CONTAMINATED SITES STATISTICAL APPLICATIONS GUIDANCE DOCUMENT NO. 12-10

COMPOSITE SAMPLES

A guide for regulators and project managers on the use of composite samples

This guidance document is one of a series that outlines important basic statistical concepts and procedures that are useful in contaminated sites studies. BC Environment recommends that these suggestions be followed where applicable, but is open to other techniques provided that these alternatives are technically sound. Before a different methodology is adopted it should be discussed with BC Environment.

April 2001

THE GENERAL IDEA

Samples are necessary in all phases of a contaminated site study. The money budgeted for analysis can be used more effectively if discrete samples from homogeneous areas are grouped together and combined into composite samples. For example, the separate analysis of five discrete samples is going to be about five times more costly than the analysis of a single composite sample created from the five discrete samples. If the samples are from an area that is thought to be uncontaminated, it is possible that the five separate analyses of the five discrete samples will be virtually identical, with each one confirming what we already suspected — that there is no contamination. The single analysis of a composite sample might be able to confirm the lack of contamination for a much lower cost.

The problem with such an approach is that it may complicate the task of classifying the material. For example, if we are trying to determine whether the arsenic concentration of discrete samples from a particular area ever exceeds 30 ug/g, and if a composite composed of five discrete samples has an average concentration of 12 ug/g, then how do we know if all of the five discrete samples would also have been below 30 ug/g had they been analyzed individually? It is possible that all five discrete samples had concentrations of about 12 ug/g, as shown for the first composite in Table 1, and that there is no significant contamination in the area. It is also possible, however, that some of the five samples had arsenic concentrations above the 30 ug/g threshold while others had virtually no arsenic, as shown by the second composite sample in Table 1.

Table 1 Arsenic concentrations (in ug/g).

	Discrete Samples					Composite Average
Composite 1	11	9	12	11	17	12
Composite 2	1	3	35	16	5	12

This document discusses the practice of combining discrete samples into composites and provides some recommendations on when it is appropriate; it also provides two guidelines for assessing whether a composite analysis is compliant, i.e. whether all the discrete values likely would fall below a specified threshold. The first of these is a rather strict guideline that requires only that we know the number of discretes that went into the composite. The second is less strict but requires that we know in advance the variability between discretes within the composite. Though this second approach requires additional analytical work at the outset of the project, it may be more cost effective in the long run.

There is another common usage of the term "composite" that refers to an average of the analyses from contiguous samples, usually from the same well or borehole. With this type of compositing, the individual samples have already been separately analyzed and the goal of the compositing is either to reduce the number of data that need to be handled, to reduce their variability in order to facilitate statistical interpretation or to standardize samples of varying core length to a common length. This other type of compositing is less of a problem than the compositing of discrete samples prior to analysis since the separate analyses are, in fact, still available and, if necessary, can be used in statistical analysis. This guidance document does not address this other type of compositing but focuses instead on the compositing of discrete samples prior to analysis and on the interpretation of the analysis of such a composite sample.

Other guidance documents in this series provide additional information on related issues. In particular, the document entitled *SAMPLING PLANS* discusses the analysis of spatial variability and also discusses issues related to the number of discrete samples that will be needed to achieve a desired confidence in statistical predictions.

IS COMPOSITING APPROPRIATE?

The primary goal of compositing in contaminated site studies is to keep down the cost of analysis by analyzing fewer samples. Unfortunately, for those who are planning the remediation and for those responsible for ensuring that material has been classified correctly, fewer samples means less information. Planners and regulators could be more confident of the success of the remediation if analyses were available for every discrete sample.

The key to appropriate compositing is to ensure that the samples being combined together have similar concentrations, such as those in the first composite in Table 1. When discrete samples that go into a composite have concentrations that differ considerably, such as those in the second composite in Table 1, the analysis of the resulting composite sample is of little value to anyone. Neither remediation planners nor regulators can make much use of it since the individual discrete samples may represent entirely different levels of contamination that would be classified in different regulatory categories. In such an event, it is likely that the planners or regulators (or both) will eventually need to have separate analyses for the individual discrete samples, at which point the whole exercise of compositing has actually ended up costing more than the separate analyses of the discretes would have cost in the first place.

