	0.74	0.96
18	9		—	—	—	0.24	0.81
32	16		—	—	—	0.05	0.68
50	25		—	—	—	—	0.54
8	2	4	—	—	1.00	1.00	1.00
16	4		—	—	—	0.54	0.93
36	9		—	—	—	0.07	0.72
64	16		—	—	—	—	0.53
100	25		—	—	—	—	0.35
16	2	8	—	—	0.97	1.00	1.00
32	4		—	—	—	0.33	0.90
72	9		—	—	—	0.01	0.61
128	16		—	—	—	—	0.37
200	25		—	—	—	—	0.18

*These calculations assume a 95% (or greater) chance of requiring additional clean-up when the concentration of PCB's is 25 ppm or greater. A dash (—) indicates that the chance is less than .005.

Table 14: Chance of requiring additional clean-up with a 50 ppm standard

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Chance of requiring more clean-up*				
			Actual PCB concentration				
			10	15	20	30	40
2	2	1	0.82	1.00	1.00	1.00	1.00
4	4		—	0.02	0.31	0.86	0.97
9	9		—	—	0.01	0.48	0.87
16	16		—	—	—	0.22	0.79
25	25		—	—	—	0.07	0.70
4	2	2	—	0.16	1.00	1.00	1.00
8	4		—	—	0.07	0.74	0.96
18	9		—	—	—	0.24	0.81
32	16		—	—	—	0.05	0.68
50	25		—	—	—	—	0.54
8	2	4	—	0.77	1.00	1.00	1.00
16	4		—	—	—	0.54	0.93
36	9		—	—	—	0.07	0.72
64	16		—	—	—	—	0.53
100	25		—	—	—	—	0.35
16	2	8	—	0.27	0.97	1.00	1.00
32	4		—	—	—	0.33	0.90
72	9		—	—	—	0.01	0.61
128	16		—	—	—	—	0.37
200	25		—	—	—	—	0.18

*These calculations assume a 95% (or greater) chance of requiring additional clean-up when the concentration of PCB's is 50 ppm or greater. A dash (—) indicates that the chance is less than .005.

Table 15: Chance of requiring additional clean-up with a 100 ppm standard

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Chance of requiring more clean-up*				
			Actual PCB concentration				
			20	30	40	60	80
2	2	1	0.82	1.00	1.00	1.00	1.00
4	4		—	0.02	0.31	0.86	0.97
9	9		—	—	0.01	0.48	0.87
16	16		—	—	—	0.22	0.79
25	25		—	—	—	0.07	0.70
4	2	2	0.16	0.98	1.00	1.00	1.00
8	4		—	—	0.07	0.74	0.96
18	9		—	—	—	0.24	0.81
32	16		—	—	—	0.05	0.68
50	25		—	—	—	—	0.54
8	2	4	—	0.77	1.00	1.00	1.00
16	4		—	—	—	0.54	0.93
36	9		—	—	—	0.07	0.72
64	16		—	—	—	—	0.53
100	25		—	—	—	—	0.35
16	2	8	—	0.27	0.97	1.00	1.00
32	4		—	—	—	0.33	0.90
72	9		—	—	—	0.01	0.61
128	16		—	—	—	—	0.37
200	25		—	—	—	—	0.18

*These calculations assume a 95% (or greater) chance of requiring additional clean-up when the concentration of PCB's is 100 ppm or greater. A dash (—) indicates that the chance is less than .005.

The sample standard deviation, s , was approximated as described in Section 6.1. For example, if $c = 4$, $m = 8$, $\mu_0 = 50$, and $\mu_a = 15$, then

$$\frac{\mu_0 - \mu_a}{s/\sqrt{q}} - t_{.05} = \frac{50 - 15}{8.90/\sqrt{2}} - 6.31 = -0.75.$$

Consulting a table of the standard normal distribution shows that

$$1 - \Phi(-0.75) = \Phi(0.75) = .77$$

as shown in Table 14.

Alternative Significance Levels. The effects of changing the significance levels to 10% and 1% are shown in Tables 16 and 17, respectively. A standard of 50 ppm is assumed in both tables. As compared with a 5% significance level (i.e., Table 14), the chance of requiring additional clean-up is substantially smaller with a 10% significance level when μ_a is small, although there is little difference as μ_a approaches 50. Similarly, when the significance level is 1%, the chance of requiring further clean-up is much greater when μ_a is small, but there is little difference when μ_a is close to 50.

