Sampling Error Associated with Collection and Analysis of Soil Samples at TNT-Contaminated Sites

T. F. Jenkins,[1] C. L. Grant,[1] G. S. Brar,[1] P. G. Thorne,[1] P. W. Schumacher,[1] T. A. Ranney[2]
[1] *U.S. Army Cold Regions Research and Engineering Laboratory, 72 Lyme Road, Hanover, NH 03755*
[2] *Science and Technology Corporation, Hanover, NH 03755*

Received 5 August 1996; revised 23 October 1996; accepted 24 October 1996

Abstract: This study assessed short-range spatial heterogeneity of TNT concentrations in surface soils at explosives-contaminated sites. Discrete and composite samples were analyzed by both on-site colorimetric techniques and standard laboratory protocols. Three locations were sampled at each of three installations, and the results were used to estimate the relative contributions of analytical error and sampling error. The major contaminant at seven of the nine sampling locations was TNT, and the on-site colorimetric method provided results that were in excellent agreement with laboratory results from the use of SW846 Method 8330. At four of the seven TNT locations, short-range concentration variations were modest and analyte distribution was sufficiently Gaussian to allow application of normal distribution statistics to fractionate the total error variances. For these four locations, standard deviations due to sampling were greater than the corresponding standard deviations due to analysis by factors ranging from 2.6 to 22.8. This relationship held whether characterization was done with the use of on-site analysis or laboratory analysis. For the other three TNT locations, enormous short-range spatial heterogeneity was encountered and sampling error overwhelmed analytical error. To improve estimates of mean concentrations, sampling error was reduced by the use of composite sampling strategies. Overall, this study indicates that characterization of explosives-contaminated sites with the use of a combination of composite sampling, in-field sample homogenization, and on-site colorimetric analysis is an efficient method of obtaining accurate and precise mean concentration estimates that are representative of the area. © 1997 John Wiley & Sons, Inc. Field Analyt Chem Technol 1: 151–163, 1997

Keywords: analytical error: heterogeneity/spatial variability; on-site analysis; sampling error; soil; TNT; composite sampling

Introduction

Accurate chemical characterization of a hazardous-waste site requires the development and implementation of a well-designed sampling plan. After defining the area of interest [target population(s)], which might be an entire site or several defined areas within a site, samples are collected according to one of several possible schemes. Distributions of contaminants are very site specific, depending on the manner in which the contamination occurred, the physical and chemical properties of the contaminants involved, soil type, and the geology and hydrogeology of the site. Because of these site-specific characteristics, many references recommend that one perform a preliminary study before devising a sampling plan.[1–5]

Explosives are solids at ambient temperature, dissolve slowly and sparingly in aqueous solution, and have low vapor pressures. These properties limit modes of mobility compared with other contaminants such as fuels or solvents. Thus areas of high concentrations remain at or near the surface where deposited, unless the soils themselves are moved.

Historically, most studies of hazardous waste sites have relied on shipping samples to off-site laboratories for analysis. Besides the high cost and potential for sample contamination or degradation of labile analytes, this arrangement does not lend itself to timely decisions that are necessary in a stepwise plan. Recently this problem has been addressed with the development and promotion of on-site analytical methods.[3,5,6–10] Inexpensive, on-site methods for the most common explosives in munitions-contaminated soils have been developed and are now in common use.[8] These procedures appear to be adequate for mapping locations of contamination and, if a sufficient number of samples are analyzed, they can provide estimates of spatial contaminant heterogeneity. Sequential modifications in sampling plans are also feasible because data become available while sampling is in progress.

On-site analytical methods are sometimes criticized as having inadequate precision, accuracy, and specificity. With respect to specificity, we agree that the QA/QC plan must include laboratory-based confirmatory measurements on selected samples. Accuracy should also be verified against reference methods for an appropriate number of samples. The precision issue, however, is a different matter. Historically, analytical precision estimates for methods used in

This publication reflects the personal views of the authors and does not suggest or reflect the policy, practices, programs, or doctrine of the U.S. Army or Government of the United States.

Correspondence to: T. F. Jenkins

hazardous waste characterization have received an inordinate amount of attention compared to sampling error. Contaminated soils are often extremely heterogeneous, which causes the major error source to be sampling and subsampling. No amount of improvement in analytical precision can significantly reduce total measurement error when the analytical error is a minor contributor to the total. Williams[5] noted that the newly released U.S. EPA DQO guidelines focus on the uncertainty of a specific decision rather than the individual parameters that contribute to the overall uncertainty. This is an encouraging change.

The optimization of a sampling plan can only occur after the process of obtaining representative samples has been adequately addressed. Numerous variations have been offered to describe the qualifications of representative samples.[1,6,11,12] We are partial to the Gilbert definition, "A representative unit is one selected for measurement from the target population in such a way that it, in combination with other representative units, will give an accurate picture of the phenomenon being studied.[1]" According to Barnard, "Representativeness is a statistical concept that is a measure of how well a data set of sample measurements yields information concerning the population.[6]"

In this study we focus on how to obtain representative samples from surface soils contaminated by munitions residues. Too often, local spatial heterogeneity is ignored in favor of grab sampling on the theory that heterogeneity will be averaged out if sufficient numbers of samples are taken. Although there is validity in this position, it hardly qualifies as cost-effective, especially when analysis cost often outpaces sample collection cost by orders of magnitude. Consistent with our experience, several authors have reported large local spatial heterogeneity, often of the same magnitude as present on a much larger scale.[4,13-15] To address this problem, others have used or recommended composite sampling.[1,2,16-21] Composite sampling is sometimes discouraged because it eliminates information regarding the variability of the individual samples composited. When applied to large areas, this limitation may represent a valid concern, especially when concentrations are near a regulatory limit. However, when used on localized areas in lieu of grab sampling, we believe it is an attractive option to improve representativeness of samples. We decided to investigate the feasibility of this approach coupled to both on-site analysis and conventional laboratory analysis. This was done by conducting sampling and analysis studies at a number of explosives-contaminated sites that varied in explosives analytes present, mode of contamination, soil type, and geohydrology.

Experimental Methods

Throughout this manuscript the term *installation* will refer to the government facility where sampling was conducted, *location* will refer to any one of the nine areas (three at each installation) sampled, and *sample position* (or *sample number*) will refer to the specific spatial position where a discrete sample was collected.

Sampling Sites

Sampling studies were conducted at (a) Monite, a Bureau of Land Management (BLM) installation near Sparks, Nevada; (b) Hawthorne Army Ammunition Plant (AAP), Hawthorne, Nevada; and (c) Volunteer AAP, Chattanooga, Tennessee.

The Monite installation is a small former industrial area with about 1.5 acres of land contaminated with TNT (2,4,6-trinitrotoluene) and DNT (2,4-dinitrotoluene). Explosives from out-of-date military munitions were reportedly reclaimed here, but because the site was abandoned many years ago, the history of contamination is largely unknown. We collected preliminary soil samples that were analyzed using the EnSys on-site colorimetric method.[7] This initial work revealed three locations that were selected for intensive sampling and analysis. One had TNT concentrations in the thousands of μg/g (location 1), one had similar levels of DNT (location 2), and a third had low μg/g levels of TNT (location 3).

