T-3676

The Application of Sampling Theory and Field Analytical Techniques to Environmental Investigation

by

Charles A. Ramsey

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Master of Science (Mineral Resource Ecology).

Golden, Colorado

Date 4-5-89

Signed: Charles A. Ramsey

Approved: Dr. Ronald H. Cohen
Thesis Advisor

Golden, Colorado

Date 4-6-89

Dr. John C. Emerick
Associate Professor and Head,
Environmental Sciences Dept.

ABSTRACT

The concerns of environmental sampling necessitate application of field analytical techniques. Unless sampling theory is considered in such investigation, the analytical results may be invalid.

Of all possible sampling errors, the Fundamental Error and the Long Range Heterogeneity Error cause the majority of the total sampling errors. It is only through understanding of the causes of these errors that proper environmental investigation can be conducted. At present, most environmental investigators lack the expertise to properly handle these errors in the investigative process.

This study analyzes inherent problems that exist in environmental sampling, presents an experiment in which the sampling theory of Pierre Gy is utilized, discusses the actual use of field analytical techniques, and develops a model for improving environmental investigation by measuring and minimizing sampling errors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

PURPOSE

The purpose of this study is to apply sampling theory to a realistic environmental sampling situation in order to develop a model for future application in environmental investigation of solid materials. This thesis does not consider the sampling of liquids or gases.

The application of sampling theory makes it possible to determine sampling errors which may occur in environmental investigation. Therefore the magnitude of the sampling errors can be assessed; this is vital because certain error levels will invalidate any conclusions obtained from the analytical results. This prior knowledge is necessary to guarantee that the accuracy and precision goals of a sampling episode or investigation will be met.

Because sampling theory is usually not considered, results of those investigations are often based on faulty or incomplete assumptions. This discussion will demonstrate that field analytical techniques can be used to measure certain characteristics of the material to be sampled, thereby eliminating the need to rely on these assumptions.

The specific sampling theory applied in this study is that of Pierre Gy. The case examined is that of sampling the waste material on two railroad cars to determine if the contents are hazardous.

Chapter 2

DEFINITIONS

The scientific sampling of waste materials involves many environmental and statistical definitions. These are explained below to ensure an understanding of this study.

Sample: A sample is an experimental unit taken from a larger group (population) of individual units. Samples are removed from a material under investigation so that the results obtained from this investigation may be used to draw statistical inferences about the population as a whole. To make accurate inferences about the target population (the entire set of units) from the samples (experimental units) taken, certain rules must be followed to ensure the validity of the calculated results. These are randomness and independence; of these, randomness must be defined for the purposes of this study. Independence is self-explanatory. (see Walpole and Myers, 1985)

Randomness: This is one of the underlying assumptions of classical statistics. For a sample to be random it must have the same probability of selection as any other possible sample taken from the target population. Given this equal chance of selection, certain unbiased generalizations about the population can be made and expressed quantitatively through application of mathematical probability. Random sample selection allows characterization of the population with a calculable degree of accuracy and precision.

Representative: In addition to a sample being random, it must be representative of the target population. If the sample does not have properties close to the properties of the target population, it cannot be used to infer the properties of the target population. Collection of a perfectly representative sample is impossible, except by luck. If it were possible, only one sample would be necessary. The closer the sample properties are to population properties, the fewer samples are necessary to estimate the population properties and the more accurate the prediction will be.

Confidence level: In this study confidence level relates to the confidence that the mean (average level) of the target population is above the regulatory threshold. If the same statistical procedure was performed a large number of times, the confidence level is the percent probability that actual mean will be above the regulatory threshold. The generally accepted cutoff for the level of scientific certainty is 95 percent, though this percentage will vary according to the needs of the investigation (Natrella, 1963).

Sampling strategies: This is a determination of the locations within the target population from which the samples will be collected. (See section on sampling strategies for further explanation).

Regulatory threshold: The level of contamination which the Environmental Protection Agency (EPA) has determined to be a violation of environmental laws.

Extraction Procedure: A test promulgated by the EPA to simulate leaching of waste material in the environment. The leachate from this test is analyzed and the results are compared to the regulatory threshold. If the leachate values are higher than the regulatory threshold, the waste is considered hazardous. (40 CFR)

Sample composite: The combining of a number of increments (experimental units) collected from the same waste to make one sample. This can have advantages and disadvantages, depending on the goals of the research. Small-scale sample compositing can be used to minimize grouping and segregation error (see Section 7.8.2). Large-scale sample compositing is used to reduce long-range variability, which might or might not be desirable, depending on whether this long-range variability is an important aspect of the investigation.

Chapter 3

INTRODUCTION

Sampling theory is usually not considered in sampling of the environment. There are several causes for this lack of theoretical application. First, there is a lack of knowledge of sampling theory itself. Second, there is a lack of awareness of the appropriateness of this theory for sampling the environment. Third, even if the theory is applied, the assumptions that must be made during the process may result in the collection of an unnecessarily large sample mass because little or nothing is known about the sample prior to the actual sampling episode.

This study will demonstrate that field analytical techniques based on Pierre Gy's sampling theory can be used to measure characteristics of the material to be sampled, and assumptions requiring unnecessarily large sample mass need not be relied upon. If the characteristics of the material are known, then the required sample can usually be reduced greatly and the sampling episode becomes more feasible. This can be accomplished without any reduction of sampling accuracy or precision.

In environmental sampling, where application of Pierre Gy's sampling theory has as its objective the evaluation of chemical content of waste materials, some inherent problems exist which do not occur with other types of sampling.

First is the possibility of limited site access due to non-cooperation by the site owner. Second is the consideration of safety for the investigator. These problems restrict accessibility to the sampling site and might only allow one opportunity for sampling and analysis. In addition, the contamination might be mobile, as a result of water runoff, erosion, or similar natural or man-made intrusions on the environment. The contaminant can also be transformed in the environment by chemical processes which alter its toxicity.

The solution to sampling a dynamic environment is to apply field analytical techniques. This allows collection of pre-sampling information in a timely manner. Conventional methods of obtaining this information result in outdated information which might invalidate the investigation.

In addition to these special environmental investigation considerations, problems which are the norm in all sampling continue to exist. These are: the number of samples to collect, the sampling strategy to employ, and the mass of sample needed. The latter, which is addressed in this study of sampling theory and its applications, is often ignored.

In most investigations, sampling strategy is the only consideration. This can result in invalid conclusions concerning the properties of the sampled population because calculation of sample mass and number of samples to collect is not possible without prior information concerning the material under investigation. It is only with application of sampling theory that these calculations can be accomplished. Ideally, these analyses are performed on site immediately prior to actual sample collection.

Although sampling strategy should not be the only consideration, it is a vital strategy to employ in environmental investigation. There are two types of sampling strategy, random (which is probabilistic) and authoritative (non-random, non-probabilistic, search, or purposive). Probabilistic sampling is preferred for characterization of the average property of hazardous waste. However, in some cases only the presence of contamination must be determined, and authoritative sampling can be utilized.

These considerations of sampling strategies are vital in order to match them with the desired purpose and goal. For cases where the presence of contamination is the issue, non-probabilistic sampling is adequate. The advantage of this type of sampling is the lack of need for prior information vital for proper random sampling. However, randomness is required when the goal of the investigation is to quantify the contamination level of the target population.

There are three probabilistic (random) sampling strategies that are commonly used for environmental investigation. Simple random sampling is the easiest of these to implement. This is the selection of an adequate number of samples from the target population, with each sample having an equal probability of being sampled. The problem with simple (as is also true with stratified) random sampling is the need for an estimate of the mean (average level) and standard deviation of the contaminant in question in order to calculate the appropriate number of samples to be taken (Cochran, 1963). These estimates are often not available in environmental work. Therefore much guess work takes place on the part of the sampling plan designer. There is no way for this individual to know if the goals of the sampling plan will be met until all the samples are analyzed unless sequential sampling is employed (Ghosh, 1970), and at this point resampling may not be a possibility.

Stratified random sampling is useful when the population contains two or more non-randomly distributed heterogeneous batches of material with respect to the chemical property of interest. In addition, this strategy can identify individual subgroups or strata and random sampling from each of these. Stratified random sampling is often able to provide a more accurate estimate of the property of the materials to be evaluated, but only if there is sufficient stratification of the material. This technique gives an accurate estimate of the chemical property, as does the simple random technique, but estimates of the mean (average level) and standard deviations are, once again, necessary to determine the number of samples required. Stratified random sampling is precise only when the material can be divided into the proper strata. Otherwise it has less precision than simple random sampling Cochran, 1963).

Systematic random sampling is used to collect data while the factors under analysis, such as time, temperature, location, or chemical composition, are changing. The first sample is taken at a random starting place and all subsequent samples are taken at specified intervals of the predetermined factor. These samples are usually treated as representing a separate, discreet population which was in existence when the specific conditions existed (Cochran, 1963). Care should be taken to analyze the collected data, with special consideration given to checking for systematic bias and the impact of unusual events during the sampling times before any statistically significant conclusions may be drawn. Although systematic sampling is useful because it allows maximum sampling efficiency once a particular process is understood, it is somewhat limited in drawing conclusions for the possibility of different conditions existing at the time or place of sampling. This technique is good when the population of interest is essentially random or only slight stratification exists. It is often employed for the sake of convenience, such as when automatic samplers are used. This technique should be used with extreme caution in environmental investigations (Pitard, 1988).

The value of non-probabilistic or authoritative sampling in environmental investigation is to show the presence or absence of a contaminant, not to quantify the extent and level of contamination. For example, this method can be used to confirm that a chemical spill has taken place. Accuracy and precision usually are not measurable nor are they required when authoritative sampling is employed (Pitard, 1988). Generally conclusions drawn apply only to the samples themselves and not to a larger population. Authoritative samples are more correctly referred to as specimens.

Probabilistic sampling techniques are used when the goal is to collect a representative sampling. This is the collection of samples which represent the properties of the population with specified levels of precision and accuracy. In some cases the collection of a representative sample is no more work than collection of a authoritative sample; in other cases it is more difficult. The only difference between representative and judgment samples is employment of a sampling plan based on sound sampling theory (Pitard, 1988). It is impossible to collect a perfectly representative sample, except by luck, so predetermined measures of accuracy and precision are necessary to define representativeness. The necessary degree of accuracy and precision depends on the goal of the investigation. The more accuracy and precision desired, the greater the sample mass required. The goal of the sampling plan is to reduce the sampling errors to a manageable level consistent with the objectives of the study. The selection of a sampling strategy alone is not sufficient for environmental investigations because the sample mass also affects accuracy and precision and therefore needs to be considered (Gy, 1982).

One of the shortcomings present in most environmental sampling is collection of an inadequate mass of sample, resulting in two errors. The first error occurs when the sample mass is too small to contain a representative portion of fragments of all sizes in the parent population. This can result in bias or increased sampling error if all particles do not contain the same level of contamination. The second error can occur because the element in question is found in trace element quantities (Clifton et al, 1969).

The result of too small a sample mass and trace element quantities is a Poisson distribution of the analytical results. A Poisson distribution is used to model rare, random occurrences in space or time (Davis, 1986). If the chemical constituent is congruent with a Poisson distribution, the application of classical statistics based on normal distributions are invalid.

One attribute of the Poisson distribution is that the mode is less than the mean (Walpole and Myers, 1985). This difference will result in underestimation of the average level of the contaminant in the target population if statistics based on a normal distribution are used. A Poisson distribution results when there are fewer than six occurrences of contaminant in each sample.

If, on the other hand, there are at least six occurrences of contamination in every sample, the Poisson distribution is no longer very skewed and, in fact, approaches a normal distribution (David, 1982). It is this normal, or approximately normal, distribution that is necessary for the use of classical statistics. To increase the number of occurrences in the sample, the sample mass needs to be increased.

These two potential errors (particle size representativeness and fewer than six occurrences of contamination) cause much of the lack of accuracy and precision associated with sampling and are the main focus of this thesis.

Given all of the above considerations, it is apparent that probabilistic sampling is best suited for environmental sampling when the average property of a waste material is in doubt since any sampling bias from the sample selection scheme must be removed (Keith, 1988). Without some type of random sampling, the possibility exists that the sample locations become dependent on the judgement of the sampler. This could mean that the results of an investigation could be contingent on the sampler's perspective. Biases of this type could invalidate the results of the study. This does mean, however, that the sampler should not use good judgement to locate the target population to representatively sample.