In situ characterization of the site is the best basis for delineating areas within which the material can be composited. The

prediction of contaminant concentrations for unsampled material will be more accurate and reliable for *in situ* material than for material that has been disturbed or stockpiled. As long as the material remains *in situ*, models of the spatial distribution of a contaminant, such as contour maps, geostatistical simulation, or the results of flow simulation, will be able to benefit from historical, geological and hydrogeological information. If the material is carefully tracked as it is excavated and stockpiled, then *in situ* characterization will be useful in determining the sample-to-sample variability in the stockpiles. If stockpiled material cannot be traced to its *in situ* location, or if a careful *in situ* characterization was never performed, then the only way to assess the sample-to-sample variability in excavated and stockpiled material is through extensive (and costly) sampling.

The fundamental motivation for compositing is to reduce the money spent on analysis. We get the most for our sampling dollar when the samples we analyze are informative about all of the unsampled material that we could not analyze. We would be foolish to squander the opportunity to analyze and interpret *in situ* samples. An *in situ* sample is much more likely to be representative of its immediate surroundings than is a sample taken from excavated or stockpiled material. With *in situ* samples and *in situ* characterization, we can make more reliable predictions about the areas that will be sufficiently homogeneous to warrant compositing.

If we know that compositing is eventually going to be considered, and that *in situ* variability will become a key issue, we should attempt to document the spatial variation of the *in situ* material. In this series of guidance documents, there is one entitled *SAMPLING PLANS* that discusses methods for describing and documenting the spatial variability of *in situ* material.

Since homogeneity is the key to the technical and economic success of compositing, it is important to check on a regular basis the discrete samples within a composite to ensure that their values do not fluctuate too much. One in every ten composites should be chosen at random to have all of its discrete samples analyzed individually. As long as the information gathered from these regular checks of the within-composite variability continues to confirm that composites are homogeneous, then the compositing of samples can continue. If these regular checks demonstrate that there is much more within-composite variability than was originally assumed, then the compositing should stop and the discrete samples should be analyzed individually. Compositing should not resume until the reasons for the lack of homogeneity are well understood and documented.

HOW TO USE COMPOSITE ANALYSES

If composite homogeneity is not documented

If we are trying to decide whether any of the N individual discrete sample values might be above a specified threshold, T, and if no information exists on the variability of individual discrete sample values within a composite, then the only prudent approach is to compare the analytical value from the composite to T ÷ N. This T ÷ N rule is justified by the fact that if any single discrete value in the composite is larger then T, then the average of N equally-weighted discretes will be larger than T ÷ N. When we observe a composite average that is less than T ÷ N, we can be sure that none of the contributing discrete samples had a value greater than T (as long as each of the discrete samples contributed the same amount of material to the composite sample).

This particular approach is very strict in the sense that it frequently leads to false positive errors — cases in which we incorrectly reject a composite as non-compliant when all of its individual discretes were, in fact, compliant. The first composite shown in Table 1, for example, would have to be treated as non-compliant under this rule. With five discrete samples contributing to this composite, and with the threshold for the arsenic concentration in any individual sample being 30 ug/g, our composite would have to produce an analytical value of 6 ug/g or less before it would be considered as compliant according to the T ÷ N rule.

A further drawback of the T ÷ N rule is that it discourages compositing large numbers of discrete samples regardless of their homogeneity. Once the number of discrete samples in the composite reaches about ten, it becomes virtually impossible to satisfy the T ÷ N requirement. With many contaminants, the thresholds that define contaminated material are low enough that T ÷ N rapidly approaches the detection limit of the best available analytical procedures.

Documenting composite homogeneity

The strictness of the T ÷ N rule stems from the fact that it does not accommodate information about the variability (or lack of it) in the N discrete values that go into each composite. If we have gathered information on the actual variability of discrete sample values within a composite, then we can use this in a less strict rule.

To measure the variability of individual discrete sample values within a composite, we need to compare several discrete sample values to the analytical value of their corresponding composite. If we have N composite samples, each one consisting of M discrete samples, then we can calculate the variance of discrete sample values within the same composite and the corresponding standard deviation:

$$s^2_{\text{within composite i}} = \frac{1}{M}\sum_{j=1}^{M}\left[D_{i,j} - C_i\right]^2$$

$$s^2_{\text{within}} = \frac{1}{N}\sum_{i=1}^{N} s^2_{\text{within composite i}}$$

$$s_{\text{within}} = \sqrt{s^2_{\text{within}}}$$

where $C_1, \ldots, C_N$ are the N composite analyses and $D_{i,j}$ is the analytical value of the j-th discrete sample in the i-th composite.

