Available online at www.sciencedirect.com

SCIENCE DIRECT®

Analytica Chimica Acta 490 (2003) 159–168

ANALYTICA
CHIMICA
ACTA

www.elsevier.com/locate/aca

Gy sampling theory in environmental studies
2. Subsampling error estimates

Robert W. Gerlach[a,∗], John M. Nocerino[b], Charles A. Ramsey[c], Brad C. Venner[d]

[a] Lockheed Martin Environmental Services, 1050 E. Flamingo Road, Suite E120, Las Vegas, NV 89119-7427, USA
[b] National Exposure Research Laboratory, United States Environmental Protection Agency,
P.O. Box 93478, Las Vegas, NV 89193-3478, USA
[c] EnviroStat Inc., P.O. Box 636, Fort Collins, CO 80522, USA
[d] National Enforcement Investigations Center, United States Environmental Protection Agency, Denver Federal Center,
Building 53, P.O. Box 25227, Denver, CO 80225, USA

Received 7 October 2002; received in revised form 8 May 2003; accepted 12 May 2003

Abstract

Sampling can be a significant source of error in the measurement process. The characterization and cleanup of hazardous waste sites require data that meet site-specific levels of acceptable quality if scientifically supportable decisions are to be made. In support of this effort, the US Environmental Protection Agency (EPA) is investigating methods that relate sample characteristics to analytical performance. Predicted uncertainty levels allow appropriate study design decisions to be made, facilitating more timely and less expensive evaluations. Gy sampling theory can predict a significant fraction of sampling error when certain conditions are met. We report on several controlled studies of subsampling procedures to evaluate the utility of Gy sampling theory applied to laboratory subsampling practices. Several sample types were studied and both analyte and non-analyte containing particles were shown to play important roles affecting the measured uncertainty.

Gy sampling theory was useful in predicting minimum uncertainty levels provided the theoretical assumptions were met. Predicted fundamental errors ranged from 46 to 68% of the total measurement variability. The study results also showed sectorial splitting outperformed incremental sampling for simple model systems and suggested that sectorial splitters divide each size fraction independently. Under the limited conditions tested in this study, incremental sampling with a spatula produced biased results when sampling particulate matrices with grain sizes about 1 mm.

Keywords: Particulates; Sampling; Subsampling; Representative; Heterogeneous; Fundamental error; Pierre Gy

---

1. Introduction

The act of sampling may introduce more uncertainty than all of the subsequent steps in the measurement process [1–3]. This is especially true for heterogeneous particulate samples. The goal of most sampling procedures is to obtain a representative sample. In terms of characterizing the level of contaminants, the ideal representative sample would have the same concentration of contaminants as the original sample. Procedures for obtaining representative samples from particulate matrices have been studied for some time in the minerals industry and a comprehensive sampling theory originated by Pierre Gy has been developed

---

[∗] Corresponding author. Tel.: +1-702-897-3293;
fax: +1-702-897-6640.
E-mail address: rgerlach@lmepo.com (R.W. Gerlach).

0003-2670/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0003-2670(03)00568-3

[4–7]. Gy sampling theory was developed as a general tool for any particulate matrix, but has been principally used for mineralogical problems. The US Environmental Protection Agency (EPA) has an interest in determining how applicable Gy sampling theory is for environmental matrices [8,9]. Environmental matrices are very diverse and individual polluted sites may contain several distinct types of matrices. Hazardous waste sites may contain a heterogeneous matrix of particles from several sources, of which only a few may be contaminated. However, sometimes the contaminant may be present to some extent in all the particles. Some sites may contain several types of hazardous components present in a variety of concentration and particle distributions, requiring different sampling practices for different analytes [10,11]. The purpose of this work is to verify the applicability of, and experimentally confirm, Gy sampling theory methods for laboratory sample analysis practices.

Gy sampling theory [4,5] is usually presented as an accounting of all uncertainty components, and seven types of sampling error are traditionally identified [4,6,7] (Table 1). The Gy error components are the result of assuming that the goal of the study is to estimate the average amount of analyte over the entire lot of the target mass to be represented. The results reported here investigate errors that arise when subsampling for laboratory analysis. While more restrictive than site characterization, the sampling issues are very similar for both applications. However, special concerns related to site characterizations are not addressed here. For laboratory subsampling, no long-range trends or large-scale periodicity issues are assumed to be present.