However, the problem with random sampling is the lack of prior information to determine the number of samples necessary to achieve prescribed precision and accuracy. Currently, the only solution to this dilemma is to be able to estimate by some other means the sampling error prior to sampling.

Another solution to this dilemma can be found through the use of Pierre Gy's theory of sampling. He developed a sampling theory that can measure errors, particularly when at least some information is available about the waste material. If these errors are too large to achieve the goals of an environmental investigation, application of sampling theory can provide a method to reduce these errors. A discussion of Pierre Gy's theory follows, and its application will be demonstrated in real environmental investigation.

Chapter 4

PIERRE GY'S SAMPLING THEORY

Pierre Gy, a French engineer, adapted portions of empirically derived sampling theory developed by his peers during this century and formulated a mathematically based, comprehensive theory applicable to sampling situations of any nature. His theory is explained in his book (translated from the French): Sampling of Particulate Materials; Theory and Practice (Gy, 1982), which will be referred to in this chapter.

Application of Gy's sampling theory allows collection of a representative sample by addressing the particle size and number of contaminant occurrences in each sample of the waste material. Essentially, the required sample mass. increases as the particle size increases. There must also be at least six occurrences of contamination in every sample. With the Pierre Gy method, all errors can be measured, reduced, or eliminated.

Pierre Gy has developed a sampling theory which quantitatively determines the amount of sample necessary to achieve the prescribed accuracy and precision of the investigation. Gy has broken down his sampling theory into individual error types. The first four errors are sampling errors and are caused either directly or indirectly by the compositional heterogeneity of the different sized particles. Other errors that are not sampling errors, but must be considered in the overall sampling scheme as they add to the resulting uncertainty of the inferred parameter of the population, are the preparation and analytical errors.

4.1 Fundamental Error

The Fundamental Error results when fragments of the material to be sampled have different chemical compositions (compositional heterogeneity) of the analyte under investigation. Fundamental Error is the only error that exists when the sampling is perfect because it is a property of the material to be sampled and cannot be eliminated by any method other than reducing the particle size of the sampled population or increasing the sample weight. When sampling, it is necessary to take a representative sample of all particle sizes. If a 100 gram sample is taken when the largest particles weigh 200 grams, it is impossible to represent the 200 gram particles in the sample. This can be a source of extreme bias if the contaminant is contained only within the 200 gram particles. The Fundamental Error can be calculated for any sample weight and particle size relatively easily. The Fundamental Error occurs at many points during the analytical process, including field sampling, field splitting, sample weighing and laboratory sub-sampling.

4.2 Grouping and Segregation Error

Grouping and Segregation Error is caused by the non-random distribution of the particles throughout the target population (distribution heterogeneity) and the impossibility of collecting samples one particle at a time. This non-random distribution can be caused by many factors, including fragment density heterogeneity, fragment size heterogeneity, fragment shape heterogeneity, air turbulence, gravity, vibrations, or anything that might tend to separate the fragments.

The magnitude of the Grouping and Segregation Error is, by definition, bounded by zero and the value of the Fundamental Error (Smith and James, 1981). The Fundamental Error is a measure of the compositional heterogeneity. If there is no compositional heterogeneity (all particles have the same composition) there can be no distribution heterogeneity.

Two ways to reduce the Grouping and Segregation Error without increasing the sample mass are compositing the sample from smaller increments and homogenizing the target population prior to sampling. The first technique of error reduction is often simpler to implement than the latter due to the practical inability of homogenizing the entire target population. In addition, complete homogenization will never be achieved (Ingammels and Pitard, 1986).

4.3 Long Range Heterogeneity Error

Long Range Heterogeneity Errors occur when there is a different concentration of the contaminant of interest in different locations within the sampled material. Long Range Heterogeneity Error can also be termed a variation, because it is not always considered to be error. This classification depends on the goals of the investigation. If the purpose of the investigation is to determine the average value of a contaminant, then any spatial variation in contaminant concentration will cause variability in the analytical results; this would be considered error. If, on the other hand, the distribution of the contaminant within the sampled area or the locations of "hot spots" within the sampled area is of concern, the spatial variability is not an error. However, the goal of the investigation will be to measure this variability. In the later case Long Range Heterogeneity Error can be considered as a natural variation and not an error.

Long Range Heterogeneity Error is caused by spatial trends in the level of the contaminant contained in the waste material or cyclic changes in the level of contamination. This error is very difficult to determine and measure, but its magnitude can be reduced by use of the proper sampling plan and collection of many sampling increments over the entire sampling population, if indeed reduction is desired. Reducing this error by compositing over a wide area will mask the variability of the property in question, which might or might not be a problem.

Geostatistics is a robust, available technique to measure Long Range Heterogeneity Error if adequate numbers of representative samples are taken. If the error is great enough to cause a problem, there are many techniques available to reduce this error. The error is not of great concern for the characterization of the average content of the total waste stream if the magnitude of this error is within the prescribed limits of the study and does not affect the resulting accuracy or precision. If the error is large, the results of a geostatistical analysis will allow the investigator to stratify the waste for stratified random sampling. This approach necessitates the sampler to have knowledge of the characteristics of the waste material prior to sample plan design.

4.4 Materialization Error

While Materialization Error is simple in concept, it is the cause of many biases that occur in sampling. A theoretical boundary exists in all sampling instruments. This boundary can easily be seen if a coring device is imagined. The inside diameter of the circular metal coring device is the outer boundary of the sample. All particles whose center of gravity lies within this boundary when the coring device punctures the target population should become part of the sample, and those whose center of gravity is outside this boundary should not become a part of the sample. When this center of gravity principle is broken, what results is a part of the Materialization Error, also called Delimination Error. A rule of thumb that can be used when selecting the proper coring device is to be sure the inside diameter of the coring device is at least three times the diameter of the largest particle in the target population (Weise, 19XX).

The other type of Materialization Error is the Increment Extraction error. Once the sample is properly deliminated, the entire sample must be extracted--no more and no less--to the sample container.

Materialization Error is caused by either improper selection of sampling device or improper use of the correct sampling device. Some devices (often including devices sold for that purpose and always including devices such as coffee cans and pop bottles) are fundamentally inappropriate and should never be used for sample collection. Even if proper Increment Delimination is achieved, proper Increment Extraction may never be. The opposite is also true.

4.5 Preparation Error

Preparation Error results from the loss or addition of the analyte in question during the preparation of the sample for analysis. These errors cause a bias that is impossible to detect without use of quality control samples and standard reference material as part of every preparation stage. Preparation Error occurs during sample transferring, weighing, drying, filtering, comminuting, screening. Proper analysis of quality control data will detect these errors. The extraction portion of the EP Toxicity test is considered a Preparation Error.

4.6 Analytical Error

Analytical Error results from lack of instrument and analytical precision and accuracy during the determination of analytes in the sample. This source of error is the most understood and controlled of any error discussed thus far. With advances in instrumentation and proper use of quality control programs, analytical errors are minimal, easily measurable and controlled.

Quantification of the errors discussed above prior to the investigation will allow the investigator to know before the sampling and analysis commence whether the goals demanded of the investigation can be met despite the error level that has been calculated. It is only through the application of Pierre Gy's sampling theory that this can be accomplished.

Pierre Gy developed sampling theory for the mining and metallurgical industries (Weiss, 19xx), where economics necessitate representative sampling. This sampling theory can just as easily be applied to environmental sampling.

The Fundamental Error is the main focus of this study because it is the major cause of environmental sampling error and is difficult to comprehend. Long Range Heterogeneity Error can also be a major source of error in environmental investigation, even though this is a more readily understood concept. For this reason, the Fundamental Error is discussed in detail in the next chapter, with appropriate mention of Long Range Heterogeneity Error as it applies to this study.

Chapter 5

FUNDAMENTAL ERROR

The only error which is inherent in a property of the material to be sampled is the Fundamental Error (FE). This error can be reduced only by increasing the sample weight or by reducing the particle size. The Fundamental Error will be the main concern of this study as most other sampling errors can be reduced or eliminated by proper sampling technique or by compositing many increments to make a sample. One possible exception to this is the Long Range Heterogeneity Error which can be a large source of error if not properly addressed in the sampling plan.

Long Range Heterogeneity Errors are difficult to measure and eliminate if field analytical techniques are not employed. One reason for this is the small data sets available in environmental sampling. Another reason is that their existence cannot be determined nor their error estimated until the samples are already taken and analyzed. By this time it is too late to reduce these errors, if their reduction is necessary to achieve the goals of the investigation.

Unless Long Range Heterogeneity Errors can be measured and included in the total acceptable error for the investigation, there exists the possibility that the study will be inconclusive. Most other errors do not need to be quantified in order to know how to reduce them. However, they should always be reduced, whether or not they are quantified, and this can usually be accomplished with ease.

The Fundamental Error is the nucleus of Pierre Gy's sampling theory and therefore the calculation of this error should be discussed.

The equation for the Fundamental Error (FE) according to Pierre Gy is:

$$S^2_{(FE)} = f*b*c*g*d^3/M_S \qquad (5.1)$$

where:

$S^2_{(FE)}$ = variance of the sample collected caused by the Fundamental Error

f = shape factor

b = liberation factor

c = mineralogical factor

g = granulometric factor

d = size of largest particle

M_S= mass of the sample

The Shape Factor (f), by definition, is 1.00 for perfect cubes and can be assumed to be .5 for spherical particles. It can be more than 1.00 for fibrous material (e.g., asbestos). The Shape Factor is a correction factor when the particle in question is not a perfect cube for determining the volume of a particle. This volume is used in conjunction with the density to determine the weight of the particles.

The Granulometric Factor (g) is assumed to be:

.25 for non-calibrated material
.55 for calibrated material (between two screens)
.75 for naturally calibrated materials
1.00 for material exactly the same size

The Granulometric Factor is a correction factor because not all particles are of the largest particle size. The smaller sized particles have a smaller error, but the diameter used in the error calculation is that of the largest particles. The granulometric factor accounts for this.

The Mineralogical Factor (c) is the density (gm/cm^3) of the material divided by the average content of the element (A_L) in question, expressed as a part of one.

$$\text{density (gm/cm}^3\text{)}/A_L \tag{5.2}$$

The Liberation Factor (b) is the size of the material when the element in question becomes "liberated" from the guange material. It is also a measure of homogeneity. The bounds of the liberation are zero and one. Approximate values for b are:

very homogeneous	$b = 0.05$
homogeneous	$b = 0.1$
average material	$b = 0.2$
heterogeneous	$b = 0.4$
very heterogeneous	$b = 0.8$

The Liberation Factor can be calculated by the following formula:

$$b = (A_{MAX} - A_L)/(1 - A_L) \tag{5.3}$$

where:

A_{MAX} = the maximum analyte concentration of the largest particle

A_L = the average analyte concentration of the largest particle

The Liberation Factor can also be measured by viewing the material with a microscope to see at what particle size the analyte of interest becomes "liberated" from the guange material. If the Liberation Factor cannot be determined and there is no reason to believe that it is very large, an estimate of 0.05 can be used.

The diameter (d) is the largest particle in the sample that is obtained when 95 percent of the sample passes a screen of size "d" openings.

The preceding equation (5.1) for the Fundamental Error is applicable when nothing is known about the material to be sampled and assumptions regarding characteristics of the material are made in order to employ sampling theory. If this is the case, then the regulatory threshold of the contaminant is substituted for the average concentration (A_L) in the calculation of the Mineralogical Factor. The density of the contaminant will also have to be estimated if the analyte of interest is not known (there are many contaminants that might cause a violation), but a value of approximately eight for heavy metal contamination will be adequate.