Hawthorne AAP was established in 1928 and was operated for many years as a load, assemble, and pack facility for the Navy. In 1977 it was transferred to Army control. Here too, three locations were selected based on preliminary results. The first location was under a conveyor belt that took empty boxes and crates from the inside of a melt facility out for disposal. TNT was present with soil concentrations in the thousands of μg/g (location 4). The second location was at an open burning area that was free of vegetation and had concentrations of TNT in the hundreds of μg/g (location 5). The final location was a disposal lagoon, where the surface soils were visually contaminated with an intense yellow crystalline material that we believed to be ammonium picrate (location 6).

Volunteer AAP is a TNT and DNT production facility, although it has not actively produced these munition compounds since 1977. Here again we selected three sampling locations based on preliminary studies. The first location was at a loading area located adjacent to a TNT production building (location 7 and location 7R). This area was also contaminated by wash water from the facility, and concentrations of TNT in the soil were in the thousands of μg/g. Location 7R was offset from location 7 by 15 cm, so like-numbered samples from the two locations were all 15 cm apart. Sampling of location 7R was necessitated by a malfunction of the automatic pipette used to dispense extracting solvent when discrete samples from location 7 were analyzed with the use of the on-site colorimetric method. The second sampling location was within a drainage ditch that received spills of TNT production wastewater (location 8). Individual samples collected within the ditch had elevations that differed by only a maximum of 25 cm; however, TNT concentrations varied from 500 to 30,000 μg/g. The third

location was an area initially thought to be free of contamination, but preliminary analyses showed TNT concentrations in the 4–40 μg/g range (location 9).

Soil Sampling Procedure

At all nine locations a plastic template was placed on the ground with the center at the selected sampling location with sample numbers 2 and 5 oriented north–south (Figure 1). Seven samples were collected in a circular pattern (radius 61 cm), with sample number 1 in the center. Each sample was collected from 0 to 15 cm deep with the use of a manual 5.0-cm-diameter stainless steel hand auger. Vegetation, when present, was removed. Cores were transferred to plastic Zip-Lock™ bags and taken to a processing area. At the Monite site, processing was conducted outdoors in the shade, to minimize the possibility of photodegradation. At Hawthorne and Volunteer, soil processing was conducting in air-conditioned buildings.

On-Site Soil Processing

A summary of the entire sampling design is shown in Figure 2.

Discrete Samples. Soil samples from the Monite installation and Hawthorne AAP were dry and mostly consisted of a mixture of sands and gravels. Each sample was emptied from its Zip-Lock™ bag into a separate 23-cm-diameter aluminum pie pan. The soil was dispersed by breaking up large clumps with gloved hands; large rocks were removed. The pans were covered with a second pie pan and the soil swirled and shaken vigorously to disperse and homogenize the material, which was then coned and quartered. Subsamples of approximately 5 g were removed from each quarter and combined to produce a sample of about 20 g for colorimetric on-site analysis. The bulk sample was remixed, coned, and quartered again, and a duplicate 20-g sample for on-site analysis was removed, as described above. The sample was remixed a third time, and another 20-g sample was removed and placed in an amber 40-ml glass vial for subsequent laboratory analysis. The remaining sample was returned to its original Zip-Lock™ bag and saved for preparation of a composite sample for that sampling location.

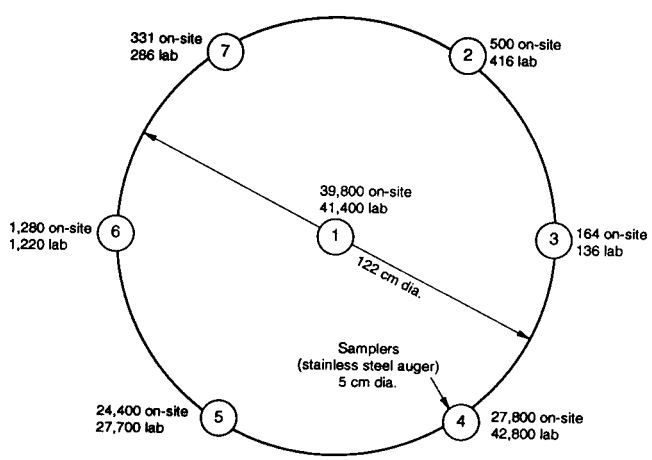

FIG. 1. Sampling pattern with analytical results (μg/g) for TNT for sampling location 1.

Soils from Volunteer had a higher moisture content and a higher percentage of fine-grained material than soils from either Monite or Hawthorne. This made field homogenization more difficult and time-consuming. At Volunteer, soil samples were placed in Zip-Lock™ bags and initially kneaded by hand to break up large clumps. Soil was then deposited in aluminum pie pans and further disaggregated by hand until approximately pea-sized or smaller pieces were produced. For soil from sampling location 7, rocks greater than 0.5 cm were removed and weighed. Soils were then coned and quartered and further processed as described above.

Composite Samples. For composite samples at Monite and Hawthorne AAP, the soil remaining after discrete samples were removed for each of the seven individual samples within a location was combined in a large aluminum roasting pan. Although the individual portions used to make the composite for Monite and Hawthorne were not individually weighed, they were approximately equal in weight. The soil was homogenized by hand mixing and then coned and quartered. Approximately 5-g sample aliquots were removed from each quarter and combined to produce a 20-g sample for on-site analysis. The soil was coned, quartered, and sampled six more times to produce a total of seven replicates for on-site analysis. The soil was dispersed, coned, and quartered one final time, and a 50-g sample was removed to an amber glass bottle for subsequent laboratory analysis.

At Volunteer, a similar procedure was used, except that equal weights of each individual sample (100 or 600 g each, depending on location) were used to prepare composites. Otherwise, samples were processed as above.

On-Site Colorimetric Analysis for TNT

The 20-g soil samples were extracted in 150-ml plastic bottles by adding 100 ml of acetone and shaking vigorously.[8] Extracts were analyzed with the use of the EnSys TNT method.[7] The acetone contained 3% water to ensure that adequate water was present for the chemical reaction that produces color development. An extraction-rate study was conducted on the soil from each installation. A 3-min extraction time was adequate for soils from the Monite site and Hawthorne AAP, but soils from Volunteer AAP were extracted by using 3 min of shaking, a 30-min rest time, and a second 3-min shaking period. After allowing the soil to settle for at least 15 min, an aliquot of each extract was removed with the use of a Plastipak syringe and filtered through a Millex SR (Millipore) membrane. Extracts were