Of the seven Gy error components, the fundamental error is the only subsampling error that can be estimated before analysis. It represents the uncertainty associated with randomly selecting particles from the sample and is related to the physical and chemical constitution heterogeneity between particles. For a well-constructed sampling program using "correct" sampling methods, the other error sources may be reduced such that the fundamental error may become the most significant contributor to the overall sampling uncertainty [4]. The fundamental error, which represents a lower bound for the measurement error, is often estimated as:

$$\sigma_{\mathrm{FE}}^2 = \left(\frac{1}{M_\mathrm{S}} - \frac{1}{M_\mathrm{L}}\right)\mathrm{IH}_\mathrm{L} = \left(\frac{1}{M_\mathrm{S}} - \frac{1}{M_\mathrm{L}}\right) fgcld^3, \tag{1}$$

where M_S is the sample mass, M_L the mass of the lot, IH_L the constant factor of constitution heterogeneity, f the shape factor, g the granulometric factor, c the mineralogical factor, l the liberation factor, and d is the largest particle diameter [4].

The mineralogical factor is maximum when the analyte is completely liberated, i.e. when the analyte consists of individual particles free of other materials. It is minimum when the analyte is uniformly distributed throughout the sample material. For samples with only two constituents, c can be estimated as [4]:

$$c = \lambda_\mathrm{M}\frac{(1-a_\mathrm{L})^2}{a_\mathrm{L}} + \lambda_\mathrm{g}(1-a_\mathrm{L}), \tag{2}$$

where λ_M is the density of the analyte particles, λ_g the density of the non-analyte material, and a_L the mass fraction (in units of decimal fraction) of the analyte.

The development of Eq. (1) assumes that all of the fragments are collected one-by-one at random. Eq. (1) also assumes that the average analyte concentration is associated with particles with an approximately normal particle size distribution. Other particles are assumed to have negligible analyte levels [4]. Based on

Table 1
Error components in Gy sampling theory

Error component	Feature/description
(1) Fundamental	Independent particle selection variability, statistical effect
(2) Grouping and segregation	Analyte heterogeneity, physical effect largely due to gravity
(3) Long-range heterogeneity	Concentration trends across a site or through time
(4) Large-scale periodicity	Periodic concentration changes across a site or through time
(5) Delimitation	Incorrect, non-representative, sample volume selected
(6) Extraction	Removing non-target sample components
(7) Preparation	Sample degradation, analyte loss/gain, contamination, etc.

Eq. (1) the analyst has two ways to reduce the fundamental error. One way is to increase the sample mass and the other is to crush the sample to reduce the particle diameter. Either option should be considered if the fundamental error estimate suggests that the study will not produce data with the accuracy and precision (representativeness) required to make a statistically supportable scientific conclusion.

Most samples are very large compared with the mass required for laboratory analysis. Continued efforts at lowering sample preparation costs, reducing waste generated from an analysis, and to provide results when small samples are all that are available have led to lower and lower mass requirements for a typical analysis. However, few studies show subsampling procedures that provide an accurate and precise, or representative, estimate of the original concentration. Previous studies [12–16] have shown that one of the best subsampling methods for particulate samples is a sectorial splitter. For demonstration purposes, experimental results in this study are limited to those from sectorial splitters or incremental sampling. All other methods, such as riffle splitting, fractional shoveling, and coning and quartering are expected to produce poorer uncertainty estimates.

Incorrect sampling may result in very uncertain values. Variability (bias and imprecision) depends on the mass of the subsample, the subsampling method, and the distribution of the analyte(s) within a sample. This behavior was verified experimentally by Starr et al. [17] who noted smaller sample masses were associated with "... smaller means, greater skewness, and higher variances ...". Lame and Defize [18] showed that the grouping and segregation error may overwhelm the fundamental error and require a minimum sample mass to obtain results with an acceptably low uncertainty. Results dependent on sample mass are also expected if the analyte particles are rare [4]. Walsh et al. [19] provide an example where performance was unaffected by an increased sample mass but was highly improved by a particle size reduction.