Table 16: Chance of requiring additional clean-up with a 50 ppm standard with 90% confidence

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Chance of requiring more clean-up*				
			Actual PCB concentration				
			10	15	20	30	40
2	2	1	0.01	0.47	0.85	0.99	1.00
4	4		—	—	0.11	0.64	0.88
9	9		—	—	—	0.31	0.75
16	16		—	—	—	0.12	0.65
25	25		—	—	—	0.03	0.55
4	2	2	—	0.12	0.63	0.97	1.00
8	4		—	—	0.01	0.47	0.84
18	9		—	—	—	0.12	0.67
32	16		—	—	—	0.02	0.52
50	25		—	—	—	—	0.38
8	2	4	—	0.01	0.31	0.93	0.99
16	4		—	—	—	0.27	0.79
36	9		—	—	—	0.02	0.55
64	16		—	—	—	—	0.37
100	25		—	—	—	—	0.21
16	2	8	—	—	0.09	0.86	0.99
32	4		—	—	—	0.12	0.72
72	9		—	—	—	—	0.43
128	16		—	—	—	—	0.22
200	25		—	—	—	—	0.10

*These calculations assume a 90% (or greater) chance of requiring additional clean-up when the concentration of PCB's is 50 ppm or greater. A dash (—) indicates that the chance is less than .005.

Table 17: Chance of requiring additional clean-up with a 50 ppm standard with 99% confidence

Total samples collected <i>m</i>	Number of composites analyzed <i>q = m/c</i>	Subsamples in each composite <i>c</i>	Chance of requiring more clean-up*				
			Actual PCB concentration				
			10	15	20	30	40
2	2	1	1.00	1.00	1.00	1.00	1.00
4	4		—	0.54	0.95	1.00	1.00
9	9		—	—	0.08	0.84	0.99
16	16		—	—	—	0.52	0.95
25	25		—	—	—	0.25	0.90
4	2	2	1.00	1.00	1.00	1.00	1.00
8	4		—	0.07	0.75	1.00	1.00
18	9		—	—	—	0.63	0.97
32	16		—	—	—	0.20	0.91
50	25		—	—	—	0.04	0.81
8	2	4	1.00	1.00	1.00	1.00	1.00
16	4		—	—	0.30	0.99	1.00
36	9		—	—	—	0.32	0.95
64	16		—	—	—	0.03	0.82
100	25		—	—	—	—	0.65
16	2	8	1.00	1.00	1.00	1.00	1.00
32	4		—	—	0.04	0.96	1.00
72	9		—	—	—	0.10	0.91
128	16		—	—	—	—	0.69
200	25		—	—	—	—	0.45

*These calculations assume a 99% (or greater) chance of requiring additional clean-up when the concentration of PCB's is 50 ppm or greater. A dash (—) indicates that the chance is less than .005.

REPORT DOCUMENTATION PAGE	1. REPORT NO. EPA 747-R-93-010	2.	3. Recipient's Accession No.
4. Title and Subtitle Sampling Guidance for Scrap Metal Shredders Technical Background			5. Report Date August 1993
			6.
7. Author(s) James Bethel, Westat, Inc.			8. Performing Organization Rept. No.
9. Performing Organization Name and Address Westat, Inc. 1650 Research Blvd. Rockville, MD 20850			10. Project/Task/Work Unit No.
			11. Contract (C) or Grant (G) No. 68-02-4293
12. Sponsoring Organization Name and Address U.S. Environmental Protection Agency Office of Prevention, Pesticides and Toxic Substances Washington, D.C. 20460			13. Type of Report & Period Covered Technical Report
			14.
15. Supplementary Notes			
16. Abstract (Limit: 200 words) <p>The purpose of this document is to provide basic instructions for collecting and statistically analyzing samples of materials that are produced as a result of shredding automobiles and other metal objects, since the by-products of these recycling operations may contain concentrations of polychlorinated biphenyl's (PCBs). Shredders are large machines that convert light metal objects into fist size or smaller pieces of scrap metal. PCBs enter the shredder output when materials containing PCB-bearing fluids are shredded. Large concentrations of PCBs have been identified in some samples that have been collected at some recycling sites. Thus agencies may wish to collect data at shredder sites in order to study the situation in their locality. The sampling procedures described in this document are intended to produce representative samples of fluff that will give reasonably accurate estimates of the overall concentration of PCBs in the material being sampled. The document discusses sample selection, laboratory testing, and statistical procedures for analyzing the data. This volume provides additional detail on the assumptions, background data, and numerical calculations that support the Field Manual (EPA 747-93-009).</p>			
17. Document Analysis a. Descriptors Environmental contaminants, scrap metal recycling b. Identifiers/Open-Ended Terms PCB, sampling, statistical analysis c. COSATI Field/Group			
18. Availability Statement	19. Security Class (This Report) Unclassified	21. No. of Pages 50	
	20. Security Class (This Page) Unclassified	22. Price	