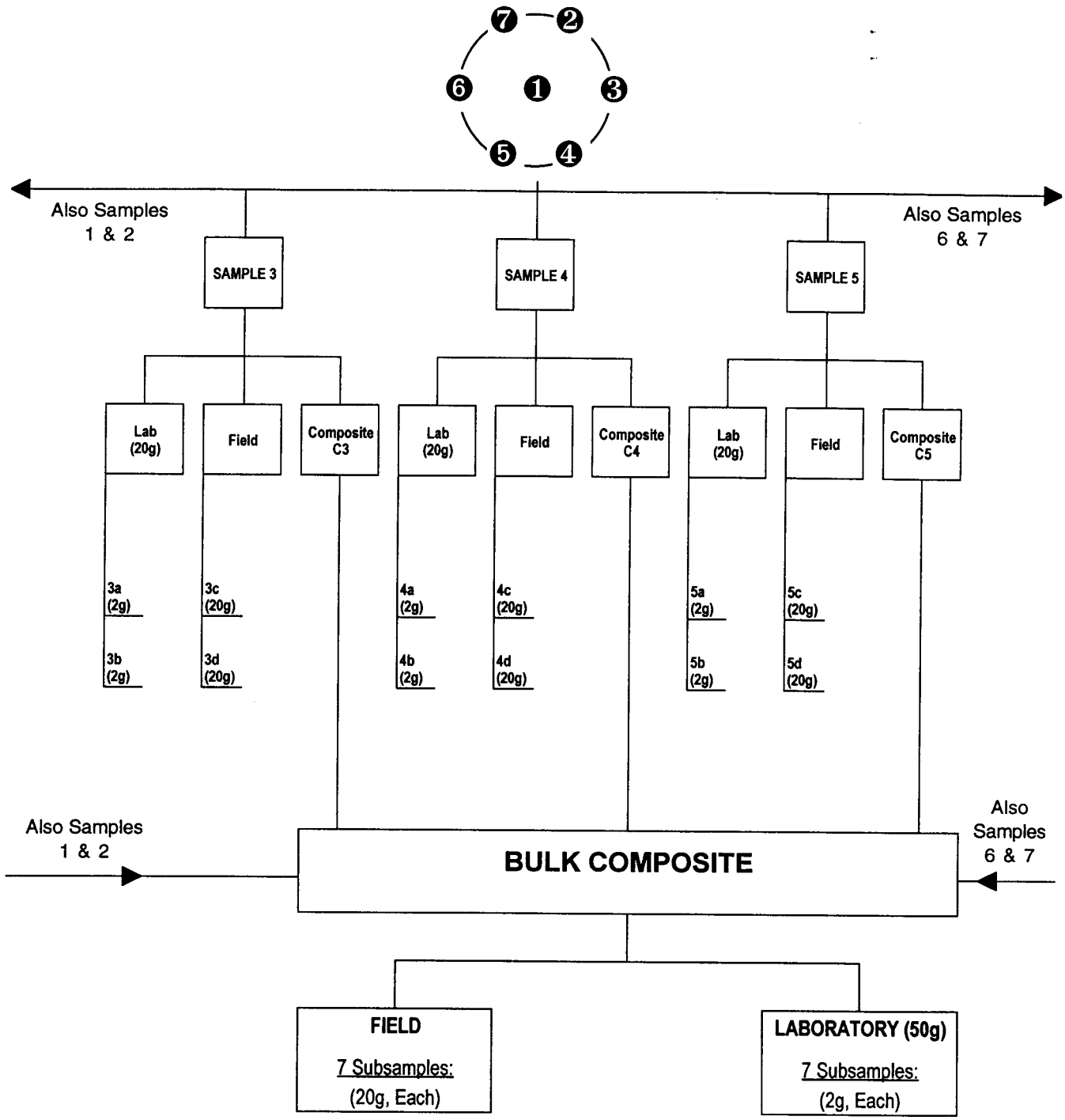

FIG. 2. Sample processing design, illustrated for samples 3–5. Samples from positions 1, 2, 6, and 7 processed identically.

diluted as appropriate such that absorbances after reaction with the EnSys reagent were less than 1.0. For extracts containing mainly TNT, the color intensity prior to reaction with the EnSys reagent served as a rough guide for sample dilution. Because soil concentration varied by such a large amount, with concentrations in excess of 100,000 μg/g, acetone extracts had to be diluted by ratios as high as 1:5000 to provide analyte concentrations in the linear range of the method (0–4 mg/l). In the field these dilutions were performed with the use of glass microliter syringes and graduated cylinders. This type of dilution procedure introduces a minor amount of imprecision, but when it was evaluated, RSDs (relative standard deviations) were less than 3%.

For seven of the nine sampling locations, extracts became reddish when reacted with the EnSys reagent, indicating the likely presence of TNT. For sampling location 2 at the Monite site, a blue-purple color indicated that DNT was

the likely contaminant, rather than TNT. At sampling location 6, acetone extracts were fluorescent yellow in color, indicating the presence of ammonium picrate as the primary contaminant. Samples from locations 2 and 6 were analyzed by appropriate procedures, and the results are presented elsewhere.[22]

Calibration for quantitation was achieved by reacting a known standard of TNT in acetone (containing 3% water) with the EnSys reagent for samples from locations 1, 3, 4, 5, 7, 7R, 8, and 9. Absorbance was measured at 540 nm with a battery-operated spectrophotometer (Hach Model DR/2000). Correction for background color in the soil extracts was obtained by measuring the absorbance of each extract prior to the addition of the EnSys reagent, doubling the value, and subtracting it from the final absorbance after addition of the reagent. Doubling the initial absorbance prior to subtraction takes into account the increased absorbance caused by reaction of humic organic matter in the extract with base, as discussed elsewhere.[8]

Laboratory Analysis for TNT and Other Neutral Nitroaromatics and Nitramines

All soil samples were returned to the laboratory in coolers by overnight carrier and maintained at 4 °C until they were processed. Samples were placed in plastic weighing boats, plant debris and other debris were removed, and the remaining sample was air dried in the dark at room temperature until a constant weight was achieved, usually within 48 hours or less. Weight loss upon drying was used to calculate percent moisture, which was then used to correct field-measured analyte concentrations to a dry-weight basis for comparison with laboratory results. Stones larger than 2 mm were removed from dried samples, which were ground with a mortar and pestle to a fine powder. The weight of stones removed from each sample was recorded. Except for locations 7 and 7R, the amount of stones removed before laboratory analysis did not significantly modify the composition of soil samples from those analyzed in the field. For locations 7 and 7R, the amount removed was large, and this had an effect on the agreement of results from on-site and laboratory analysis, as will be discussed later.

Duplicate 2.00-g subsamples from each discrete soil sample and seven replicate 2.00-g subsamples from composites were extracted and analyzed according to the SW846 Method 8330.[23] Primary analysis was conducted on a Supelco LC-18 column eluted with 1:1 methanol/water at 1.5 ml/min. Absorbance was recorded at 254 nm on a Spectra Physics Model 8490 variable wavelength detector, and peaks were recorded on a Hewlett Packard 3396 Digital Integrator operated in the peak height mode. Selected samples were subjected to second-column confirmation on a Supelco LC-CN column with the use of either 35:65 methanol/water or 23:12:65 acetonitrile/methanol/water, depending on the specific analytes detected in the primary analysis.[24]

Chemicals and Reagents

All standards for TNT and DNT were prepared from Standard Analytical Reference Materials (SARMS) obtained from the U.S. Army Environmental Center, Aberdeen Proving Ground, MD. Standards of TNT and DNT in acetone were prepared with the use of OmniSolv-grade acetone from EM Science.