If it is determined that the sample contains at least six occurrences of contamination, a simplification can be made to equation 5.1. This simplification only considers the representativeness of the particle size and ignores any errors associated with trace element problems. The equation is:

$$S^2_{(FE)} = 18*f*e*d^3/M_S \qquad (5.4)$$

$S^2_{(FE)}$ = variance of the sample collected caused by the Fundamental Error

f = shape factor (assumed 0.5)

e = density of material (gm/cm^3)

d = size of largest particle (cm^3)

M_S= mass of the sample (gm)

An example of the calculation of the Fundamental Error using equation 5.4 is:

assumptions:

$f = 0.5$

$e = 2.5$ gm/cm^3

$d = 1$cm

$M_S = 100$ grams

$S^2_{(FE)} = 18*0.5*2.5*1^3/100 = 0.225$

$S = 0.47$ (47 percent)

As has been previously stated, a way to reduce this error would be to reduce the particle size of the sampled material prior to sampling or increase the mass of the sample to be collected. For example, if the sample mass is increased to 1,000 grams, the variance will decrease to 0.0225 (S = 15 percent). If the particle size is reduced to

4mm, with the same 100 gram mass, the variance will decrease to 0.014 (S = 12 percent). It is through this sample manipulation that the Fundamental Error can be altered to suit the needs of a particular study.

Sampling theory can measure and reduce the errors associated with sampling if some information is known about the waste material. The proper application of field analytical techniques is invaluable in obtaining this type of preliminary information about the population(s) of interest. These analytical techniques need not be of the highly sophisticated sort employed in the modern analytical laboratory. They should, however, provide the investigator with timely data of sufficient quality so that proper sampling related decisions are made. Examples include portable X-ray fluorescence, infrared spectroscopy, potentiometry, and conductimetry. The following section describes the experiments used in the development of a field analysis protocol that allows proper sample plan design.

Chapter 6

EXPERIMENT

As has been discussed, Pierre Gy developed a theory of sampling particulate matter (Gy, 1982) that will allow collection of a representative sample by addressing the particle size and number of contaminant occurrences in each sample of the waste material. Essentially, the required sample mass increases as the particle size increases but in addition there must be at least six occurrences of contamination in every sample. With the Pierre Gy method, all errors can be measured, reduced or controlled.

This study will address sampling of a lead-containing waste material prior to application of the Extraction Procedure (EP) (40 CFR) test for determining whether the waste material exhibits hazardous characteristics. One of the most important parameters when designing environmental sampling plans is the regulatory threshold of the contaminant(s) of interest. Regulatory threshold is the level that constitutes a violation of the regulations.

Other considerations when designing a sampling plan are the mass of the sample to be taken and the size of the target population. In the case of the Extraction Procedure, 100 grams is required (40 CFR) as a minimum sample size.

The total mass of the sampled material can also be a factor in any sampling investigation. When sampling to determine if a waste possesses hazardous waste characteristics a minimum of 1,000 kg of material must be characterized as a hazardous waste in order for a particular generator to be classified as a large quantity generator (40 CFR). This classification is important as more stringent regulations apply to this class of generators.

These concerns must be addressed prior to the design of a sampling plan.

6.1 Initial Sampling Episode

The initial problem was evaluation of the results of six samples taken from two railroad cars to determine if a waste material contained in the railroad cars exhibited the Resource Conservation Recovery Act (RCRA) hazardous waste characteristic of Extraction Procedure (EP) toxicity (40 CFR) due to the lead constituent. EP lead concentrations for the six samples ranged from 6.2 mg/L to 15.1 mg/L, averaging 9.8 mg/L (See Table 6.1). The maximum allowable concentration for EP toxicity under RCRA is 5 mg/L of lead in the leachate. This RCRA limit will be referred to as the regulatory threshold. Samples were collected from two open gondola railroad cars by taking a surface scrape at the front, middle, and end of each car.

TABLE 6.1

EP Toxicity Lead Results for the Waste Material in the Railroad Cars

Sample Number	Sample Location	EP Lead Value
1	Railroad Car 1	15.1 mg/L
2	Railroad Car 1	7.4 mg/L
3	Railroad Car 1	6.2 mg/L
4	Railroad Car 2	6.7 mg/L
5	Railroad Car 2	14.4 mg/L
6	Railroad Car 2	9.2 mg/L

Average = 9.8 mg/L

Standard Deviation = 3.9 mg/L

Relative Standard Deviation = 39.8%

The EP toxicity lead results for the samples were statistically analyzed. Three different statistical techniques were used to determine if the average value of EP lead in the target population is above the regulatory threshold and at what confidence level. The first technique, normal hypothesis testing, tested the hypothesis that the mean was greater than 5 mg/L. This technique assumes that the EP toxicity lead data set was drawn from a population having a Gaussian (normal) distribution. The second technique, the Wilcoxon Sign Rank test, tested the hypothesis that the median was greater than 5 mg/L. This technique makes no assumptions about the distribution other than the distribution is symmetrical about the median (Conover, 1980) and allows use of the test in situations where other parametric methods would not apply (e.g., non-normal distribution). The third technique is called the sign test. It also tests the hypothesis that the median is greater than 5 mg/L, but makes no assumptions about the parent population distribution. Thus, this test may be applied to any data set regardless of the distribution.

No statistical evidence existed to suggest rejection of any of the hypotheses at the 95 percent confidence level. Therefore it can be concluded that the average concentration of EP lead in the parent population is above the regulatory threshold of 5ppm. Results of these tests are summarized in Table 6.2.

TABLE 6.2

Results of Statistical Analysis of
the Initial Six Samples
(mg/L)

Statistical Test	Result	Level of Confidence
Normal	Mean > 5	97%
Wilcoxon	Median > 5	98%
Sign	Median > 5	98%

From the results of the initial sampling shown in the previous table it was determined that the six samples were insufficient to make the necessary conclusions. There were two reasons the six samples were deemed insufficient. First, the samples were taken only from surface material. This means that the sampled population was taken from the top few inches of the material. With such a small sampled population it might not be possible to prove that 1000 kg of material was made available to be sampled. This would not meet the minimum quantity of hazardous waste necessary. Second, the samples were not collected with any particular random sampling strategy. This means that sampler bias might have affected the results of the investigation. If the results of the initial sampling were not so close to the regulatory threshold, this bias and the fact that sampling error was never considered might not be a problem, but the 1000 kg minimum weight would still be a problem. Therefore additional sampling was required.

6.2 Second Sampling Episode

Before further sampling could be implemented, it was necessary to obtain additional site information in order to design a proper sampling plan. This information was collected from telephone interviews with the original sampling team.

The material was contained in two open gondola railroad cars. Each railroad car had a length of 52' 5", a width of 9' 6", and a depth of 5', and a capacity of 88,000 kg. The waste material was described as a dry, soil-like material. The optimum target population was the entire contents of both railroad cars. Samples were to be taken with a split spoon sampler.

Before sampling could begin, the number of samples to be taken had to be determined. In this case there was information available from the initial sampling episode upon which to base the number of samples that needed to be collected.

As stated previously prior information is usually not available or not representative of the current environmental situation. The number of samples to be taken was calculated from the EPA manual entitled "Test Methods for the Evaluation of Solid Waste, Physical/Chemical Methods," EPA Publication SW-846 (Second and Third Editions). This formula should be applied only when the sample weight is adequate.

For this sampling episode, the adequate sample weight was not known. The reason for this was lack of knowledge of proper sampling theory at that time. However, this lack of knowledge did not affect the results of the study in any way. The formula used to calculate the number of samples is:

$$N = t^2 * S^2 / D^2 \tag{6.1}$$

where:

N = the number of samples
t = the appropriate student t value
S = the standard deviation of the samples
D = the regulatory threshold minus the average

The standard deviation previously calculated from the six samples was used (3.9 mg/L). The value of D is the difference between the regulatory threshold and the mean from the previously collected samples (9.8 mg/L - 5 mg/L = 4.8 mg/L). The calculated number of samples was eight at an alpha of .01 (99 percent confidence). Due to the closeness of the EP lead value to the regulatory threshold and the possibility that the data already collected might be biased, additional precautions were taken. It was decided that it would be prudent to be able to determine the difference between eight and five mg/L (D = 3) in case the true mean is as low as 8 mg/L. The results of the formula (equation 6.1) provided by the EPA manual suggested that 20 samples would be necessary provided the samples were of the correct weight. The decision was made that 25 samples would be collected as a margin of safety in case there was any error or bias in the initial results. The next process was to randomly locate the sample sites within the target population.

6.2.1 Design of sampling plan

The target population excluded the material within 3" of the sides of the railroad cars. This was done to avoid possible contamination from the sides of the railroad cars. A grid of one foot spacings was superimposed over the surface area of the target population. One foot spacings were used to circumvent disturbance of adjacent possible sampling points. Sample coordinate points were obtained by a computerized random number generator for the simple random sampling scheme. Table 6.3 lists the sample coordinate points within each railroad car. The coordinates represent the center of the one foot squares from which the samples were collected.

6.2.2 Sample Plan Implementation

After designing the sampling plan, implementation was the next phase. For this sampling plan, there was no consideration given to any of the previously discussed errors, only that random location of the sampling points was necessary. Upon inspection of the site, it was determined that the entire depth could not be cored with a split spoon. The maximum depth that could be sampled was one foot. A calculation was done to determine if the sampling of the top one foot of the material contained in the railroad car was enough to meet the requirement of at least 1,000 kg of waste material. This was necessary to ensure that the generator would be classified as a large quantity generator.

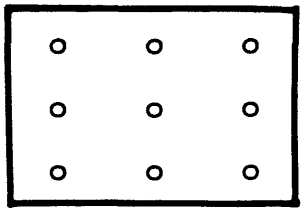
Since one foot was found to be adequate, sampling proceeded. Each sample was made up of nine one-foot cores taken in the following pattern as shown in Figure 6.1.

TABLE 6.3

Sampling Coordinates

Railroad Car 1			Railroad Car 2		
Station	X Value	Y Value	Station	X Value	Y Value
1	3	3	14	3	4
2	4	9	15	8	1
3	5	8	16	10	2
4	10	4	17	10	6
5	12	0	18	12	6
6	13	4	19	14	4
7	13	9	20	24	5
8	20	5	21	24	9
9	26	8	22	26	2
10	36	0	23	35	4
11	36	7	24	40	3
12	37	4	25	50	0
13	41	0			

Note: Grid value 0,0 was located 3" back from the front of the railroad car and 3" in from the side of the railroad car. All other values were measured in feet from this location. The sample masses collected are listed in Table 6.19.



Core Locations Within the One Foot Square Sample Area
(not to scale)

Figure 6.1

The samples were collected and homogenized in the field by manual agitation. After homogenization, the sampled material was split into two portions by the alternate shoveling technique (Pitard, 1988). One portion was maintained for analysis, the other was made available to the designated company representative. It was assumed that the Fundamental Error associated with the splitting process was minimal because of the proper splitting technique and sample weight. In actuality, these errors can become quite large if proper splitting techniques are not used (Pitard, 1988).

Quality control samples were collected at the site. This sampling included the collection of equipment contamination rinsate samples, collected from the rinsing of the split spoon between each of the collected samples. Three field triplicates (13,22,24) were collected to determine sampling precision. The field triplicate consisted of nothing more that collection of more sample mass to enable three EP analysis.

6.3 Extraction Procedure and Analysis

EP Toxicity testing was conducted in accordance with the test methods cited in the Code of Federal Regulations, Title 40, Part 261, July 1, 1987. The test methods are contained in "Test Methods for the Evaluation of Solid Waste, Physical/Chemical Methods", EPA Publication SW-846 [Second Edition, 1982 as amended by Update I (April, 1984), and Update II (April, 1985)]. The extractions were conducted by SW-846 Method 1310. SW-846 Method 3010 was used to digest the extracts and SW-846 Method 7420 (Direct Aspiration, Atomic Absorption Spectroscopy) was used to measure the lead in the digested extracts. The method of standard additions was used to quantify the lead in the extracts.

Results of the accuracy and precision of the analysis are reported in Table 6.4A,B.

For three sampling points (13,22,24), three separate aliquots of the sample were extracted. The replicate extract values indicated the variability due to sampling error, sample preparation error, and analytical error. For three extracts (06,08,11), three separate aliquots of each

TABLE 6.4A

Precision of EP Analysis

Extraction Triplicate Data

Sample No.	Average (mg/L)	%RSD
13	22	9.8
22	17.5	10.7
24	12.4	6.7

Digestion Triplicate Data

Sample No.	Average (mg/L)	%RSD
06	4.5	3.4
08	22.8	3.8
14	13.3	6.3

Measurement Triplicate Data

Sample No.	Average (mg/L)	%RSD
07	9.87	1.3
11	15.9	3.5
20	11.1	3.6

%RSD = percent relative standard deviation
%DEV = percent deviation of found value from the actual value
%REC = percent recovery of spike

TABLE 6.4B

Accuracy of EP Analysis

Control Sample Data

Sample ID	Type	True Value (mg/L)	%DEV
EPA-283-2	Digest	8.0	- 4.9
EPA-283-1	Measure	22.0	- 6.8

Spiked Sample Data

Sample No.	Type	True Value (mg/L)	%REC
06	Digest	2.5	106.
08	Digest	20.	99.5
14	Digest	10.	99.0
07	Measure	10.	100.
11	Measure	20.	99.5
20	Measure	10.	103.