Gy sampling theory provides one mechanism to predict the uncertainty based on the sample properties and to adjust the sampling protocol if necessary. We report on three studies designed to test Gy sampling theory and representative sampling for environmental matrices. Each study was conducted with laboratory-prepared samples where the true concentration was known or predictable. Study one compares sectorial splitting sampling to increment sampling, with implications for all subsampling practices. Study two shows the effect of large particles on analytical variability, even when those particles have no analyte. The third study assesses Gy sampling theory predictions for mixtures of various particle sizes when the analyte is present in a non-traditional form, such as a thin film or coating. The matrix is non-traditional in the sense that typical controlled studies usually use matrices with discrete analyte particles.

2. Experimental

2.1. Study one, sectorial versus incremental sampling

Study one compares incremental sampling with sectorial splitting sampling. Samples consisted of 0.200 g coarse salt (Morton) ($\lambda_M = 2.165\,\mathrm{g\,cm^{-3}}$, $d = 0.05\,\mathrm{cm}$) and 39.8 g sand (ASTM C-778) ($\lambda_g = 2.65\,\mathrm{g\,cm^{-3}}$, $d = 0.06\,\mathrm{cm}$). Incremental sampling took place after first mixing the sand/salt mixture by tumbling it end-over-end for 60 s. Half of the sample was poured in a back and forth pattern across a 20 cm × 16 cm Pyrex pan. The pan was rotated 90° and the process was repeated with the rest of the sample. Eight 5 g subsamples were made by combining 10 increments taken with a stainless steel spatula using a stratified random sample selection. The sample was exhaustively subsampled. The sampler had access to the entire sample, a requirement of correct sampling design with respect to delimitation [4].

Sectorial splitting for Study one used a Model RR-4 Rotary Micro Riffler (QuantaChrome Corp., Syosset, NY) with a capacity to split a sample into eight fractions. The vibration of the feed hopper was adjusted to process a 40 g sample over several minutes. A slow feed rate reduces errors related to particle distribution heterogeneity and is preferable to high rotational rates for generating representative subsamples [20]. The remaining studies used a six-fold Model PT 100 sectorial splitter (Retsch, Hahn, Germany), but used the same feed system as described above.

Subsamples were added to distilled deionized water and the salt content was determined by conductivity analysis. Conductivity was measured with a Radiometer C.M. 83 Conductivity Meter with a type CDC 304T probe (Radiometer, Copenhagen). Standards were prepared from an NEIC primary KCl stock solution and a Radiometer conductivity standard (Radiometer Analytical, S.A., Villeurbane, France) of 0.01 D KCl. Concentrations were converted to [NaCl] by dividing with 0.84 [21], resulting in a maximum bias of 0.5% due to non-linear effects. Bias adjustments, based on quality assurance standards run periodically over the experimental period, were also done, though they were not significant with respect to the conclusions. Replicate analyses were run on randomly selected samples. Precision ranged from 1.3% relative standard deviation (R.S.D.) at low salt concentrations (5×10^{-4} M) to 0.7% R.S.D. at higher salt concentrations (5×10^{-3} M). The precision of the conductivity results was comparable to the uncertainty reported in prior studies [21].

2.2. Study two, the effect of large particles

The effect of coarse inert materials on subsampling uncertainty was evaluated in this study. The coarse component consisted of sandstone particles with maximum diameters of 0.5 cm. Twelve grams (72 particles) of coarse sandstone were combined with 23 g of sand and 1 g of table salt (Morton). The 36 g sample was split six-fold with a sectorial splitter and a randomly selected subsample was split to generate six 1 g subsamples. The subsample mass was determined with and without the coarse sandstone.

2.3. Study three, sampling mixtures

The main objective of this study was to evaluate subsampling uncertainty related to particle size. A second objective was to assess the performance when the analyte was present as a surface coating instead of as separate particles. A particle matrix with analyte coating the surface was prepared as follows. Quartz stones were crushed and sieved into three sizes; 0.018–0.071, 0.071–0.2, and 0.2–0.4 cm, labeled small, medium, and large, respectively. For each size fraction, approximately 50 g of material was placed in a separate filtration flask and covered with 0.34 M NaCl for approximately 5 min. Excess solution was aspirated from each filter flask, the coated particulates were dried overnight, and the dried mass was gently crushed to disaggregate the particles cemented together as the salt crystallized. Twelve grams of each size fraction was split into 2 g subsamples with the six-fold sectorial splitter. These subsamples were analyzed to obtain mean and variability estimates for each size fraction. A 6 g composite sample was prepared by combining 2 g of material from each size fraction, and the composite sample was split with the sectorial splitter to produce six 1 g subsamples.