All acetone used in the field for soil extraction and glassware cleaning was hardware grade, and was obtained locally at each site. For these on-site colorimetric methods, reagent-grade acetone is not necessary, and the use of locally obtained acetone avoids the need to transport a highly flammable solvent. Acetonitrile and methanol used in the laboratory for soil extraction and preparation of HPLC eluents were Baker, EM, or Mallinckrodt HPLC Grade. Water used in the field for cleaning and for addition to extracts was distilled water obtained from local food stores. Laboratory reagent-grade water used for preparation of HPLC eluents was prepared from a Millipore Milli-Q Type 1 reagent-grade water system.

Statistical Analyses

To test for significant concentration differences among sample positions at each sampling location, results from both methods of analysis were subjected to one variable of classification, completely randomized analysis of variance (ANOVA) with the use of CoStat version 1.03 software (CoHost Software, Inc.). For locations 1, 3, and 8, where concentration variations were extremely large, variances were clearly heterogeneous. In these instances, concentrations were log transformed prior to ANOVA. When the ANOVA indicated significant differences among sample positions for a given sampling location, least-significant differences (LSDs) were computed to identify specific differences.

For sampling locations 4, 5, 7, 7R, and 9, concentration ranges were less extreme and variances were adequately homogeneous for ANOVA. A specific test for homogeneity, such as Bartlett's test, was not conducted because there was only one degree of freedom available at each sample position. However, our experience with similar data involving many more degrees of freedom was the basis for this decision. In these cases variances of untransformed results were fractionated to yield estimates of the standard deviations for subsampling plus analysis (S_A) and for the field sampling (S_S). Henceforth, all references to analytical error include contributions from mixing and subsampling, extraction, dilution, measurement, and concentration computations, and sampling error refers to spatial heterogeneity at the sampling location. CoStat software was also used to compute means and standard deviations of duplicates, overall means of the seven duplicates, plus means and standard deviations of composites. Analytical precision of the seven duplicates for each sampling location and each analysis method were expressed as the pooled RSDs.

One-way ANOVA was also used to compare on-site versus laboratory analyses of composites. A paired *t* test and correlation analysis were used to compare on-site versus laboratory analysis for sets of seven samples for a given sampling location. These tests were done with Sigma Stat (Jandel Scientific). In addition to the linear least-squares model with intercept, correlations were also computed for the linear zero-intercept model on untransformed data. When intercepts are close to zero, the correlation coefficient for the zero-intercept model approaches the value for the model with intercept. As an intercept departs from zero, the correlation coefficient *r* for the zero-intercept model will decrease relative to the value for the model with intercept, thereby giving an indication of the significance of the intercept.

For all on-site versus laboratory comparisons, the sum of TNB, TNT, and 2,4-DNT laboratory concentration estimates were compared with on-site measurements. The Janowsky ions formed by the reaction of TNT and TNB with the EnSys reagent both have wavelengths of maximum absorption around 540 nm, and their molar absorptivities at that wavelength are similar.[25] There is a peak with higher absorptivity at lower wavelength, but high humic background makes measurement at this peak wavelength prone to interference. In any case, the on-site TNT method will record the sum of TNT and TNB.[8] The absorptivity of the Janowsky ion from 2,4-DNT is not maximum at 540 nm, but it is significant. However, DNT reacts more slowly with the EnSys reagent than TNT and TNB, and the rate of color formation varies with water concentration in the extract. Because the contribution of DNT to the on-site TNT estimates will depend on analysis conditions, corrections are impractical, so we decided to use the total of these three analytes to represent laboratory concentration estimates.

One further aspect of the statistical analysis requires mention. It has already been noted that total absolute variances for the seven sample positions in some sampling locations were heterogeneous. Furthermore, they were computed without regard to the presence of variable amounts of spatial correlation between positions. It was observed that the spatial correlations were irregular in some cases, in contrast to a regular gradient, such as the directional concentration change that one might find on the edge of a plume of highly mobile compounds. For example, see the pattern of TNT concentrations observed for sampling location 1 (Figure 1). Spatial correlation undoubtedly introduces some bias in the variance estimates, but we believe that the magnitude of this effect is insufficient to have any practical impact on the conclusions.

Results and Discussion

A complete data set is presented for one of the seven TNT locations (Table 1). Results for the other six TNT locations as well as those for the DNT and ammonium picrate locations are available elsewhere.[22] Results for location 1 at the Monite installation are representative of those sites exhibiting extreme localized spatial concentration variations. Location 9 at Volunteer AAP yielded results typical of those with more moderate localized variations.

Monite Installation—Location 1

TNT was the major analyte present at location 1, with concentrations varying from sample to sample by over $2\frac{1}{2}$ orders of magnitude (Table 1). Acetone extracts for on-site analysis were highly colored even before reaction with the EnSys reagent. Extracts for sample positions 2, 3, and 7 were yellow in color; extracts from sample 6 and the composites were orange; and extracts of samples 1, 4, and 5 were dark brick red. These colors are due to phototransformation products of TNT in these surface soils. The intensity of color before reaction with the EnSys reagent correlated very well with the TNT concentrations obtained by the on-site colorimetric method. Reaction of the acetone extracts with the EnSys reagent resulted in the development of red-colored solutions indicative of the presence of TNT. Substantial dilutions (as high as 1:2000) were required to obtain absorbances in the linear range (0.0–1.0 absorbance units) at 540 nm.

Duplicate on-site analyses at sampling location 1 were in excellent agreement (pooled RSD was 5.4%), indicating that homogenization of discrete samples was adequate. Duplicate laboratory analyses varied to a greater extent than on-site analyses (pooled RSD was 14.4%), probably due to the smaller sample size used for lab analysis (2 versus 20 g).

Because TNT concentrations varied by such a large amount from sample to sample, the data were not normally distributed, and absolute variances were not homogeneous. Data were transformed by taking the logarithm of individual values for both the on-site and laboratory results, and an analysis of variance (ANOVA) was performed on both sets of data. For the on-site analyses, the *F* ratio was 233, indicating that a significant difference was detected among the seven discrete samples at greater than the 99.9% confidence level. Results of a least-significant-difference test (LSD) indicated that all seven discrete samples were significantly different from each other at the 95% confidence level. Similar results were obtained when ANOVA was conducted on the laboratory results. Thus, for sampling location 1, very similar conclusions were reached regarding the nature of the analyte distribution with the use of either the results of on-site analysis or results of laboratory analyses.