%DEV = percent deviation of found value from the actual value
%REC = percent recovery of spike

extract were digested. The replicate digestion values indicate variability due to sample handling after extraction and analytical variability. For three digested extracts (07,11,20), three separate aliquots were analyzed for lead. These replicate values indicated the variability due to the measurement process.

Precision indicated by these quality control measures was quite good (less than seven percent) and more than adequate to meet the objectives of the investigation. Accuracy measures included the analysis of samples of known value and analysis of spiked samples. A standard reference material of known lead content was digested and analyzed to access bias due to the digestion and measurement procedures. A second standard reference material of known lead content was used to verify the accuracy of the calibration standards.

The extracts, prior to digestion, were screened for lead levels by SW-846 Method 6010 (Inductively Coupled Plasma Atomic Emission Spectroscopy) (3rd Edition) to determine appropriate spiking levels for SW-846 Method 3010 digestion spiked samples. Three digestion spiked samples and three post-digestion (measurement) spiked samples were analyzed. Accuracy of the determinations indicated by these quality control measures was quite good (less that seven percent deviation) and more than adequate to meet the objectives of the investigation.

Field equipment rinses, extraction blanks and digestion blanks were analyzed to verify that sample handling contributed no significant contamination. No lead was detected in these blanks at detection limits of 0.2 mg/L for SW-846 Method 7420 and 0.05 mg/L for Method 6010. Results of the analysis for EP lead are presented in Table 6.6.

The same statistical tests were performed on the 25 samples as on the six samples and, as expected, the confidence level of the 25 samples was much higher. These results are summarized in Table 6.5.

Table 6.5

Statistical Results from
the 25 Samples

Statistical Test	Result	Level of Confidence
Normal	Toxic	99.99%
Wilcoxon	Toxic	99.99%
Sign	Toxic	99.98%

Table 6.6

EP Lead Determination
(mg/L)

Sample	EP Lead	Sample	EP Lead
01	8.5	14	13.3
02	9.9	15	9.1
03	7.7	16	9.0
04	8.1	17	4.4
05	6.3	18	4.3
06	4.5	19	9.8
07	9.9	20	11.1
08	22.8	21	10.5
09	10.5	22	17.5
10	14.8	23	23.0
11	15.9	24	12.4
12	12.1	25	7.6
13	7.0		

Average = 10.8 mg/L
Standard Deviation = 4.9 mg/L
Relative Standard Deviation = 45.4%

In addition, other statistical tests were performed. These included two tests for normality, the Chi Square Test (Chi Square = 3.36) and the Kolmogorov-Smirnov test (estimated overall statistic DN = 0.164). Neither test rejected the normality assumption necessary for the confidence interval test. Other tests were performed to determine if the two results of the two sampling episodes were from the same population. From the results of these tests it can be concluded that the two sampling episodes were from the same population (see Table 6.7). This suggests that there was no bias during the first sampling episode.

Table 6.7

Comparison of EP Data Sets

Test	Conclusion
T-Test of Means	Same Population
Mann Whitney	Same Population
Kolmogorov-Smirnov	Same Population

The data infer that the mean of the target population of the toxic constituent EP lead is above 5ppm. All underlying assumptions possible to verify have been checked and do not violate any conditions of the statistical tests used. The results of the six sample and the 25 sample analysis give similar conclusions.

6.4 Fundamental Error of Samples Collected

The maximum Fundamental Error associated with the sampling program, if only particle size representativeness is considered (assuming that there are at least six occurrences of contamination in each sample), can be calculated from the basic Pierre GY equation:

$$S^2_{(FE)} = 18*f*e*d^3/M_S \qquad (6.2)$$

where:

$S^2_{(FE)}$ = Fundamental Error
f = shape factor (assumed 0.5)
e = density of material (2.5 gm/cm^3)
d = size of largest particle ($1cm^3$)
M_S= mass of the sample

The calculated Fundamental Error from equation 6.2 is 47 percent. This means that if only one 100 gram sample were taken, the error from the Fundamental Error alone would be 47 percent. This error does not include any of the other errors previously mentioned (such as Grouping and Segregation Error, Increment Extraction Error and Long Range Error). If the Fundamental Error is greater than 17 percent, it means that an inadequate amount of sample was taken and that the population distribution is deviating from normal (Pitard, 1988). If the distribution is no longer normal, the results of the sampling episode will not accurately describe the parent population. Clearly, 47 percent is much larger than 17 percent, which indicates that a 100 gram sample was too small a mass to representatively sample the parent population if the Fundamental Error is at a maximum. If this is the case, more than 100 grams of material needs to be extracted to reduce the error to 17 percent or below per sample. It is impossible to know if the Fundamental Error was indeed as large as 47 percent unless some type of field analysis was performed. If field analysis is not performed, the 47 percent error must be assumed and the sample weight increased to lower the Fundamental Error to 17 percent or less. In calculating the 47 percent Fundamental Error, the reduced form of the Fundamental Error equation was used. This reduced form of the Fundamental Error equation (5.4) assumed that there were at least six occurrences of EP lead contamination in the 100 gram sample. If there are not at least six occurrences of EP lead contamination in the 100 gram sample, the complete form of the Fundamental Error equation (5.1) should be used, which will increase the Fundamental Error. This assumption should also be checked before the reduced form of the

Fundamental Error equation is used. The only way to check this assumption before sampling commences is also with some type of field analysis.

The first experiment (X-ray) will be to determine the heterogeneity of the constituent total lead with respect to the different particle sizes. In addition, X-ray analysis can also determine if there are at least six occurrences of total lead in each sample. With the results of this experiment, the actual Fundamental Error can be measured. The Fundamental Error can be determined from the factors that make up the Sampling Constant (C). The Sampling Constant (f*b*c*g) is defined as the product of the Shape Factor (f), the Liberation Factor (b), the Mineralogical Factor (c), and the Granulometric Factor (g).

6.5 X-Ray Experiment to Measure Fundamental Error

From the 25 samples taken to determine the EP lead content, five were chosen for extensive X-ray analysis. The five samples used for the X-ray experiment were samples 02, 06, 07, 08, and 21. It was determined from observation of the samples that the maximum size sieve allowing 95 percent of the particles to pass through was one centimeter. This was also the approximate size of the largest particles in the sample. The samples were weighed prior to any analysis. After weighing, the samples were sieved with a No. 18 (1mm) plastic screen. From the +1mm fraction, the 1cm particles were also weighed. From this, sampling error of the 1cm and the 1mm particle sizes can be determined. The sample fraction weights are presented in Table 6.8.

Table 6.8

Sample Weights for X-Ray Analysis
(grams)

Sample	Initial Weight	<1mm Weight	+1mm Weight	1cm Weight
02	696	370	312	75
06	1095	735	330	97
07	865	374	471	143
08	431	149	282	45
21	915	298	717	156

From each of the <1mm samples, three 5 gram subsamples were taken. These samples were placed in plastic containers for X-ray analysis. From each of the 1cm fractions, three subsamples were also taken. The particles were reduced in size with a spatula to approximately 1mm and placed in a plastic sample cup. This reduction was necessary to ensure a relatively flat sample surface required for accurate X-ray analysis.

Analyses were performed with a Kevex 8000 Energy Dispersive X-Ray Florescence (XRF) Excitation and Detection System. This particular instrument is not portable but yields similar results to a portable X-ray.

For accuracy quality control standards, two known reference materials were analyzed with the samples--PACS and NBS-1648 (see Table 6.9).

Table 6.9

Accuracy of Lead
X-Ray Analysis
(mg/kg)

Standard Reference Material					
PACS			NBS-1648		
Actual Value	Measured Value	%Dev	Actual Value	Measured Value	%Dev
404	364	-10.0	6550	6550	0.0
404	378	- 6.4	6550	6900	5.3
404	358	-11.4	6550	6900	5.3
404	390	- 3.5	6550	6900	5.3
404	360	-10.9	6550	6800	3.8

For an instrument precision measurement, one subsample of each sample was measured in triplicate, one measurement on each of three consecutive days. The results are shown in Table 6.10.

The precision and control sample results were more than adequate to meet the requirements of this experiment. Results of the total lead analyses are presented in Table 6.11.

6.6 Conclusions From X-Ray Analysis

There are four possible approaches for determination of the proper sample size from the X-ray analysis. All four methods should give similar conclusions within their limitations and can be used as a check against each other. Results of the X-ray analyses were consistent with the concept that greater heterogeneity exists with larger sized particles.

6.6.1 Method One for Proper Sample Size Determination

As a first comparison, the standard deviation of a small sample mass can be compared to the standard deviation of a large sample mass. This is basically the empirical approach of Ingammels and Visman (Ingammels and Pitard, 1986) which can be used to obtain a rough idea of the heterogeneity of the waste material.

Table 6.10

Precision of Lead
X-Ray Analysis

Sample Number	Measurement One	Two	Three	Average	Standard Deviation	RSD%
02S	0.53	0.53	0.54	0.54	0.01	1.1
02L	0.90	0.89	0.87	0.89	0.02	1.7
06S	0.58	0.59	0.56	0.57	0.02	2.7
06L	0.61	0.61	0.61	0.61	0.00	0.0
07S	0.38	0.38	0.38	0.38	0.00	0.0
07L	0.34	0.34	0.33	0.34	0.01	1.7
08S	0.60	0.59	0.60	0.60	0.01	0.9
08L	0.54	0.53	0.55	0.54	0.01	1.9
21S	0.49	0.48	0.49	0.49	0.01	1.2
21L	0.53	0.54	0.53	0.53	0.01	1.1

Mote: L refers to the lcm particles of the sample.
S refers to the <lmm particles of the sample.
All measurement results are in percent lead.
Standard deviation is in percent lead.
RSD = relative standard deviation.
Weight of analyzed sample is .25 grams.

Table 6.11

Results of X-Ray Analysis.

Station Number	Subsample One	Two	Three	Average	Standard Deviation	RSD%
02S	0.54	0.51	0.47	0.51	0.04	6.9
02L	0.89	0.41	0.44	0.58	0.27	46.4
06S	0.57	0.53.	0.50	0.53	0.04	6.6
06L	0.61	0.49	0.57	0.56	0.06	10.9
07S	0.38	0.35	0.35	0.36	0.02	4.8
07L	0.34	0.51	0.61	0.49	0.14	27.9
08S	0.60	0.57	0.56	0.58	0.02	3.6
08L	0.54	0.50	0.53	0.52	0.02	4.0
21S	0.49	0.45	0.45	0.46	0.02	5.0
21L	0.53	0.34	0.45	.44	0.09	21.7

Note: L refers to the 1cm particles of the sample.
S refers to the <1mm particles of the sample.
All subsample results are in percent lead.
Standard deviation is in percent lead.
RSD = relative standard deviation.
Weight of analyzed sample is .25 grams.

There are four possible conclusions from this comparison. First, both standard deviations are large. This means that both sample sizes are too small and even larger sample sizes should be compared because no quantitative information is available from the current samples. If this trend cannot be reduced with larger samples, large scale heterogeneity, local trends or cycles are not adequately being dealt with and geostatistical methods need to be applied.

Second, the standard deviation of the small samples is smaller than the standard deviation of the large samples. If this is the case, something is wrong. Contamination may exist at very high levels and the low results might have occurred by chance. Again, more samples need to be taken at large sample weights to determine any useful information.

Third, the standard deviation of the large samples is smaller than the standard deviation of the small samples. This is the usual case for environmental sampling. If the standard deviation of the large sample is adequate for the sampling strategy, then the large sample size is correct.

Fourth is the case where the standard deviations of the large and small sample are approximately equal and both are small. In this case, if the standard deviation is acceptable, the small sample size is sufficient. This method should be used more for a guide than for a sole decision-making tool. If sampling decisions are to be based on this method, great care needs to be taken to ensure that no biases exist (e.g., no Materialization Error) and that enough samples are analyzed (approximately 30).