3. Results and discussion

3.1. Study one, sectorial versus incremental sampling

Estimates for the total amount of salt in the original sample were determined for each subsample. For incremental sampling, Fig. 1a shows the measured percent bias from each individual subsample and the percent bias for the cumulative estimate of the mean based on all of the subsamples up to that point. The first six incremental subsamples are biased low and the last two are biased high. Subsample eight is biased high enough to qualify as an outlier ($P < 0.01$) based on Dixon's outlier test [22]. The cumulative bias and relative standard deviation are remarkably consistent for runs two through seven, and the cumulative bias remains at least 16% low until the result from subsample eight is included. The mean bias with all eight results is +1.9%. In contrast, the sectorial splitter produced five high-biased and three low-biased subsamples and the cumulative bias was always less than 5% except at run two (see Fig. 1b). The mean bias after exhaustive analysis was 3.2% for the sectorial splitter results.

Incremental sampling is affected by several error components of Gy sampling theory (Table 1), including: the fundamental error, the grouping and segregation error, the extraction error, and the delimitation error [4]. The fundamental error is related to the heterogeneous nature of the particle mixture and is independent of the sampling method. The grouping and segregation errors are related to the non-uniform distribution of analyte, often resulting from gravitational effects. Only 10 increments were used while more than 30 increments have been recommended [4]. The extraction error is the difference between the planned, or delineated, portions to be sampled and the portion

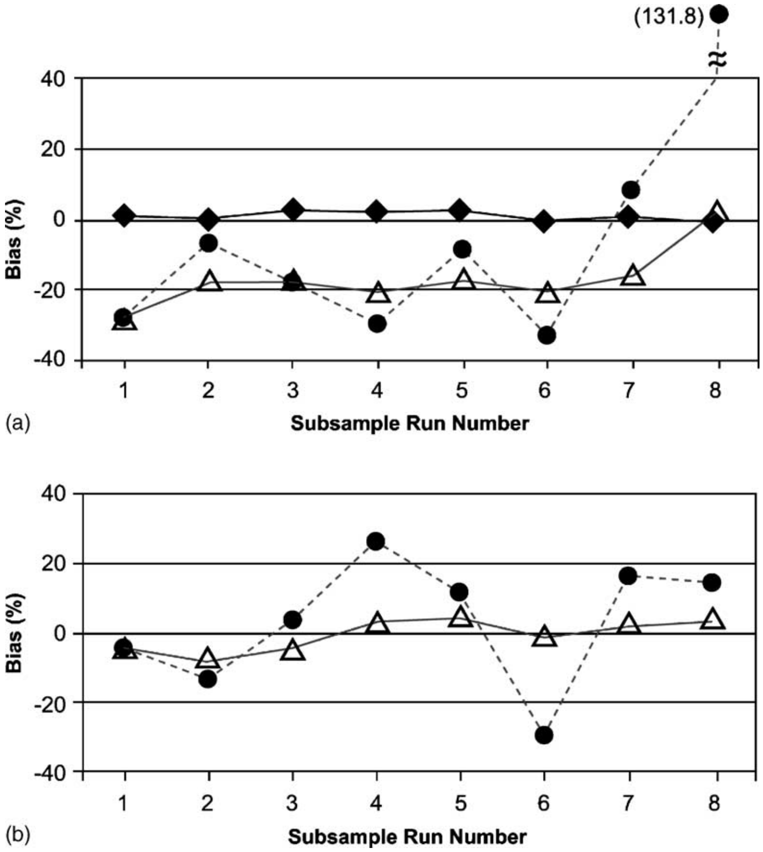

Fig. 1. Bias for individual subsample analysis (●) and for the cumulative estimate of average analyte level (△) as a function of the subsample run number for (a) incremental and (b) sectorial splitter sampling in Study one. Bias for single increment subsamples taken after ring-and-puck grinding to a fine powder ((a) only) (◆).

taken. The magnitude of this error is unknown. The delimitation error is also unknown but is expected to be small since the particle sizes are small compared with the spatula and the area involved in each increment. The corresponding errors with sectorial splitting are expected to be smaller than or equal to the errors for the incremental method.