Because the mean concentrations and absolute variances for samples from location 1 differ so drastically, partitioning variances of untransformed data with the use of normal distribution statistics is not possible. An ANOVA of the log-transformed data indicates that even the log concentrations from various samples differ significantly from one another. A simple way to compare sampling and analytical uncertainties is to compare the ratios of extreme mean concentrations obtained for the seven samples with those for duplicate analyses from the same location. For location 1, the

TABLE 1. Analytical results for Monite site, sampling location 1

Sample	TNT On-site analysis (μg/g)	Laboratory analysis (μg/g)			
		TNB	TNT	2,4-DNT	Total
Discrete samples					
1a	42,700	107	37,500	70	37,700
1b	36,900	104	45,000	–	45,100
2a	492	30	390	–	420
2b	507	30	382	–	412
3a	174	12	113	20	145
3b	154	11	116	–	127
4a	28,000	97	44,400	–	44,500
4b	27,600	–	41,200	–	41,200
5a	24,400	–	33,000	–	33,000
5b	24,400	–	22,400	–	22,400
6a	1,240	42	1,170	–	1,210
6b	1,310	33	1,200	–	1,230
7a	327	23	305	–	328
7b	334	17	227	–	244
Mean	13,500				16,300
Composites					
C1	12,900	–	11,800	–	11,800
C2	12,900	–	13,400	–	13,400
C3	13,300	–	13,600	–	13,600
C4	14,200	–	15,200	–	15,200
C5	13,000	–	13,900	–	13,900
C6	13,200	–	15,000	–	15,000
C7	12,500	–	16,100	–	16,100
Mean	13,100				14,100
Std. dev.	532				1,420
RSD	4.06%				10.1%

ratio of highest mean concentration to lowest mean concentration was 243 for the on-site analyses and 315 for the laboratory analyses. The highest ratios for duplicates were 1.16 for the on-site analyses and 1.47 for the laboratory analyses. Thus for this location, sampling error contributes many times more uncertainty than analytical error for either on-site or laboratory analysis.

Agreement of on-site and laboratory analyses of these discrete samples was compared in two ways. Linear correlation analysis was conducted with the use of the untransformed data with and without intercept and for the log-transformed values with intercept. Correlation coefficients were 0.973, 0.973, and 0.999 for nontransformed data with and without intercept and the log-transformed data, respectively. The correlation coefficient for the zero-intercept model is identical to that for the model with nonzero intercept, and we interpret this to indicate that the intercept is not significantly different from zero and the accuracy of the on-site method relative to the lab method can be estimated from the slope of the best-fit zero-intercept linear least-squares line (81.5%). The excellent correlation for the log-transformed data demonstrates the equivalency of the results for the two methods over several orders of magnitude of concentration.

A paired *t* test was also conducted on the seven log-transformed mean data for the two methods of analysis. The *t* value was 0.07, which is not significant at the 95% confidence level. Results of the paired *t* test agree with those from correlation analysis; that is, the laboratory and on-site results compare very favorably.

Results of the analyses of the composite samples at sampling location 1 were also quite interesting. The mean and standard deviation of the on-site analyses for the composite were 13,100 ± 532 μg/g. By comparison, the mean of the seven discrete samples was 13,500 μg/g (Table 1). Clearly, these means are in excellent agreement, and this was true for results at all locations. For the laboratory analyses, the mean and standard deviation of the seven composites were 14,100 ± 1,420 μg/g, whereas the mean of the results for the seven discrete samples was 16,300 μg/g. These results do not agree quite as well as those for the on-site analyses, but they appear to be quite adequate when compared with the wide range of concentrations found for the discrete samples. ANOVA was conducted to compare the laboratory and on-site results for the composite samples. The *F* ratio of 3.05 indicates that the results of the composite laboratory and on-site analyses for this sampling location were not significantly different at the 95% confidence level. This is true even with the good precision (RSDs of 4.1 and 10.1% for on-site and laboratory analyses, respectively) obtained for the analyses of these composite samples. Thus for this location, a good indication of the degree of contamination

could be obtained with the use of a combination of composite sampling and colorimetric on-site analysis.

Volunteer AAP—Location 9

Acetone extracts from soils at location 9 were light yellow in color, suggesting that, if TNT was present, it was in low concentration. When undiluted extracts were reacted with the EnSys reagent, pink to reddish colored solutions were produced, indicative of the probable presence of TNT. Laboratory analyses confirmed the presence of TNT in these soils at concentrations ranging from 7 to 40 μg/g.

Analytical precision was excellent for both the on-site and lab analyses for samples from location 9. The pooled RSD for the on-site analyses was 4.7% for the discrete samples, and the RSD from replicate analysis of the composite was 9.0%. Likewise, the pooled RSD for lab analysis of the discrete samples was 5.9%, and the RSD from replicate analysis of the composite was 2.8%.

Like sampling locations 4, 5, and 7, results from location 9 appeared to be sufficiently normally distributed to conduct ANOVA without log transformation. When this was done, F ratios of 217 and 321 were obtained for on-site and lab results, respectively, indicating highly significant differences among discrete samples. LSD tests showed that nearly all of the discrete samples were significantly different from one another. When variances were fractionated into analytical and sampling error, the standard deviation for analysis was 1.0 μg/g for both the on-site and laboratory methods. Sampling error estimated from the on-site analysis data was 10.4 μg/g, and from the lab data it was 12.4 μg/g, indicating that sampling error again dominated the total error.

Correlation analysis of the on-site and lab data from location 9 gave a best-fit linear relationship with a slope 0.990, a Y intercept of 0.856, and a correlation coefficient r of 0.984 (Figure 3). The best-fit zero-intercept model had a slope of 1.032 and an r of 0.982, indicating that the intercept was probably not significant. A paired t- test confirmed that the two methods did not yield significantly different results at the 95% confidence level. However, the good analytical precision did produce a significant difference in the replicate analyses of the composite, even though the mean concentrations of the on-site and lab results were 16.6 μg/g and 14.9 μg/g, respectively. The good agreement of these means for sampling location 9 is particularly encouraging, because the range of concentration encountered is quite low (4–40 μg/g).

Comparison of Locations 7 and 7R

Because of the close proximity (15 cm) of corresponding samples from locations 7 and 7R, it seemed worthwhile to compare results, despite the fact that some of the variation in location 7 on-site results can be ascribed to faulty measurement of extracting solvent volume. Concentrations of TNT ranged from about 55,000 to 112,000 μg/g for location 7 and from about 40,000 to 119,000 μg/g for location 7R. The pooled RSD for the on-site analyses for location 7R was 7.1%, compared with 17.2% for location 7. Laboratory results were unaffected by this problem and very similar pooled RSDs were found for locations 7R and 7 (7.0% and 9.6%, respectively).

When these samples were processed in the laboratory, a large percentage of the remaining material proved to be smaller stones (2–5 mm), which were removed prior to laboratory analysis. The material excluded in the laboratory ranged from 51 to 64% and 47 to 67% for locations 7 and 7R, respectively. Samples of the segregated stones were extracted and analyzed in the same manner as the soil. TNT concentrations obtained for the stones ranged from 6025 to 8150 μg/g, but the corresponding soil for these samples had TNT concentrations over 100,000 μg/g. Because the small stones had much lower concentrations of TNT than the soil, their exclusion from the material originally analyzed on site with the colorimetric method accounts for the higher concentrations observed in the laboratory analyses. Although these results for the discrete samples do not offer a valid comparison of on-site versus laboratory accuracy, they did give an identical picture of analyte distribution at this location. Composite samples produced nearly identical concentration estimates for locations 7 and 7R: 57,000 μg/g versus 55,200 μg/g for on-site analysis and 107,000 μg/g for both from the laboratory analyses. The nearly identical results for the composites from locations 7 and 7R give further evidence of the ability to prepare representative composite samples, even when substantial short-range heterogeneity is present.

Summary of Results

To assess the overall performance of the on-site TNT colorimetric method for the three installations, numerical on-site results for sampling locations 1, 3, 4, 5, 8, and 9 were correlated with the corresponding laboratory results.