From the X-ray data, it is necessary to artificially composite the samples to make larger samples as only one sample weight can be analyzed at a time with the X-ray (0.25 grams). In order to accomplish this, the three subsamples (0.25 grams each) will be averaged into one result. The standard deviation of the five 0.75 gram samples can be compared with the fifteen 0.25 gram samples for both the large and small particles. The standard deviations are presented in Table 6.12

The previous discussion indicates that larger variances occur with the largest particles, which is indeed true as indicated by Table 6.12. From standard deviation data, it can be seen that, for the small particles, the 0.25 gram sample is large enough if a RSD of 17 percent can be tolerated. For large particles, the result is different. Samples of 0.75 grams are needed to reduce the RSD to the same level as the small particles. Since the error of the

TABLE 6.12

Standard Deviations of Large and Small Samples

Large Particles

Weight	Average	Std. Dev.	%RSD
0.25	0.518	0.133	25.6
0.75	0.518	0.056	10.8

Small Particles

Weight	Average	Std. Dev.	%RSD
0.25	0.488	0.080	16.4
0.75	0.488	0.084	17.1

Note: Weight in grams.
Average in is percent lead.
Std. Dev. in percent lead.

large particles can be reduced to the error of the small particles with only 0.75 grams, the error of a 100 gram sample would be too small to measure. While this method gives an idea as to the heterogeneity of the sample, it has a serious practical drawback. If the standard deviation that was deemed acceptable is not reached until a sample weight of 35 grams is achieved, it will take 140 0.25 gram X-ray analyses to make each 35 gram sample. Since the largest particles contain the most error, it is only necessary to do this comparison with the largest particles.

If the largest particles are very large, it will be difficult to analyze the content at a 0.25 grams at time.

6.6.2 Method Two for Proper Sample Size Determination

A second method of particle size analysis to determine adequacy of sample size is a technique developed by Visman. The Visman technique determines the difference of the standard deviation between large and small samples. From this difference in standard deviations, the optimum sample weight can be determined to reduce this error to the desired level of precision.

This method is a mathematical expression of the previous method. If the sample weight is above this optimum weight, the error is within the prescribed limits and not a factor in the analysis. If the weight is less than the optimum weight, an error can be calculated. If this error term is too large to meet the objectives, larger samples must be analyzed. Visman allows the artificial compositing of smaller samples to obtain larger ones. This is accomplished by averaging results of single samples.

In this case five samples were composited into one large sample. This resulted in comparing 15 small samples with three larger ones. Results of this test show that the optimum weight is below the actual sample weight and indicate that only a small amount of sample is necessary.

The data for Table 6.13 and Table 6.14 are from the results presented in Table 6.11. The data in Table 6.15 are from the analysis of all 25 small particle samples (see Table 6.18).

The Visman equations are:

$$A = W_1 * W_2 * (V_1 - V_2)/(W_2 - W_1) \tag{6.3}$$

$$B = V_2 - A/W_2 \tag{6.4}$$

$$V_3 = A/W_T + B/N \tag{6.5}$$

$$W_{OPT} = A/B \tag{6.6}$$

$$W_{MIN} = A/(AVG-L)^2 \tag{6.7}$$

$$K = A/(V^2 * X_2{}^2) \tag{6.8}$$

where:

B = segregation constant
A = homogeneity constant
W_1 = weight of small samples
W_2 = weight of large samples
V_1 = variance of small samples
V_2 = variance of large samples
V_3 = total sampling variance
W_T = total weight of all samples
N = number of increments per sample
W_{OPT} = optimum sample weight

W_{MIN} = minimum sample weight

L = background level of contaminant

K = sampling error

X_2 = average value of large particles

If the Segregation Constant (B) is small compared the Homogeneity Constant (A), there is a minimum segregation error and the minimum weight can be used instead of the optimum weight. If the Segregation Constant (B) is large compared to the Homogeneity Constant (A), the optimum weight should be used. From the above results the minimum weight that should be used, taking into account the analysis from the first five samples is 3.44 grams. Results from Table 6.14 and Table 6.15 might be a little misleading because of the small number of samples used. If all 25 samples are considered, in which case a better approximation of the optimum sample weight can be made, the optimum weight is only 0.44 grams.

The Visman results for the 25 samples only used the total lead concentrations for the small particles, but this was appropriate since the small and large particle results of total lead concentration differ only slightly. The estimation of the optimum sample from only five samples is very conservative and can be used to estimate the result of

Table 6.13

Small Particle Visman Results

Sample Weight	0.25
Grouping Factor	5
Sample Composite 1 Average	0.524
Sample Composite 2 Average	0.436
Sample Composite 3 Average	0.504
Average of all Samples	0.488
Weight of Original Samples	0.25
Weight of Composited Samples	1.25
Variance of Original Samples	0.0064
Variance of Composited Samples	0.0021
Homogeneity Constant (A)	0.0013
Segregation Constant (B)	0.0011
Optimum Weight	1.24
Minimum Weight	0.07
Weight for 1% Sampling Error	55.7
Weight for 10% Sampling Error	0.56

Note: All weights are in grams.

All averages are in percent lead.

Table 6.14

Large Particle Visman Results

Sample Weight	0.25
Grouping Factor	5
Sample Composite 1 Average	0.568
Sample Composite 2 Average	0.514
Sample Composite 3 Average	0.470
Average of all Samples	0.517
Weight of Original Samples	0.25
Weight of Composited Samples	1.25
Variance of Original Samples	0.0175
Variance of Composited Samples	0.0024
Homogeneity Constant (A)	0.0047
Segregation Constant (B)	0.0014
Optimum Weight	3.44
Minimum Weight	0.15
Weight for 1% Sampling Error	177.
Weight for 10% Sampling Error	1.8

Note: All weights are in grams.

All averages are in percent lead.

Table 6.15

Small Particle Visman Results

Sample Weight	0.25
Grouping Factor	5
Sample Composite 1 Average	0.422
Sample Composite 2 Average	0.486
Sample Composite 3 Average	0.472
Sample Composite 4 Average	0.512
Sample Composite 5 Average	0.588
Average of all Samples	0.496
Weight of Original Samples	0.25
Weight of Composited Samples	1.25
Variance of Original Samples	0.0076
Variance of Composited Samples	0.0037
Homogeneity Constant (A)	0.0012
Segregation Constant (B)	0.0027
Optimum Weight	0.44
Minimum Weight	0.06
Weight for 1% Sampling Error	49.2
Weight for 10% Sampling Error	0.5

Note: All weights are in grams.

All averages are in percent lead.

the 25 samples. This method, as well as the other methods presented here, are only an approximation of the desired sample weight and should be used accordingly.

6.6.3 Method Three for Proper Sample Size Determination

The third method to calculate the amount of sample necessary to adequately represent the parent population is direct calculation of the Sampling Constant (C) in the Pierre Gy equation for the Fundamental Error (equation 1.1). Two variables can be calculated from the X-ray results and substituted into the equation, the Liberation Factor (b) and the Mineralogical Factor (c), and two that will not, the Shape Factor (f) and Granulometric Factor (g). It is possible to calculate the Shape Factor (f) and the Granulometric Factor (g), but those calculations are beyond the scope of this study and the values for these factors do not significantly affect the end result. The Granulometric Factor (g) will still be assumed to be 0.25. The Shape Factor (f) will still be assumed to be 0.5. The Mineralogical Factor (c) can be calculated by dividing the density of the contaminant by the average total lead content of the 25 samples (expressed as a mass fraction). To determine density of the contaminant, several forms of the most probable compound of lead were considered. Their densities are shown in Table 6.16.

Table 6.16

Density of Common Lead Compounds

Compound	Density (mg/cm^3)
Lead Chloride ($PbCl_2$)	5.9
Lead Oxide (PbO)	9.3
Lead Carbonate ($PbCO_3$)	6.6
Lead Sulfate ($PbSO_4$	6.3

The most probable lead density is approximately 7 mg/cm^3, so that density was used. The average lead content of the large particles was 0.517 percent. From this, the mineralogical factor was 7/.00517 = 1354. The Liberation Factor (b) can also be calculated (equation 5.3) from the X-ray results. The highest concentration in any of the particles was 0.89 percent, which can used for A_{MAX}. From these values the Liberation Factor can be calculated.

$$b = (.0089 - .005)/(1 - .005)$$

$$b = .0038$$

The Fundamental Error can now be calculated (5.1):

$$S^2_{(FE)} = 0.5*0.0038*0.25*1354*1/100 = 0.0064$$

$S = 0.080$

Fundamental Error is 8.0 percent.

6.6.4 Method Four for Proper Sample Size Determination

The fourth method to determine adequate sample size from the X-ray results is to estimate the Sampling Constant (C) for the Pierre Gy method by looking only at the combination of the Factors in the Fundamental Error equation (5.1). The empirical estimation of the Sampling Constant eliminates the inaccuracies of the assumptions. While this method gives an empirical estimation of the Sampling Constant, the individual Factors cannot be determined. Once this Sampling Constant is estimated from the 0.25 gram samples, it can be used to estimate the Fundamental Error for the 100 gram samples. The Fundamental Error is a property of the material and therefore should not change with the different sized samples. Once the Fundamental Error, the diameter of the particles, and the mass of the sample are known, the Sampling Constant can be algebraically determined from equation 6.9.

$$C = (f*b*g*c) = S^2_{(FE)}*M_S/d^3 \qquad (6.9)$$

The error for a certain weight at a size of 1cm has already been determined (the maximum error is 46.4 percent for 0.25 grams). From this error and sample weight, the sampling constant can be determined. The resulting Sampling Constant is 0.0538 for the largest particles (which contain the most error). The variance for a 100 gram sample is very small (0.00054). This yields a Fundamental Error of 2.3 percent for a 100 gram sample. This is quite different than the 45.4 percent calculated from the results of the 25 EP samples and suggests that there is some other form of error not related to the total lead Fundamental Error.

6.7 Comparison of Fundamental Error Estimations

While there is some discrepancy between the results of the methods presented to calculate the Fundamental Error (see Table 6.17), the results agree sufficiently to design a sampling plan. The larger Fundamental Error associated with the fourth method is probably due the large number of estimations made. The Visman and the Estimated Sampling Constant methods are congruent and are more direct methods than the others so they can be used with confidence. The Comparison of Variances method is a crude approximation to verify that the other methods are reasonable.

All the samples were measured for total lead on the small particles, which adequately predict the total lead concentration. The results are presented in Table 6.18

Table 6.17

Comparison of Estimated Fundamental Error

Method	Fundamental Error for 100 Gram Sample
Comparison of Variances	<1 percent
Visman	2 percent
Pierre Gy	8.0 percent
Estimated Sampling Constant	2.3 percent

Table 6.18

Total Lead Analysis (%)

Sample	Total Lead	Sample	Total Lead
01	.36	14	.48
02	.51	15	.44
03	.39	16	.52
04	.50	17	.45
05	.35	18	.46
06	.53	19	.53
07	.36	20	.60
08	.58	21	.46
09	.49	22	.60
10	.47	23	.74
11	.48	24	.57
12	.50	25	.57
13	.46		

6.8 Determination of EP Lead Sampling Constant

From the calculations and the experiments performed thus far, there are two possibilities to explain the difference between the calculated Fundamental Error (approximately 2 percent) and the observed total error (45.4 percent). One possibility is that the total lead sampling constant is different than the EP lead sampling constant. The other is that there is some other error affecting the results.

If the Sampling Constant for the total lead is the same as the Sampling Constant for EP lead, it would be possible to predict the EP lead Sampling Constant from the total lead Sampling Constant. The easiest prediction method would be regression analysis between the EP lead results and the total lead results. The R-Squared for the regression of EP lead and total lead is 36.4 percent which indicates that there is not a strong correlation. The regression analysis was checked to ensure that the residuals were normally distributed, had an average of zero, were uncorrelated, and had equal variances. These checks ensure the validity of the regression analysis (Neter, Wasserman, and Kunter, 1985) and were performed on all the regression analyses in this thesis. This means the total lead Sampling Constant cannot be used to predict the EP lead Sampling Constant and that the two differ. This difference in the Sampling Constants should at least partially explain the difference between the observed total error and the calculated Fundamental Error.