The fundamental error for an individual sample analysis predicted using Eqs. (1) and (2) with $f = 0.75$ (very cubic shaped), $g = 0.55$ (a screened sample), $a_L = 0.005$, $l = 1.0$ (complete liberation of analyte from the matrix), and $d = 0.06\,\mathrm{cm}$ is $\sigma_{FE} = 8.2\%$. This represents 46% of the 17.7% sample analysis error found experimentally from the population of individual sectorial splitter results. Eq. (1) required several assumptions and is considered to be an approximate result, with the total experimental error accounting for additional uncertainty from other sources.

Significant operational factors differentiate the sectorial and incremental results. The sectorial splitter subsamples are generated simultaneously, are statistically equivalent, and could have been analyzed in any order. The incremental sampling results depend on the subsample run order. Any bias or error associated with one subsample may affect the results for subsequent subsamples. In this example, all early incremental subsamples had large negative biases and the result from the last incremental subsample was crucial in generating a summary result with a low bias.

If incremental sampling as described in this study was used for routine analysis, then at most a few subsamples would be taken and any reported concentration would have a large negative bias (Fig. 1a). For the sectorial splitter method, only a few subsamples had a large positive or negative bias (Fig. 1b). If only a few randomly chosen sectorial splitter subsamples were analyzed, then the expected bias would be much smaller than the bias in the incremental sampling method. For sample types similar to those used in this study, sectorial splitting is expected to outperform incremental sampling.

It is also of interest to compare the performance from the above methods to the results from the analysis of one additional sample. The additional sample consisted of 100 g of the same materials used above and in the same proportion. This sample was ground to a fine powder in a ring-and-puck mill before taking individual 0.5 g grab samples from random locations. The relative biases for eight of these samples are plotted in Fig. 1a. The cumulative (mean) bias for these samples was 1.1% with a coefficient of variation of 1.4%. Despite having 1/10th the mass of the Study one subsamples and using no incremental fractions, the particle size reduction from grinding lowered the sampling error well below the error found with the original matrix by either sampling method. This is in agreement with Gy sampling theory, where a change in particle diameter is expected to have a large effect on the fundamental error (Eq. (1)). Using Eq. (1) the diameter for the ground powder is predicted to be a factor of 12 smaller than the initial particle size. This is consistent with the silt-like visual and tactile characteristics of the powder [23]. However, there is a cost associated with this procedure. It consumed the entire sample and required a lengthy cleanup effort, substantially increasing the price per sample.

3.2. Study two, the effect of large particles

The large particles in each nominal 1 g subsample were counted. The frequency distribution of the large particles agreed with a Poisson distribution based on a χ^2-test [24] where all of the subsamples with $n \geq 4$ large particles were pooled ($\chi^2 = 2.54$; critical $\chi^2_{0.90,3} = 6.25$), indicating that the sectorial splitters distribute all of the particles independently. The distribution of the fine components appeared visually uniform and independent of the distribution of the coarse particles.

If one ignores the presence of the large particles, then the experimental R.S.D. for NaCl in the salt/sand portion of the samples is 17.2% (Table 2). The uncertainty may be primarily attributable to the distribution of salt rather than sand. The relative error is only slightly different from the error for NaCl and the error for sand is comparatively small. When the mass of the large particles is included, the measured R.S.D. for NaCl is 35.1%. If the large sandstone particles were converted to sand particles before sectorial splitting, the predicted R.S.D. for NaCl in sand is 11%. This value is based on the assumption that the sand particles created from the large particles were distributed with the same variability associated with the original sand component. This result is a more appropriate value to compare with the large particle R.S.D. because the subsample masses are equal for each case. Thus, the increase in uncertainty associated with larger inert particles is at least a factor of 2 and more likely to be larger than a factor of 3.

Gy sampling theory can be applied to the subsampling of just the sand and salt if the particle distributions are assumed to be independent of particle

Table 2
Effect of large inert particle on subsampling variance in Study two

Subsample	Sand (g)	Large particles (g)	Total inert (g)	NaCl (g)	NaCl in sand (g/g)	NaCl in total (g/g)
LP-21	0.64	0.73	1.37	0.0335	0.0500	0.0239
LP-22	0.61	0.11	0.72	0.0243	0.0385	0.0328
LP-23	0.65	0.55	1.20	0.0233	0.0348	0.0191
LP-24	0.63	0.11	0.74	0.0356	0.0531	0.0457
LP-25	0.68	0.00	0.68	0.0347	0.0488	0.0488
LP-26	0.56	0.45	1.01	0.0320	0.0542	0.0307
Mean	0.626	0.325	0.951	0.031	0.047	0.034
R.S.D.[a] (%)	6.4	90.1	30.1	17.6	17.2	35.1

[a] R.S.D.: relative standard deviation.