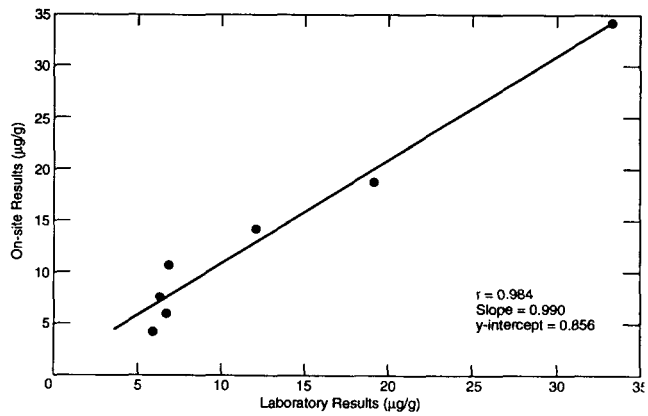

FIG. 3. Correlation analysis for on-site and laboratory analyses for TNT for sampling location 9 at Volunteer AAP.

Data from location 7 were not used because about 50% by weight of each sample processed on site was excluded prior to laboratory analysis, thereby introducing a large bias between methods.

The results for the untransformed means of duplicates for the seven discrete samples at each of the six remaining sampling locations yielded an $r = 0.979$ between the on-site and laboratory results, with a slope of the best-fit linear regression line of 0.867. Because concentrations vary over about five orders of magnitude, the data were plotted on a log-log basis, so that the characteristics of the relationship can be seen equivalently at different absolute concentrations (Figure 4). Clearly the log-log plot shows that the linear relationship between on-site and lab results is very strong for lab values above a log of about 0.6 (corresponding to a concentration of about 4 μg/g). Data below this value are all from sampling location 3, and it is not clear whether the poor correlation for these low concentration samples is specific to location 3 or is simply due to inaccuracy of the method at very low concentrations.

An excellent correlation was also found for the on-site versus lab results for the composite samples for these same six sampling locations; the r value was 0.989, with a slope of the best-fit relationship of 0.999. Each point in this relationship represents a mean of seven on-site and seven lab determinations. For both the discrete samples and the composites, the correlation coefficients for the best-fit linear relationships with zero intercept were equal to those with nonzero intercept, which we interpret to indicate that the Y intercepts were not significantly different from zero and that the slope (of the zero-intercept model) can be considered an overall measure of the accuracy of the on-site method relative to the lab method. With the use of this interpretation and the computed slopes from the zero-intercept models, the accuracy across concentrations varying from near the detection limit of 1 μg/g to over 40,000 μg/g was estimated to be 87.6% for the discrete samples and 100.5% for the composites. Clearly, use of the on-site colorimetric TNT method is justifiable from nearly any conceivable data quality objective, particularly where we have direct evidence of the short-range heterogeneity present in soil concentrations.

The data from this study can also be used to put in perspective the uncertainty introduced in results by analysis relative to that from sampling. Although random grab sampling is appealing from a cost perspective, it may be totally inadequate for decisions about the need for, or adequacy of, remediation. To provide data that can satisfy this need with a high level of confidence, the total uncertainty associated with site characterization must be understood and reduced to acceptable levels. Until now, little or no information has been available where the components of error have been quantified for soil characterization at explosives-contaminated sites. For some of the sampling locations studied here, analyte distributions exhibited such extremes that use of untransformed data to fractionate the error was not a valid approach. For locations 4, 5, 7, 7R, and 9, however,

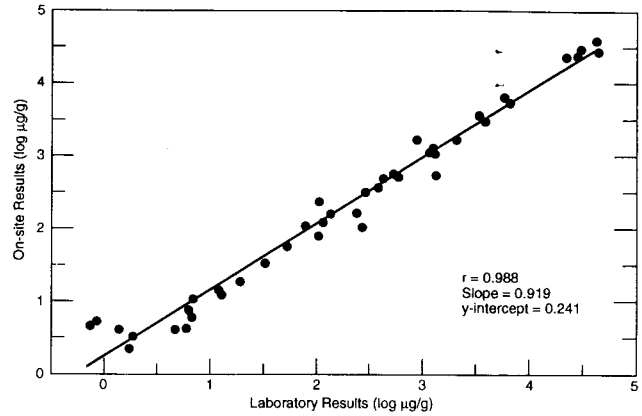

FIG. 4. Correlation of on-site and laboratory analyses for TNT from sampling locations 1, 3–5, 8, and 9.

we were able to fractionate the total error variances because concentration variations were modest (Table 2). For these four locations, standard deviations due to analysis, whether on site or laboratory, were always much lower than the corresponding standard deviations due to sampling, and hence total error was dominated by sampling error. For the other locations, sampling error was even greater and so overwhelmed analytical error that this type of fractionation would only be possible with the use of asymmetric (logarithmic) limits. Clearly, if we want to make a significant improvement in the quality of site-characterization data, the major effort should be placed on reducing sampling error. Single grab samples are totally inadequate.

To reduce sampling error, the material analyzed must be more representative of average concentrations within the area the sample is supposed to represent. For the data here, if we assume that the mean analyte concentration of the seven samples taken from a circle with 122-cm diameter is the true concentration, we can assess the difficulty in achieving representativeness by looking at the ratio of highest to lowest values in the group of seven mean determinations. These ratios, presented in Table 3 under the heading of local heterogeneity, range from 3.8 to 243 for the on-site TNT method and 3.0 to 315 for the lab method. Much larger grids than the areas we sampled are typically used for site characterization, and this would only serve to further increase uncertainties due to sampling. Analysis of composite samples, however, gave results that were good estimates of the means of the seven discrete samples, with low standard deviations (Table 4). It is also useful to note that standard deviations for the on-site analysis of all of the composite samples are low relative to mean concentration (low RSDs), indicating that the in-field homogenization procedures used were adequate. Thus, the number of analyses of the composite required to produce data with a high degree of confidence is low.

Compositing is an effective way to reduce intersample variance caused by heterogeneous distribution of contaminants. The total variance for the formation and analysis of

TABLE 2. Fractionation of total error into analytical and sampling components.

Sampling location	Standard Deviation				Ratio	
	Analytical		Sampling		Sampling/Analytical	
	On site	Lab	On site	Lab	On site	Lab
Hawthorne location 4	217	265	1,970	2,150	9.1	8.1
Hawthorne location 5	5.3	11.0	121	131	22.8	11.9
Volunteer location 7	*	7,680	*	19,800	*	2.6
Volunteer location 7R	5,120	6,320	24,700	27,600	6.1	4.4
Volunteer location 9	1.0	1.0	10.4	12.4	10.4	12.4

* Data are considered unreliable, as discussed in the text.

composites can be expressed as

$$(C_T)^2 = \frac{(C_S)^2}{n} + \frac{(C_A)^2}{k},$$

where C_T is the total percent relative standard deviation, C_S and C_A are the percent relative standard deviations of sampling and analysis, respectively, n is the number of discrete samples formed into a composite, and k is the number of replicate analyses performed on the composite. In Table 5 we show values of C_T for various combinations of C_S, C_A, n, and k. The values chosen for C_S and C_A are typical of those found here for on-site or laboratory analyses of TNT. There would be nothing to prevent using larger values of n, but there is almost no benefit in using larger values of k, given the relationship of C_S to C_A. If desired, plots of C_T versus n could be formed for various values of C_S, C_A, and k. We should also remember that the values of C_T are for a single composite. Uncertainty in a mean of several composites would be reduced by $\frac{1}{\sqrt{N}}$, where N is the number of composites averaged.