It was necessary to determine what controlled the EP lead Sampling Constant. The first procedure was to perform a particle size analysis to determine if there was a different concentration of material in the larger and smaller particles that was controlling the total lead Sampling Constant. Regression analysis was used to determine if any correlation exists. The weights of the various fractions are listed in Table 6.19.

From regression analysis it can be determined that there was no significant correlation between particle size and EP lead (R-squared = 15.4 percent) concentration.

The only other information easily available was the pH values obtained during the extraction procedure. There were three pH values of potential importance in the analysis: the initial pH at the beginning of the extraction procedure, the minimum pH obtained at any point during the extraction procedure, and the final pH at the end of the procedure. All pH measurements were taken with a Beckman 70 pH Meter. The measurements are as presented in Table 6.20.

Table 6.19

Sample Weights of 25 EP Samples
(grams)

Sample Number	Initial Weight	-1mm Fraction	+1mm Fraction	+2mm Fraction
01	700	407	293	225
02	696	370	312	135
03	918	193	725	521
04	825	380	435	346
05	512	264	238	193
06	1095	735	330	205
07	865	375	471	263
08	431	149	282	107
09	692	260	402	287
10	817	196	612	491
11	508	147	351	262
12	1179	303	876	715
13	380	234	146	96
14	596	236	360	284
15	579	267	312	250
16	317	181	136	100
17	718	332	386	300
18	376	191	185	138
19	977	464	513	381
20	867	240	627	470
21	915	270	645	448
22	530	211	319	256
23	685	198	487	385
24	480	217	263	207
25	1520	887	633	477

Table 6.20

Extraction Procedure pH Values

Sample	Initial	Minimum	Final	Expini	Expmin	Expfin
1	10.00	5.17	5.38	1.0E-10	6.8E-06	4.2E-06
2	9.83	5.14	5.38	1.5E-10	7.2E-06	4.2E-06
3	9.77	5.11	5.48	1.7E-10	7.8E-06	3.3E-06
4	10.09	5.09	5.54	8.1E-11	8.1E-06	2.9E-06
5	9.90	5.14	5.50	1.3E-10	7.2E-06	3.2E-06
6	10.19	5.08	5.67	6.5E-11	8.3E-06	2.1E-06
7	9.98	5.15	5.37	1.0E-10	7.1E-06	4.3E-06
8	9.90	5.00	5.04	1.3E-10	1.0E-05	9.1E-06
9	9.85	5.06	5.31	1.4E-10	8.7E-06	4.9E-06
10	9.65	5.18	5.18	2.2E-10	6.6E-06	6.6E-06
11	9.99	4.94	5.26	1.0E-10	1.2E-05	5.5E-06
12	10.09	5.01	5.33	8.1E-11	9.8E-06	4.7E-06
13	10.11	5.21	5.49	4.3e-10	7.0E-06	3.3E-06
14	10.17	5.09	5.33	6.8E-11	8.1E-06	4.7E-06
15	10.10	5.04	5.54	7.9E-11	9.1E-06	2.9E-06
16	10.06	5.14	5.49	8.7E-11	7.2E-06	3.2E-06
17	9.69	5.59	5.63	2.0E-10	2.6E-06	2.3E-06
18	10.20	5.17	5.76	6.3E-11	6.8E-06	1.7E-06
19	10.60	5.08	5.36	2.5E-11	8.3E-06	4.4E-06
20	10.20	5.21	5.17	6.3E-11	6.2E-06	6.8E-06
21	9.90	5.19	5.27	1.3E-10	6.5E-06	5.4E-06
22	9.96	5.23	5.16	3.5E-11	6.0E-06	6.9E-06
23	9.97	5.23	5.11	1.1E-10	5.9E-06	7.8E-06
24	10.10	5.24	5.43	6.4E-10	6.5E-06	3.8E-06
25	10.17	5.17	5.59	6.8E-11	6.8E-06	2.6E-06

Note: Expini is the initial hydrogen ion concentration (M).
Expmin is the minimum hydrogen ion concentration (M).
Expfin is the final hydrogen ion concentration (M).

Regression analysis showed that there was a correlation between the final hydrogen ion concentration and the EP lead concentration. The R-squared was 83.2 percent. The model is EP lead = Expfin * 2.438E6. The observed, predicted and residual values are presented in Table 6.21.

If the cause of the soil alkalinity (which is the only thing that can cause a difference of final pH) could easily be determined, then perhaps the EP lead sampling constant could be determined with a portable X-ray. The easiest measure of soil alkalinity would be to determine the concentration of the alkali and alkaline earth elements in the EP extracts. The analyses were performed by the same method used to screened the EP extracts (SW-846 Method 6010). Results of these analyses are shown in Table 6.22.

Regression analysis was performed with these elements and the EP lead values. The only significant element was Ca, which had a R-squared of 66.8 percent with EP lead. This was not an adequate model to predict EP lead so two other analyses were performed--total sulfate and total carbonate.

Sulfate was measured (SW-846 Method 6010) as it might precipitate out the lead during the extraction procedure into the sludge and thus not be available in the leachate.

Table 6.21

Prediction of EP Lead from Expfin

Sample	Actual Value	Predicted Value	Residual
1	8.5	10.2	-1.7
2	9.9	10.2	-0.3
3	7.7	8.1	-0.4
4	8.1	7.0	1.1
5	6.3	7.7	-1.4
6	4.5	5.2	-0.7
7	9.9	10.4	-0.5
8	22.8	22.2	0.6
9	10.5	11.9	-1.4
10	14.8	16.1	-1.3
11	15.9	13.4	2.5
12	12.1	11.4	0.7
13	7.0	8.0	-1.0
14	13.3	11.4	1.9
15	9.1	7.0	2.1
16	9.0	7.9	1.1
17	4.4	5.7	-1.3
18	4.3	4.2	0.1
19	9.8	10.6	-0.8
20	11.1	16.5	-5.4
21	10.5	13.1	-2.6
22	17.5	16.9	0.6
23	23.0	18.9	4.1
24	12.4	9.2	3.2
25	7.6	6.3	1.3

Table 6.22

ICP Analysis of EP Extracts for Alkali and Alkaline Earths

Sample	Element		
	Ca	K	Na
1	1150	10	991
2	918	11	1170
3	1060	10	1090
4	1040	10	1190
5	1140	10	1180
6	1110	11	1200
7	1050	10	926
8	739	10	1160
9	942	11	1030
10	577	10	1280
11	723	10	1230
12	1040	10	1260
13	1080	10	1070
14	801	10	1310
15	1030	10	1070
16	927	10	1290
17	1110	37	1110
18	1160	10	1080
19	934	11	1180
20	804	18	1270
21	973	17	1080
22	795	12	1285
23	570	10	1350
24	926	9	1273
25	817	11	1350

Note: ALl results are in mg/L. Other metals that might contribute to the soil alkalinity (Mn, Mg, Ba) were also measured, but the concentrations were so small that they were ignored. Quality Control was performed on the analysis with acceptable results of +/- 10 percent relative standard deviation.

Regression analysis showed no correlation with the sulfate concentration in the EP leachate and the EP lead values (R-Squared = 28.7 percent). The results of the EP sulfate analysis are presented in Table 6.23

Multiple regression techniques were also used, but no combination of analytes were more significant than simple linear regression between calcium and EP lead.

It appeared that though the EP lead values were related to the soil alkalinity, there was not, at least in this case, an easy method to measure the soil alkalinity. Perhaps some type of multivariate analysis on the X-ray results could directly lead to the variability in soil pH. Field analysis of the soil pH could probably be performed but would be more time consuming than analysis with a portable X-ray.

Interpretation of these experiments for the development of a sampling model will be discussed in the next section.

Table 6.23

ICP Analysis of EP Extracts for Sulfate (mg/L)

Sample	EP Sulfate	Sample	EP Sulfate
01	197	14	290
02	234	15	234
03	195	16	288
04	233	17	207
05	215	18	212
06	260	19	246
07	198	20	274
08	250	21	228
09	232	22	263
10	266	23	350
11	243	24	274
12	201	25	291
13	224		

Chapter 7

DISCUSSION

7.1 Introduction

The use of sampling theory for environmental investigations can be quite complicated. Because of this, the preceding experiments were conducted in order to develop a strategy for sampling theory implementation. Following is an explanation of this strategy.

This study does not cover all the errors associated with sampling theory. The errors given the most attention are the Fundamental Error and the Long Range Error, as they are the most significant and are usually not considered in most environmental investigations. All other errors have been minimized through proper sampling technique and will not be discused (e.g., Grouping and Segregation Error).

Before implementation of any method, it must be determined if there is a problem for the method to solve. This is also the case with environmental investigations. If the characteristics of the waste material are such that the investigator can easily determine whether or not a representative sample can be taken without consideration of certain errors, the sampling becomes much easier. For instance, if the particles of the waste material are micron in size, relatively homogeneous, and contain high levels (e.g., factor of 100 times the regulatory threshold) of leachable contaminants, proper sampling can be accomplished with a minimum of consideration given to the sampling theory and practice described herein. If the waste does not possess these ideal characteristics, proper sampling is a problem that only sampling theory can solve.

7.2 Conclusions From Experiments

The types of sampling errors which affect the accuracy and precision of any sampling episode can be determined with field techniques. This will allow the implementation of a sampling plan that achieves the goals of any investigation without prior knowledge of the waste material. Errors that affect the results of the sampling episode are discussed in this section.

7.2.1 Fundamental Error

The Fundamental Error for this investigation was indirectly measured by the three field triplicates taken during the second sampling episode. Analysis of these triplicates contained all the error associated with sampling theory except for the Long Range Error. Fortunately, it is possible to separate these errors and measure the Fundamental Error directly. The relative errors had to be converted to variances before they could be separated. The variances presented in Table 7.1 are the average variances for the three triplicates.

Table 7.1

Calculation of Sampling Precision
(mg/L)

Extraction Triplicate Variance	Digestion Triplicate Variance	Sampling Error Variance	Sampling RSD(%)
1.24	.37	.87	8.6

Note: Sampling Error is Fundamental Error.

If the field analyses are not available, no assumptions about the waste material characteristics can be made and the calculated Fundamental Error would be:

$$S^2_{(FE)} = f*b*c*g*d^3/M_S$$

where:

$f = 0.5$

$g = 0.5$

$l = .05$

$c = 7/0.0001 = 70{,}000$

$$S^2_{(FE)} = 0.5*0.05*70{,}000*0.25*1/100 = 4.38$$

$S = 2.09$

Fundamental Error equals 209 percent!

In order to have a Fundamental Error of 10 percent the amount of material required to be sampled would be:

$M_S = f*b*c*g*d^3/S^2{}_{(FE)}$

$M_S = 0.5*0.05*70{,}000*.025*1/0.01$

$M_S =$ 43,700 grams of material

If it can be assumed that there are at least six occurrences of contamination in every sample, the reduced form Pierre Gy equation (5.4) can be used. It has already been shown (Chapter 4) that the Fundamental Error for the samples collected in this investigation would be 47 percent. In order to reduce the Fundamental Error from 47 percent to 10 percent, the amount of material to be sampled would be:

$M_S = 18*f*e*d3$

$M_S = 18*0.5*2.5*1/0.01$

$M_S =$ 2,250 grams of material

In this example, if the worst case Sampling Constant were used, the sample mass required for a 10 percent Fundamental Error would be:

$$M_S = (f*b*c*g)*d^3/S^2{}_{(FE)}$$

$$M_S = 0.0538*1/0.01$$

$$M_S = 5.38 \text{ grams of material}$$

It can be seen that there is significant improvement because of the use of X-ray analysis. If a sampling plan were designed from these results, the collection of one 100 gram sample made up of many increments would be adequate. It would, however, be wise to take more than one sample in order to be able to use classical statistics on the EP results. It is only through the use of classical statistics that the calculation of the Fundamental Error can be verified if the investigator is only estimating the Fundamental Error from the X-ray results.

If the Fundamental Error were calculated from the results of the equation where nothing was known (5.1), it would suggest collection of 437 one hundred gram samples to achieve a guaranteed Fundamental Error of 10 percent or less. If some material characteristics are known, the number of 100 gram samples would be reduced to 23. If the Sampling Constant can be determined by X-ray analysis, then only one sample need be collected. While 437 one hundred gram samples is a worst case situation that will most likely never exist in the real world, it must be planned for if some type of field analysis is not available.