Table 3
Fundamental error estimates for Study two

Parameter[a]	Salt and sand		Large particles added	
	Step 1	Step 2	Step 1	Step 2
M_S (g)	4	0.667	6	1
M_L (g)	24	4	36	6
f	0.75	0.75	0.5	0.5
g	0.55	0.55	0.55	0.55
λ_M (g cm^{-3})	2.165	2.165	2.165	2.165
λ_g (g cm^{-3})	2.65	2.65	2.65	2.65
a_L	0.0416	0.0416	0.0278	0.0278
$c = \lambda/a_L$	50.3	50.3	76.2	76.2
l	1	1	1	1
d (cm)	0.06	0.06	0.5	0.5
σ_{FE} (%)	3.1	7.5	60.0	148
Total σ_{FE} (%)[b]	8.1		160	

[a] Parameter definitions—M_S: mass of the sample (the subsample); M_L: mass of the lot (the sample); f: shape factor; g: granulometric factor; λ_M: analyte particle density; λ_g: non-analyte particle density; a_L: concentration as decimal proportion; c: mineralogical factor; l: liberation factor; and d: largest particle diameter.

[b] Fundamental error (as R.S.D.) due to both Steps 1 and 2.

size. Since the mass of the large particles is known, it can be subtracted and Eq. (1) can be applied to the two-step process summarized in Table 3. Columns 2 and 3 of Table 3 show the parameter estimates and resulting fundamental error predictions for the sand/salt fraction alone. The fundamental error predicted for the two-phase salt/sand subsampling protocol is 8.1%. This is 47% of the measured value of 17.2%. The fundamental error is very similar to the 17.7% results from Study one.

Eq. (1) was also applied to the complete subsample, including the large particles, with parameter settings and results shown in the last two columns of Table 3. The fundamental error for the two-phase subsampling protocol is 160%, much larger than the measured value of 35%. Eq. (1) provides a poor estimate when applied to the complete subsample mainly due to the sample characteristics not conforming to the assumptions. Eq. (1) was developed with the assumption that the diameter is representative of particle sizes from a compact distribution. The large particles represent 33% of the sample weight while the rest of the sample consists of particles that are much smaller. The shape factor change from 0.75 to 0.5 reflects the more sphere-like nature of the sandstone particles compared to the rather cubic shape of the sand particles. However, the sample particle diameters do not form a continuous distribution, but a bimodal discrete distribution. Since the diameter, d, in Eq. (1) is identified with an upper bound for a normal distribution of particle sizes, poor predictions are not surprising here.

3.3. Study three, sampling mixtures

Study three involves subsampling a mixture composed of three size fractions where each size fraction contributes an equal mass. Evaluation of the mean and standard deviation for each particle size fraction reveals an interesting result. Despite having the least amount of NaCl, the large particle fraction had the largest standard deviation (Table 4). The large fragment fraction has more than five times the error in absolute concentration than the small fragment fraction has, though it contains only about one-third as much NaCl. A simple statistical error estimate based on the variability for the individual size fractions gives an 8.3% error for the mixture sample results compared with the measured 12% value.

The above results are in overall agreement with Gy sampling theory, which characterizes sample particle size effects by using a diameter representative of the largest particles. Larger particle sizes are associated with larger uncertainties. As shown above, this is the case not only with particles associated with the analyte of interest, but also with the analyte-free particles.

Table 4
Measured composition and uncertainty for [NaCl] in crushed quartz splits in Study three compared to the uncertainty for the fundamental error predicted from Gy sampling theory

Size fraction	Size range (cm)	Mean [NaCl] (g/g)	S.D.[a] [NaCl] (g/g)	Measured error[b] (%)	Propagated error[b,c] (%)	Gy theory fundamental error[b,d] (%)
Large[e]	0.2–0.4	0.00107	0.000257	24.0	–	–
Medium[e]	0.071–0.2	0.00247	0.000161	6.5	–	–
Small[e]	0.018–0.071	0.00324	0.000046	1.4	–	–
1:1:1 Mixture[f]	0.018–0.4	0.00263	0.000320	12.0	8.3	8.2

[a] S.D.: standard deviation.
[b] Error as relative standard deviation.
[c] Mixture error predicted from size fraction errors.
[d] Fundamental error.
[e] Final subsample mass of 2 g.
[f] Final subsample mass of 1 g.