From Table 5 it is very obvious that improved reliability of concentration estimates can only be realized by reducing the magnitude of C_S relative to C_A. On-site analysis is just as reliable as laboratory analysis for TNT in surface soils, because the analysis step does not contribute much error in either case. When we look at the cost estimates (Table 5) for on-site versus laboratory analysis and combine that with the fast turnaround of on-site analysis, the advantages of on-site analysis are clear. In arriving at the cost of on-site analyses, all materials and their disposal were included, along with capital equipment costs and labor. An allowance was also made for 10% of the samples to be sent for confirmatory laboratory analysis. Clearly, the cost of compositing is relatively small compared to the benefits. Unless C_S is much lower than found for the sites studied so far, the major justification for performing replicate analyses of composites would be to identify procedural mistakes.

The approach to characterizing a new site should involve a preliminary field survey to obtain information on the magnitude of both short- and long-range spatial heterogeneity. From these results, a flexible sampling plan would evolve with the understanding that it was subject to modifications (if necessary) as results accumulate.

TABLE 3. Comparison of measures of analytical precision, accuracy, and discrete sample representativeness.

Sample location	Precision				Accuracy	Local Heterogeneity	
	Pooled RSD of Duplicates		Largest Concentration Ratio of Duplicates		Slope of 0-Intercept Model	Ratio of Highest Mean Concentration vs. Lowest for Discrete Samples	
	On site	Lab	On site	Lab	On site vs. Lab	On site	Lab
1	5.4	14.4	1.16	1.47	0.815	243	315
3	22.4	7.5	1.82	1.19	1.464	50.0	98.1
4	16.1	19.6	1.70	1.99	0.911	69.0	58.1
5	4.1	5.7	1.13	1.16	0.847	28.9	29.5
7R	7.1	7.0	1.27	1.21	0.677	3.8	3.0
8	23.3	5.6	1.73	1.19	1.070	53.1	55.6
9	4.7 (pooled)	5.9 (pooled)	1.13	1.17	1.032	8.2	5.7
Mean	14.7	11.3	1.42	1.34	0.974	65.1	80.7

TABLE 4. Comparison of TNT results for discrete and composite soil analysis.

Installation	Sampling location	On site or lab	Discrete samples (mean ± SD*)	Composite samples (mean ± SD)
Monite	1	OS	13,500 ± 16,800	13,100 ± 532
		L	16,300 ± 20,200	14,100 ± 1,420
Monite	3	OS	19.8 ± 42.0	12.6 ± 1.2
		L	12.9 ± 29.0	4.16 ± 0.7
Hawthorne	4	OS	1,970 ± 1,980	1,750 ± 178
		L	2,160 ± 2,160	2,000 ± 298
Hawthorne	5	OS	156 ± 121	139 ± 16.6
		L	168 ± 131	193 ± 7.7
Volunteer	7	OS	84,900 ± 33,400	57,000 ± 2,600
		L	89,100 ± 20,500	107,000 ± 9,230
Volunteer	7R	OS	57,500 ± 25,000	54,800 ± 5,840
		L	86,900 ± 27,900	107,000 ± 7,520
Volunteer	8	OS	9,920 ± 12,000	11,300 ± 2,020
		L	8,910 ± 11,600	9,620 ± 409
Volunteer	9	OS	13.7 ± 10.4	16.6 ± 1.5
		L	13.0 ± 10.3	11.8 ± 0.3

The discrete sample standard deviations for locations 1, 3, and 8 are all larger than their corresponding means because the results from these locations are not normally distributed. These results may be log-normally distributed, in which case the data should be transformed.

Conclusions

The results presented here exhibit several unifying themes that can be applied in designing future investigations of munitions-contaminated sites. First, it is clear that there was extreme heterogeneity at all sampling locations. A single sample from any of the 122-cm diameter circles could differ by orders of magnitude from the mean concentration of the small area sampled. Relative standard deviations (RSDs) for the seven discrete samples were often greater than 100%.

A second consistent finding was that composite samples of the seven discrete samples could be reliably homogenized and subsampled on site. This also opens the possibility of compositing discrete samples representing a larger area if concentration variations suggest that this approach

TABLE 5. Dependence of total percent relative standard deviation (C_T) on compositing and analysis schemes using various assumed values for sampling and analysis standard deviations.

Number of samples composited (n)	Number of replicate analyses (k)	(C_S)	Percent relative standard deviations: Sampling (C_A)	Analysis (C_T)	Total On site	Cost of procedure ($): Lab	
1	1		50	10	51.0	81	337
4	1		50	10	26.9	86	342
7	1		50	10	21.4	90	347
7	2		50	10	20.2	166	680
1	1		100	10	100.5	81	337
4	1		100	10	51.0	86	342
7	1		100	10	39.1	90	347
7	2		100	10	38.5	166	680
1	1		150	10	150	81	337
7	1		150	10	57.6	90	347
7	2		150	10	57.1	166	680
1	1		100	5	100	81	337
7	1		100	5	38.1	90	347
7	2		100	5	38.0	166	680
1	1		50	20	53.9	81	337
7	1		50	20	27.5	90	347
7	2		50	20	23.6	166	680
7	1		100	20	42.8	90	347
7	2		100	20	40.4	166	680

would be desirable. Most importantly, it permits on-site processing without elaborate apparatus.

Another major finding was that the specificity and accuracy of the on-site TNT method was quite adequate. The two locations where TNT was not the major contaminant were readily identified, and the seven locations where TNT appeared to be the primary contaminant were confirmed by the reference HPLC method. The concentration estimates from on-site analyses agreed very well with laboratory estimates, except for location 7, where a major bias was introduced by removing small stones during the grinding operation prior to laboratory analysis. For the other six TNT locations, the agreement was excellent. Admittedly, there were small but statistically significant differences in mean concentration estimates at some locations, but their magnitude was insufficient to impart meaningful differences in conclusions. Of course, each site should include some reference laboratory analyses to validate the on-site analyses.

Perhaps the most surprising finding was the consistency of the overall analysis precision for TNT. For the seven locations where TNT was the primary contaminant, RSDs for duplicate on-site analyses of subsamples of the discrete samples ranged from 4.1 to 23.3%, with a pooled value of 14.7%. Comparable laboratory results yielded RSDs from 5.6 to 19.6%, with a pooled value of 11.3%. Replicate analyses of composites produced RSDs ranging from 4.1 to 17.9% (pooled = 10.6%) for on-site results and 2.8 to 15.9% (pooled = 9.6%) for laboratory analyses. The estimates are approximately equal for composites despite the extra mixing step, probably because the wide concentration variations of discrete samples required large differences in dilutions and because of the 10-times-larger sample size used for the on-site method. Nonetheless, the consistency of the pooled estimates is both surprising and reassuring. We believe that subsampling and analysis (S_A) typically yields RSDs of about 10% for both on-site and laboratory methods.