The Fundamental Error resulting from this field analysis is 8.6 percent. If it is assumed from Pierre Gy's sampling theory (equation 5.4) that at least six occurrences of contamination exist in every sample, the worst case Fundamental Error would be 47 percent. If the Fundamental Error of 8.6 percent were known before sampling began (instead of the 47 percent calculated from sampling theory) the mass of the sample could be reduced by a factor of 23. This significant reduction of sample mass could make an otherwise impractical sampling episode feasible.

If no assumption can be made regarding the waste material--including at least six occurrences of contamination in every sample--the required sample mass will be 437 times the actual mass necessary! This demonstrates the power of using field techniques along with sampling theory for the reduction of sample mass.

The Fundamental Error associated with EP lead in this experiment was not the same as the Fundamental Error associated with total lead. The EP lead Fundamental Error, being much greater, had to be associated with a Fundamental Error from another source. EP lead error included the Fundamental Error associated with sampling the waste alkalinity. Fundamental Error from sampling soil alkalinity was much greater than the Fundamental Error associated with the total lead (8.6 percent from field triplicates vs. approximately two percent from X-ray analysis). Therefore, the Fundamental Error associated with the total lead content was insignificant (i.e., the 8.6 percent Fundamental Error was from sampling waste alkalinity and not from sampling total lead) when the Fundamental sampling error was determined. The EP lead error can always be expected to be greater than the total lead error due to the nature of the Extraction Procedure, which by design adds more variability. Total lead variability is the lower limit of EP lead variability that will never be reached.

7.4.2 Long Range Error

The field triplicates measured the Fundamental Error. However, each of the three field triplicates produced a different regression model relating EP lead to soil alkalinity (see Table 7.2). This suggested that some type of Long Range Error existed as the waste material characteristics were not homogeneous over long distances. Long Range Error was further confirmed by the analysis of the 25 samples. The expected total error would have been 8.6 percent if there had been no Long Range Error present.

However, the actual total error was 45.4 percent, thereby confirming the suspected presence of Long Range Error.

Table 7.2

Comparison of Field Triplicate Models

Sample	Regression Model	R-Squared
13	EP lead = 3.29 + 1.13 * Expfin	.99
22	EP lead = -30.19 + 6.89 * Expfin	.99
24	EP lead = 7.59 + 1.28 * Expfin	.98

Note: Expfin is the final hydrogen ion concentration (M).

The total error (45.4 percent) was so large that even the Fundamental Error associated with the sampling of soil alkalinity (8.6 percent) became insignificant. The resulting Long Range Error would still have been 44 percent if the soil alkalinity Fundamental Error had been subtracted. This Long Range Error controlled the total sampling error. Therefore, Long Range Error also had to be determined in addition to the Fundamental Error for correct sampling plan design.

7.3 Field Techniques Necessary For Sample Plan Design

There are many analytical instruments available for field analysis, but perhaps the best for sample plan design is the portable X-ray. This instrument will provide the following information:

1. Existence of a potential violation. In this study the total lead concentration had to be greater than 100 mg/kg for the possibility of EP lead to be above the regulatory threshold of five mg/L (the Extraction Procedure diluted the concentration by a factor of 20).

2. Homogeneous distribution of the contaminant in the waste material. If there are at least six occurrences of contamination in every 100 grams of waste material, the reduced form Pierre Gy Fundamental Error equation (5.4) can be used to ensure particle size representativeness.

3. Comparison of waste materials composition. This comparison will allow the separation of different waste streams into similar populations for sampling purposes. It will also allow the investigator to determine if there is any segregation in the waste to be sampled (which might change the sampling strategy).

4. The Sampling Constant for the contaminant of interest. This will determine the sample mass necessary for specific sampling precision if the alkalinity Fundamental Error is not a factor in the total error. The alkalinity is

only a factor if the waste's buffering capacity exceeds that of the acetic acid extraction agent (i.e., all the allotted acetic acid is used during the Extraction Procedure).

5. Concentration of the contaminant in the EP leachate. This can occur only if that concentration is above the limit of detection (LOD) by the portable X-ray. If the concentration is below the LOD concentration of the leachate, another analytical technique must be used. The problem is that the EP test takes 24-28 hours to complete.

6. Numerous analytical measurements of waste. These enable characterization of the entire waste and will allow for the use of geostatistics to map the entire waste area. Geostatistics provides a more accurate measurement of Long Range Error.

The X-ray analysis cannot discover the following:

1. The EP sampling Constant of the contaminant of interest. If there can be a direct measurement of the parameters that cause the soil alkalinity (e.g. Ca concentration), EP sampling constant can be quantified. The soil alkalinity, however, might be too complicated to allow such easy quantification. Fortunately, this constant can be determined indirectly with the final pH of the EP.

2. The average level of the EP for the contaminant of interest. If the level is below the limit of detection for the portable X-ray, other analytical methods must be employed. These methods might include concentration of the leachate to increase the level of contaminant.

7.4 Development of Model Based on Experiment Results

7.4.1 Preliminary Information

The first step in proper sample plan design is to determine the objectives of the sampling plan. If the only purpose is to show contamination, judgement sampling is adequate and sampling theory is not of great concern. If the purpose of the study is to quantify the level of contamination in a waste stream or define a minimum amount of hazardous waste, sampling theory must be considered. As much information as possible should be collected concerning the waste properties before a sample plan is designed.

A very important consideration is the level of confidence the investigator wants to have in the results of the analytical results of the sampling. An answer like "as confident as possible" is not adequate. The difference between a 95 and a 99 percent confidence level could triple the work required of the sampling crew, not to mention the increased analytical workload. Sampling Theory will not only increase the confidence that the sampling plan will meet the objectives of the study, but will allow for better management of resources.

Once the data quality objectives (e.g., desired confidence level) have been determined, information regarding the waste material needs to be determined. Maximum particle size is one of the most important characteristics of the waste material.

When particle size and regulatory threshold have been determined, the maximum Fundamental Error can be calculated if assumptions regarding characteristics of the material are made. If this error would necessitate the collection of a very large sample mass, field techniques can be used to measure the Sampling Constant and determine if the Fundamental Error is actually less than a worst case scenario. For the problem considered in this study, X-ray Florescence is a suitable analytical technique to incorporate into the sampling effort.

7.4.2 USE OF X-RAY TO DETERMINE SAMPLING ERRORS

The first field analysis would be use of the portable X-ray to determine if there is a potential violation. Violation determination can be difficult if the waste material is very heterogeneous with respect to the contaminant of interest. If this is the case, an average level of at least 100 mg/kg of contaminant would be necessary for a potential EP lead violation.

Results of the X-ray analysis can also be used to determine if there are at least six occurrences of contamination is every 100 gram sample. These two pieces of information can be used to determine which form of the Fundamental Error to use (equation 5.1 or equation 5.4).

If there are at least six occurrences of contamination in every sample, the sample need only be representative of the largest particle size. The Fundamental Error of the total lead can then be measured empirically by using the method described in Section 6.6.4

After Fundamental Error is measured the next step is to determine the Long Range Error of the entire waste material for the total level of the contaminant of interest. This can also be measured with the portable X-ray. The Fundamental Error can then be compared to the Long Range Error to see if either error will predominate. If the Fundamental Error is less than one-third the Long Range Error, it can be ignored and vice-versa (Kratochvil and Tylor, 1981). The next step is to determine if the total error thus far calculated is too large to achieve the accuracy and precision goals of the study.

If the resulting error at this point is too large calculating additional errors is futile. The errors must be reduced before sample plan design can continue. If the Fundamental error is the problem, it can be reduced with larger sample mass or particle size reduction.

If the Long Range Error is a problem there are two possible solutions. First, the sample compositing area could be increased, thus reducing the variability. Second, the waste could be segregated into sections that are more alike. Then stratified random sampling can be employed. The practical considerations of the sampling episode must be considered when determining which error to reduce and by how much. For instance, it might be impractical to reduce the particle size or separate the waste into similar strata.

The next step is to determine the Fundamental Error due to waste alkalinity. This additional error must be added to the errors already measured. The EP lead variability can be determined indirectly by the variability of the final hydrogen ion concentration in the leachate. This variability will give the Fundamental Error associated with the waste alkalinity. The Long range Error can be measured in a similar manner.

The sampling error can now be calculated to determine if the accuracy and precision goals of the study can be achieved. If the errors are unacceptably large, they can be reduced by techniques already mentioned.

7.5 Number of Samples to Collect

The Pierre Gy equation calculates the total amount of sample mass needed to achieve a prescribed Fundamental Error. This mass should be broken into several samples, even if it can be collected in one unit. The reason for this is threefold: First, the nature of the investigation might require the calculation of classical statistics, which is impossible with only one sample. Second, there exist many assumptions in the design process. The only check available as to the validity of these assumptions is calculation of the error from multiple samples. Third, field analysis lacks the accuracy and precision of laboratory analysis; biases and relative precision error up to 20 percent are possible. Contaminant levels might be inaccurate by as much as 20 percent.

The number of samples needed for the above three considerations depends on the closeness of the average value to the regulatory threshold and the objectives of the investigation. If more samples are taken, the confidence level of the mean decreases by the factor of $(n)^{.5}$. Conservative measures should always be used in any field calculations, but accuracy and precision still need to be verified.

7.6 Maximum Total Error Allowed

One of the factors that affects the required precision is how close the average value of the contaminant is to the regulatory threshold. It is intuitively evident that if the level of contamination is 100 times the regulatory threshold, the precision required can be much less than if the level of contamination is twice the regulatory threshold. In addition, if the precision is larger that 100 percent relative standard deviation it is impossible to distinguish the results from a value of zero at a 68 percent confidence level. If a 95 percent confidence (two standard deviations) is desired, the maximum relative standard deviation is only 50 percent (half of 100 percent). For a 99 percent confidence level, the maximum is only 33 percent. The lower confidence level is calculated as shown in equation 7.1

$$CI_{lower} = X - t*s/(n)^{.5} \tag{7.1}$$

where:

X = average level of contaminant

t = appropriate t value for specified degrees of freedom and confidence level

s = standard deviation

n = number of samples collected

The lower confidence level must be greater than the regulatory threshold in order to statistically distinguish the result from the regulatory threshold itself. In this study, the maximum amount of standard deviation error that can be tolerated is given by equation 7.2.:

$$s_{max} = (n)^{.5} * (X - RTH)/t \qquad (7.2)$$

where:

RTH = regulatory threshold

s_{max} = maximum standard deviation

In order to determine if the average value of the waste material exceeds the regulatory threshold, the standard deviation cannot exceed 11.6 (25 samples at a 99 percent confidence level). This is larger than the standard deviation of 4.96 actually calculated from the results of the 25 samples in the experiment. It is apparent that an adequate number of samples were collected to meet the goals of the investigation provided the target population is normally distributed at the sample mass used. If, however, the results from the first sampling episode are taken (the first six samples), the maximum standard deviation allowable at a 99 percent confidence level is 3.5. This is smaller than the actual standard deviation of 3.9, which implies that more than six samples had to be collected to meet the 99 percent confidence level.

7.7 Explanation of Actual Analysis

Following are the steps taken in performing the investigation described in the experiment chapter.

The first step was to determine if there is a potential EP violation, which is not a difficult task. In order for a potential violation to exist the total lead concentration must be at least 100 mg/kg. In this study the total lead concentration was approximately 5,000 mg/kg, therefore the potential existed.

The next step was to determine if there are at least six occurrences of total lead in every sample. The X-ray analysis indicated the presence of lead in every 0.25 grams of sample. From this it was concluded that there were at least six occurrences of total lead in every 100 gram sample.

The Fundamental Error for total lead should now be determined. There are several methods that can be used to determine the Fundamental Error of a 100 gram sample (see section 6.6). The method of choice was estimation of the

Sampling Constant. The sampling constant from the .25 gram sample was 0.0538. This gave a Fundamental Error of 2.3 percent for a 100 gram sample. This error was less than 17 percent so the sampling process could continue. If this error had been more than 17 percent, the error should have been reduced by methods already mentioned. Had the final EP value been very large compared to the Regulatory Threshold, more Fundamental Error would have been acceptable, depending on the goals of the investigation.