The fundamental error for the mixture sample could not be estimated using the formula in the right hand side of Eq. (1), as the sample did not contain a set of particles enriched in analyte distinct from the rest of the particles. For this sample matrix, an alternate approach based on contributions from individual particles as described in [25] was used. In this approach, the fundamental error equation can be written as:

$$s_{\mathrm{FE}}^2 = \left(\frac{1}{M_{\mathrm{S}}} - \frac{1}{M_{\mathrm{L}}}\right) \sum_i \frac{(a_i - a_{\mathrm{L}})^2}{a_{\mathrm{L}}^2} \frac{M_i^2}{M_{\mathrm{L}}}, \tag{3}$$

where the summation is over each particle, M_i is the mass for the ith particle, M_{S} the mass of the sample, M_{L} the mass of the lot, and "a" values are the analyte mass fractions.

To use Eq. (3) the summation is broken into three parts, one for each size fraction, with a_i and M_i based on approximations. The approximations are that the values of a_i and M_i are constant for each size fraction. The values for a_i are the average analyte level found for each size fraction, but scaled slightly (see Table 5) so the calculated average analyte level for the mixture matched the measured analyte level. This correction removes variance contributions due solely to bias between initial estimates. The value of M_i is the estimated average mass for a particle with average particle size for the size fraction associated with the ith particle. Average mass values were estimated using a shape factor of 1 for this matrix, allowing particle volumes to be calculated as d^3. The shape factor is higher than one might expect because the particles were observed to have generally elongated shapes rather than being generally symmetrical.

The error contribution by size fraction (Table 5) again demonstrate the importance of the larger size fraction in determining the sampling error. The constant factor of constitution heterogeneity contribution

Table 5
Size fraction components of the fundamental error calculations for Study three

Size fraction	Average dimension (cm)	Particle number[a] (N)	Adjusted a[b] (weight fraction)	Average particle mass (g)	$\mathrm{IH_L}$[c]
Large	0.3	28	0.00125	0.0714	0.00655
Medium	0.135	307	0.00288	0.00651	0.00002
Small	0.0445	8564	0.00377	0.000234	0.00147
			Average = 0.00263		Total = 0.00804

[a] Number of particles estimated with shape factor = 1 ($v_i = d_i^3$).
[b] Weight fraction sample concentrations (as a mass fraction) scaled so the average calculated mixture concentration matches the average measured mixture concentration.
[c] Constant factor of constitution heterogeneity.

for the large size fraction is 0.00655. This is substantially larger than the corresponding values for the medium and small size fractions (last column, Table 5). The overall fundamental error estimated with this approach is 8.2%, which is 68% of the measured error. This estimate would provide useful information if available at the planning stage of a study, or when determining a subsampling protocol.

4. Conclusions

Studies that use particulate samples must consider the subsampling process as an integral part of the method. Subsampling may substantially affect the variability and sampling practices can generate biased values despite the use of accurate chemical analyses. These studies show the principal error is due to sampling rather than chemical analysis procedures.

The comparison of incremental sampling with sectorial splitting sampling showed that incremental sampling results could be biased unless exhaustive analysis is carried out. Sectorial splitters outperformed incremental sampling when applied to very simple sample matrices. The incremental sampling procedure did not meet the typical benchmark of 30 increments to minimize grouping and segregation error effects. However, the results from exhaustive analysis suggest an extraction error bias would have been unaffected by the use of additional increments—as the sampling device failed to select 100% of the target when each increment is sampled. The sampling device was flat but without vertical sides, and is classified as an incorrectly delimited sampling device. While technically true, the sample depth was so shallow that parallel sides were expected to make little difference in the outcome. The sampling device was probably also incorrect with respect to extraction [4] because it was unable to remove all the fine particles associated with each increment.

Several aspects of Gy sampling theory were experimentally confirmed with laboratory generated test matrices containing discrete or large particles. Estimates of the fundamental error were consistent with the measured variability for the samples with discrete analyte particles, and represented about 47% of the total error (Study two). The presence of large, analyte-free particles was shown to play as important a role as the analyte in contributing to the measurement variability. Though finely crushing a sample resulted in dramatically reduced subsampling variability, the additional effort required for cleanup between samples is very large, making this an unattractive alternative.