The approach to a new site should involve a preliminary field survey to obtain information on the magnitude of both short- and long-range heterogeneity. From these results a flexible sampling plan could evolve with the understanding that it would be subject to modifications (if necessary) as results accumulate. It is our intention to conduct one or more such studies (demonstration projects) as the next phase of this research. Characterization with a combination of composite sampling, adequate in-field sample homogenization, and on-site colorimetric analysis is an efficient method of producing data that are not only accurate and precise, but are also representative of the area.

Acknowledgments

The authors gratefully acknowledge Clyde Murray, Bureau of Land Management, Carson City, Nevada; Douglas Peery, Quanterra Corporation, Knoxville, Tennessee; Sophie Hong, U.S. Army Corps of Engineers, Sacramento District; and Harry Craig, U.S. EPA Region 10, Portland, Oregon for providing assistance in locating appropriate sites to conduct this field sampling study. Jane Mason, CRREL, is acknowledged for assistance in preparation of this manuscript. This work was funded by the U.S. Army Enviromnental Center, Aberdeen Proving Ground, Maryland, Martin H. Stutz, project monitor.

References

1. R. O. Gilbert, *Statistical Methods for Environmental Pollution Monitoring* (1st ed.), Van Nostrand Reinhold Co., New York; 1987.
2. M. H. Huesemann, "Guidelines for the development of effective statistical soil sampling strategies for environmental applications" in *Hydrocarbon Contaminated Soils and Groundwater,* ed. by E. J. Calabrese and P. T. Kostecki, Association for the Environmental Health of Soils, Amherst, MA, 1994, Vol. 4, Chap. 4, pp. 47–96.
3. L. H. Keith, G. L. Patton, D. L. Lewis, P. G. Edwards and M. A. Re, "A New Calculation Tool for Estimating Numbers of Samples" in *Proceedings of the Eleventh Annual Waste Testing and Quality Assurance Symposium,* Washington, DC, July 23–28, 1995, pp. 550–564.
4. J. J. van Ee, L. J. Blume, and T. H. Starks, "A Rationale for the Assessment of Errors in the Sampling of Soils," (EPA Report No. EPA/600/4-90/013), Environmental Monitoring Systems Laboratory, Las Vegas, NV, 1990, p. 57.
5. L. R. Williams, "Criteria-based approach for solicitation and selection of field-measurement technologies" *Am. Environ. Lab.* **8**(2), 36–37, (1996).
6. T. E. Barnard, "Extending the concept of data quality objectives to account for total sample variance" in *Principles of Environmental Sampling* (2nd ed.) ed. by L. H. Keith (in press).
7. U.S. EPA, "Soil screening for trinitrotoluene (TNT)," SW846 Method 8515, Third Update (Draft), January 1995.
8. T. F. Jenkins and M. E. Walsh, "Development of field screening methods for TNT, 2, 4-DNT, and RDX in Soil" *Talanta* **39,** 419–428, (1992).
9. E. K. Triegel, "Sampling variability in soils and solid wastes" in *Principles of Environmental Sampling,* ed. by L. H. Keith, American Chemical Society, Washington, DC, 1994, pp. 385–394.
10. E. K. Triegel, L. Guo and S. Ward, "Cost-effective sampling approaches" presented at *American Chemical Society Symposium of Principles of Environmental Sampling,* March 16, 1994, San Diego, CA.
11. M. J. Barcelona, "Overview of the sampling process" in *Principles of Environmental Sampling,* ed. by L. H. Keith, American Chemical Society, Washington, DC, 1988, pp. 1–23.
12. F. Smith, S. Kulkarni, L. E. Myers and M. J. Messner, "Evaluating and presenting quality assurance sampling data" in *Principles of Environmental Sampling,* ed. by L. H. Keith, American Chemical Society, Washington, DC, 1988, pp. 157–168.
13. T. B. Parkin, "Soil microsites as a source of denitrification variability" *Soil Sci. Soc. Am. J.* **51,** 1194–1199, (1987).
14. W. E. Sabbe and D. B. Marx, "Soil sampling: spatial and temporal variability" in *Soil Testing: Sampling, Correlation, Calibration, and Interpretation* (SSSA Spec. Publ. No. 21), ed. by J. R. Brown, SSSA, Madison, WI, 1987, pp. 1–14.
15. J. L. Starr, T. B. Parkin and J. J. Meisinger, "Influence of sample size on chemical and physical soil measurements" *Soil Sci. Soc. Am. J.* **59,** 713–719, (1995).
16. D. R. Cameron, M. Nyborg, J. A. Toogood and D. H. Laverty, "Accuracy of field sampling for soil tests *Can. J. Soil Sci.* **51,** 165–175, (1971).
17. M. C. Fabrizio, A. M. Frank, and J. F. Savino, "Procedures for forma-

tion of composite samples from segmented populations" *Environ. Sci. Technol.* **29,** 1137–1144, (1995).
18. F. C. Garner, M. A. Stapanian and L. R. Williams, "Composite sampling for environmental monitoring" in *Principles of Environmental Sampling,* ed. by L. H. Keith, American Chemical Society, Washington, DC, 1995, pp. 363–374.
19. J. Paasivirta and R. Paukku, "Use of composited samples to optimize the monitoring of environmental toxins" *Chemosphere* **19,** 1551–1562, (1989).
20. R. S. Parrish, G. O. Ware, C. N. Smith and P. A. Banks, "Composite sampling techniques for determining pesticide concentrations" in *Proceedings of the 1990 Kansas State University Conference on Applied Statistics in Agriculture,* ed. by G. A. Milliken and J. R. Schwenke, Department of Statistics, Kansas State University, Manhattan, KS, 1990, pp. 21–32.
21. D. J. Schaeffer, H. W. Kerster and K. G. Janardan, "Grab Versus Composite Sampling: A Primer for the Manager and Engineer" *Environ. Mgmt.* **4,** 157–163 (1980).
22. T. F. Jenkins, C. L. Grant, G. S. Brar, P. G. Thorne, T. A. Ranney and P. W. Schumacher, *Assessment of Sampling Error Associated with the Collection and Analysis of Soil Samples at Explosives Contaminated Sites* (CRREL Report No. CRREL SR 96-15), U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, NH, 1996.
23. U.S. EPA, "Nitroaromatics and Nitramines by HPLC" (Second Update), SW846 Method 8330, 1995.
24. T. F. Jenkins and S. M. Golden, "Development of an improved Confirmation separation Suitable for use with SW846 Method 8330 (CRREL Report No. CRREL SR 93-14), U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, NH, 1993.
25. T. F. Jenkins, M. E. Walsh, P. W. Schumacher and P. G. Thorne, "Development of colorimetric field screening methods for munition compounds in soil" in *Proceedings of Environmental Monitoring and Hazardous Waste Site Remediation,* ed. T. Vo-Dinh and R. Niessner, Munich, Germany, 19–21 June, 1995, pp. 324–333.