Long Range Error for total lead should now be measured. X-ray results can also be used for this calculation. The relative standard deviation was 17 percent. This was the Long Range Error which was larger than the Fundamental Error. Therefore, it was the only error that needed to be considered in the total error. Long Range Error should be reduced only if the total error is not acceptable. For this study, 17 percent error is acceptable and allowed the continuation of the investigation.

The next error to calculate is the Fundamental Error of EP lead, which is correlated to the variability of the final hydrogen ion concentration. This error had already been determined as 8.6 percent (see Section 7.2.1). This error was significant when compared to the 17 percent error already calculated, but was still less that the 17 percent maximum tolerated Fundamental Error. At this point the total error was 19 percent.

The final error to determine is the Long Range EP Error. This error is also related to the variability of the final hydrogen ion concentration. The variability of the final hydrogen ion concentration was 45.4 percent. From this 45.4 percent the 8.6 percent Fundamental Error should be subtracted. However, the Fundamental Error was less than three times the Long Range Error so it could be ignored. The errors measured thus far are presented in Table 7.3.

Table 7.3

Sampling Errors of 25 Samples

Type of Error	Error (%)
Fundamental Error of Alkalinity	8.6
Fundamental Error of Lead	2.3
Long Range Error of Alkalinity	44.
Long Range Error of Lead	17.

The controlling error was the Long Range Error and the Fundamental Error was less that 17 percent. If this Long Range Error can be tolerated, then the 25 samples taken were adequate. If this Long Range Error could be reduced, the amount of samples could also be reduced when determining the average level of contamination. The only proper method of reducing the Long Range Error is through stratified random sampling.

If a specified confidence level is required, the total number of samples can be determined. For example, when a 95 percent confidence level is required, the average level of contamination is 10.8 ppm and the standard deviation is 4.96. Application of equation 7.2 yields a required sample number of four.

If stratified random sampling were used to eliminate the Long Range error, different sample mass could be determined for various acceptable Fundamental Errors. If a Fundamental Error of 15 percent is acceptable, the required sample mass would be 100 grams. For a Fundamental Error of 5 percent the required sample mass would also be 100 grams.

7.8 Other Considerations

7.8.1 Sample Compositing Area Versus Amount of Waste Toxic

The amount of EP toxic material that can be statistically determined will depend on the size of the sample compositing areas. In this study samples were collected from an area of one square foot. According to tolerance limits, the amount of EP toxic material is given by equation 73. (7.3)

$$R_{lower} = X - K*s$$

where:

R_{lower} = lower tolerance limit

X = average level of contaminant

K = value from tolerance limit table

s = standard deviation

The K factor from the table can be compared to:

$$(RTH - X)/s.$$

where

RTH = regulatory threshold

This value can be compared to K values listed in tables of "One-Sided Tolerance Limits for Normal Distributions" to obtain confidence that a certain percentage of the population will be greater than the regulatory threshold (Natrella, 1963).

Results of this study indicated at a 95 percent confidence level that at least 75 percent of the waste material was above the regulatory threshold based one foot sample areas. If samples were composited over the entire area, the portion of waste material that was above the regulatory threshold would be greater than 75 percent at the

95 percent confidence level. If samples were composited over an area less than one foot squares, the amount of material greater than the regulatory threshold would have been less than 75 percent at the 95 percent confidence level.

This variation was directly related to the varying standard deviations at the different sample sizes. The problem of the amount of material above a specified threshold, depending on the size of the sample, has been addressed quite extensively by mining industry researchers in the form of grade-tonnage curves (Clark, 1979).

7.8.2 Compositing Area Versus Long Range Error Reduction

If samples were composited from the entire target population, the Long Range Error would become insignificant. This reduction in total error would affect the study in two ways. First, fewer samples need to be collected for the same accuracy and precision as there would be less variation. Second, a larger amount of the waste material would be above the regulatory threshold from a statistical point of view (tolerance limits).

There are, however, some associated problems. If each sample were composited over the entire target population, the Long Range Error would be masked. Masking might not allow quantification of a section of hazardous material that exists within the bounds of the target population. For example, if one of the railroad cars in this study contained hazardous material and the other did not, the masking might show the entire population to be non-hazardous. This would lead the analyst to the improper conclusion that there was no hazardous material contained in the railroad cars due to dilution. Even though compositing over the entire target population area might be much less work, it should not be done (Garfield, 1984). Compositing might be necessary to reduce error if stratified sampling is impossible, but only to the point where the total error is reduced to a manageable level. Compositing beyond this level compromises the quality of the information obtained.

7.8.3 Problems with Large Fundamental Error

If the Fundamental Error is greater than 17 percent, the sampled distribution is deviating from normal. This large deviation is artificial and not a property of the material, which will lead to incorrect conclusions. The use of any statistical tests which depend on a normally distributed population will give false conclusions. The statistical literature affords detailed discussion of this matter (Koch and Link, 1980). In order to meet the underlying random distribution assumption of the statistical tests presented in this chapter the Fundamental Error must be less than 17 percent or classical statistical tests are not reliable. In addition to the lack of statistical uncertainty, the mode will be less than the mean. This results in an underestimation of the average level of contamination. This is not only important for characterization of a waste, but for the cleanup of a site. If there is an underestimation of the level of contamination, a site might be considered clean, when in fact it is not.

Chapter 8

CONCLUSIONS AND RECOMMENDATIONS

8.1 Conclusions

The purpose of an environmental investigation is to determine if the target population possesses characteristics that are hazardous to the environment. It is impractical and impossible to collect the entire target population for analysis, and therefore some type of statistical sampling strategy must be employed. From the results of this statistical sampling, certain inferences can be made regarding the target population. Since not all the target population is not analyzed, it is impossible to guarantee with 100 percent confidence that the statistical inference is correct.

However, any desired confidence level up to but not including 100 percent can be achieved through the application of sampling theory. The more accuracy is desired, the more resources must be utilized to achieve this. It is important to measure error before the investigation begins so that vital resources are not wasted. In order to take advantage of the power of sampling theory, certain characteristics about the waste material must be determined prior to sampling.

When determining the characteristics of the waste material, certain environmental limitations must be considered. These include limited site access and changing characteristics of the waste material over time. Given these limitations and the need of waste material characterization, the only viable types of analysis are those that can be performed immediately prior to sample collection. The only alternatives are portable types of field analysis such as X-ray fluorescence.

The current method of environmental sampling is inadequate because errors cannot be measured prior to investigation. If the errors are not measured, there is no guarantee that the desired goals can be met and the effort may be futile.

The current method of environmental sampling, as it is practiced at present, is also wrong because the sample mass is never considered. This inadequate sample mass is the source of many undetectable biases that occur in environmental investigations.

The combination of Pierre Gy's sampling theory and field analytical techniques will allow the environmental investigator to determine the total mass of sample necessary to achieve a particular goal. In addition to sample mass, two other considerations are also vital in environmental investigation. These are sampling strategy and the number of samples necessary to collect. Normally the latter considerations are the only ones the environmental investigator is concerned with. If this is the case and too small of a sample mass is collected, the distribution of results are positively skewed. If the distrubution is not normal, the classical statistical test normally used will provide erroneous information and will usually lead to an underestimation of the contamination level (Lichtenberg et al, 1988). A serious problem results from this underestimation. The statistical analysis would indicate that contaminated material is uncontaminated. A futher example of this problem could be in the evaluation of a site after cleanup; it is possible that sampling might suggest the site is clean when actually waste material levels are above cleanup specifications.

The only solutions to these inadequacies are the utilization of field analytical techniques, as discussed above, and sampling theory. Other benefits of field analysis are that the existence of a potential violation can be determined and characterization of waste materials into similar groups for purposes of stratified sampling.

Unfortunately many resources are expended in the laboratory and in case investigation, including months of labor and millions of dollars. Yet this effort is based on sampling accomplished with buckets and shovels in one afternoon. There is a need to shift more attention and resources toward the latter and bring them in line with technological and theoretical advances in investigation. An equalization of effort could result in improvement in environmental investigation as well as major savings of resources.

8.2 Recommendations

Pierre Gy's sampling theory and field analytical techniques must be employed if environmental investigations are to yield the desired information. No other approach will provide valid conclusions with the flexibility necessary for environmental applications.

Due to the sophisticated and integrated approach that is suggested with the employment of Gy's theory and field analytical techniques, a well trained professional should ideally be in charge of the environmental investigation. The most likely individual to be in charge of the environmental investigation is the analytical chemist. Although this Probably this is not a widely accepted suggestion, and may seem anathema to practitioners at certain functionals levels in the field, there are several reasons for the adoption of this proposition.

First, because the chemist performs sub-sampling in the laboratory, he/she should be aware of sampling theory (admitedly most chemists are not aware of sampling theory, but educating chemists in sampling theory is not difficult) and its practical application.

Second, the chemist is more aware of the necessary sample mass required to perform the analytical measurements.

Third, the chemist is more aware of appropriate holding times, proper containers, contamination sources, and preservations techniques necessary for eventual analysis of the contaminant.

Fourth, the chemist is more astute in recognizing potential chemical hazards that may exist in the field or in transporting samples back to the laboratory.

Fifth, although the chemist needs to work closely with the sampling crew because of its knowledge of tools available and their proper use, he/she should remain in charge of the process because of the sophisticated knowledge necessary for avoidance of biases and error associated with various sampling techniques.

Future developments necessary for improved environmental investigation are the following.

First, cooperation between all parties involved in any environmental investigation. This is the most valuable aspect in improving the success of environmental investigation.

Second, the need to develop a field technique that will accurately reflect the Extraction Procedure in order to more effectively determine its sampling constant.

Third, techniques need to be developed for using the same type of sampling theory and field analytical techniques for organic contaminants as used for inorganic contaminants.

Fourth, development of other technology besides portable X-ray to be used in conjunction with sampling theory that would provide a better limit of detection and sensitivity for various contaminants.

The potential for future development in the field of environmental investigation is great if Pierre Gy's sampling theory and field analytical techniques are employed. This study has discussed some aspects of this potential and some recommendations have been made.

REFERENCES CITED

Clark, I. 1979. Practical geostatistics. London: Elsevier

Clifton, H. E., Hunter, R. E., Swanson, F. J., and Phillips, R. L. 1969. Sample size and meaningful gold analysis. U. S. Geological Survey. Professional Paper 625-C. Government Printing Office.

Cochran, W. G. 1963. Sampling techniques. 2D ed. New York: John Wiley.

Code of Federal Regulations, Protection of the Environment, Title 40, Part 261, Office of the Federal Register, Washingtonm D.C.

Conover, W. J. 1980. Practical Nonparametric Statistics. Wiley: New York.

David, M. 1982. Geostatisitcal ore reserve estimation. Elsevier: Amsterdam.

Davis, J. C. 1986. Statistics and data analysis in geology. 2D ed. New York: John Wiley.

Garfield, F. M. 1984. Quality assurance principles for analytical laboratories. Association of Official Analytical Chemists: Virginia.

Ghosh, B. K. 1970. Sequential tests of statistical hypothesis. Addison-Wesley: Massachusetts.

Gy, P. M. 1982. Sampling of particulate material. Elsevier: Amsterdam.

Ingamels, C. O., and Pitard, F. F. 1986. Applied geochemical analysis. John Wiley: New York.

Kieth, L. H., ed. 1988. Principles of environmental sampling. American Chemical Society.

Koch, G. S., JR., and R. F. Link. 1980. Statistical analysis of geologic data. Dover Publications: New York.

Kratochvil, B. and Taylor, J. K. 1981. Sampling of chemical analysis. Analytical Chemistry. 53: (6) 924-938.

Lichtenberg, J. J., Winter, J. A., Weber, C. I., and Fradkin, L., eds. 1988. Chemical and biological characterization of municipal sludges, sediments, dredge spoils, and drilling muds. American Society for Testing and Materials: Philadelphia.

Merks, J. W. 1985. Sampling and weighing of bulk materials. Trans Tech: Germany

Natrella, M. G. 1963. Experimental statistics. Government Printing Office.

Pitard, F. 1988. Sampling methodologies for monitoring the environment: theory and practice. P. Gy and F. Pitard Sampling Consultants: Colorado.

Walpole, R. E., and Myers, R. A. 1985. Probability and statistics for scientists and engineers. Macmillan: New York.

Weise, N. L., ed. 1985. SME mineral processing handbook. Society of Mining Engineers of the American Institute of Mining, Metallurgical, and Petroleun Engineers: New York.