Gy sampling theory also successfully provided estimates of subsampling uncertainty when applied to particulate matrices where the analyte was distributed across all of the particles. The fundamental error estimated for Study three represented 68% of the total error. However, the calculation of the fundamental error for the surface-coated matrix sample relied on auxiliary information that would not normally be available. In addition, though the results for this matrix appear promising, the fundamental error calculation for Study three involves several approximations and should be verified using different combinations of size fractions and particulate characteristics before this method is generally applied.

Acknowledgements

The authors thank K. Eric Nottingham, US Environmental Protection Agency (EPA), NEIC, Denver, for facilitating these studies. The US Environmental Protection Agency through its Office of Research and Development funded and managed the research described here under GSA Contract Number GS-35F-4863G to Lockheed Martin Environmental Services. It has been subjected to the Agency's review and approved for publication. Mention of trade names or commercial products does not constitute an endorsement or recommendation for use.

References

[1] T.F. Jenkins, C.L. Grant, G.S. Brar, P.G. Thorne, P.W. Schumacher, T.A. Ranney, Field Anal. Chem. Technol. 1 (3) (1997) 151–163.
[2] D.M. Crumbling, Applying the Concept of Effective Data to Environmental Analyses for Contaminated Sites, EPA 542-R-01-013, US Environmental Protection Agency, Washington, DC, 2001.
[3] T. Allen, Particle Size Measurement, vol. 1, Chapman & Hall, London, 1996.
[4] F.F. Pitard, Pierre Gy's Sampling Theory and Sampling Practice, second ed., CRC Press, Boca Raton, FL, 1993.
[5] P. Gy, Sampling for Analytical Purposes, Wiley, New York, NY, 1998.

[6] J.C. Myers, Geostatistical Error Management, Van Nostrand Reinhold, New York, 1997.

[7] P.L. Smith, A Primer for Sampling Solids, Liquids, and Gases, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2001.

[8] D.M. Crumbling, Getting to the Bottom Line: Decision Quality vs. Data Quality, US Environmental Protection Agency, http://www.epa.gov/quality/qs-docs/21qa-papers.pdf (accessed August 2002).

[9] B.J. Mason, Preparation of Soil Sampling Protocols: Sampling Techniques and Strategies, EPA/600/R-92/128, US Environmental Protection Agency, Las Vegas, NV, 1992.

[10] E. Mishalanie, C. Ramsey, Nat. Resour. Environ. 13 (4) (1999) 522–527.

[11] C.A. Ramsey, J. Suggs, Environ. Test. Anal. 10 (2) (2001) 13–16.

[12] R.W. Gerlach, D.E. Dobb, G.A. Raab, J.M. Nocerino, J. Chemom. 16 (2002) 321–328.

[13] T. Allen, A.A. Khan, Chem. Eng. (London) 238 (1970) CE108–CE112.

[14] C.E. Mullins, B.J. Hutchison, J. Soil Sci. 33 (1982) 547–561.

[15] B.H. Kaye, Powder Metall. 9 (1962) 213–234.

[16] H.J. Scullion, J.E.C. Harris, Proc. Soc. Anal. Chem. 10 (5) (1973) 108–114.

[17] J.L. Starr, T.B. Parkin, J.J. Meisinger, Soil Sci. Soc. Am. J. 59 (1995) 713–719.

[18] F.P.J. Lame, P.R. Defize, Environ. Sci. Technol. 27 (1993) 2035–2044.

[19] M.E. Walsh, C.A. Ramsey, T.F. Jenkins, Chemosphere 49 (2002) 1267–1273.

[20] T.A. Hatton, Powder Technol. 19 (1978) 227–233.

[21] H.S. Harned, B.O. Owen, The Physical Chemistry of Electrolytic Solutions, third ed., Reinhold, New York, NY, 1958.

[22] W.J. Dixon, F.J. Massey Jr., Introduction to Statistical Analysis, third ed., McGraw-Hill, New York, 1969.

[23] Soil Survey Division Staff, Soil Survey Manual, Handbook Number 18, US Department of Agriculture, Washington, DC, 1993.

[24] J.H. Zar, Biostatistical Analysis, third ed., Prentice Hall, Upper Saddle River, NJ, 1996.

[25] P.M. Gy, Sampling of Particulate Materials, Theory and Practice, Elsevier, Amsterdam, 1982.