Tables 2 and 3 show an example of this calculation from six composites, each of which contains five discrete samples. Table 2 shows the 30 discrete analyses and their corresponding composite analysis; note that the composite analysis may not be the same as the mean of the corresponding discrete sample values. Table 3 shows the values of $[D_{i,j} - C_i]^2$ for all 30 discrete samples along with the within-composite variance for

each of the five composites. The average within-composite variance for these data is 7.33. The within-composite standard deviation based on these data is therefore 2.71 ug/g.

Table 2 Discrete and composite arsenic values used to calculate the within-composite standard deviation.

	Discrete Sample Analyses					Composite Analysis
Composite 1	12	7	10	12	16	11
Composite 2	10	13	15	12	15	13
Composite 3	22	16	15	16	18	18
Composite 4	7	10	2	5	10	7
Composite 5	17	9	15	12	11	12
Composite 6	8	12	7	13	6	8

Table 3 Values of squared differences for the 30 discrete analyses and their composite analysis in Table 2.

	Squared Difference from Composite Analysis					Within-composite variance
Composite 1	1	16	1	1	25	8.8
Composite 2	9	0	4	1	4	3.6
Composite 3	16	4	9	4	0	6.6
Composite 4	0	9	25	4	9	9.4
Composite 5	25	9	9	0	1	8.8
Composite 6	0	4	1	25	4	6.8
						Average within-composite variance = 7.33

This method requires at least 30 discrete samples in order to produce a good estimate of the within-composite standard deviation. As presented above, the calculation assumes that there are the same number of discrete samples in each composite. If the number of discrete samples varies from composite to composite, then the averaging of the N within-composite variances should be weighted by the number of discrete samples within each composite:

$$s^2_{within} = \frac{\sum_{i=1}^{N} n_i \cdot s^2_{within\ composite\ i}}{\sum_{i=1}^{N} n_i}$$

where n_i is the number of discrete samples in the i-th composite.

The within-composite standard deviation has several uses. Its first use is that it allows us to quantify the degree of homogeneity of the composites. A common yardstick for deciding that a population is reasonably homogeneous is to check to see if the coefficient of variation (CV) is bigger or smaller than 1. The CV is the ratio of the standard deviation to the mean; a CV less than 1 means that there are few erratic high values in the population. As long as the value of s_{within} remains less than the mean of the composited values, we have statistical support for our assumption that the material is sufficiently homogeneous to warrant compositing. For the data shown in Table 2, their within-composite standard deviation was calculated earlier as 2.71 ug/g; with the average composite value being larger than this, the coefficient of variation is certainly less than 1, and compositing of the discrete samples is acceptable.

If composite homogeneity is documented

The second, and more important, use of s_{within} is that it permits the development of a less strict rule regarding the interpretation of the composite's analytical value. Once we have an accurate estimate of s_{within}, when we are trying to decide if any of the discrete sample values in a composite might have a concentration above a threshold T, then we can compare the analytical value of the composite to the following quantity:

$$\text{Composite compliance threshold} = T - 3s_{within} \times \left[1 + \frac{1}{\sqrt{N}}\right]$$

The idea behind this rule is that we can be reasonably sure that no single discrete sample exceeds T if the mean of the discrete samples (which is assumed to be the same as the analytical value from the composite) is three standard deviations below T. There is some uncertainty on the mean, however, since we have only a few samples. The $1/\sqrt{N}$ term in the square brackets moves the composite compliance threshold a little bit lower so that even when the fluctuation on the mean is taken into account, we can still be reasonably sure that the threshold T is at least three standard deviations above the population mean.

Apart from the assumptions that the discrete samples contribute the same amount of material to the composite and that the discrete sample values are all uncorrelated with each other, there are no other assumptions hidden in this approach. If we are willing to be a little bolder, and assume that the discrete sample values follow a normal distribution (not a very defensible assumption since contaminant concentrations for discrete samples are usually quite skewed) then we can be more specific about the actual probability that a discrete sample value exceeds T when the composite's analytical value is below the compliance threshold provided by the formula. Under an assumption of normality, this probability is less than 1%. There is no particular need to assume normality, however; even with no assumption about the distribution of the discrete samples, this probability is never more than 10%. Further details on the calculation of these probability values can be found in the guidance documents entitled *DISTRIBUTION MODELS* and *NON-PARAMETRIC METHODS*.

As an example of the use of this formula, consider the situation from Table 1, where we are combining N=5 discrete samples in our composites and we have a threshold of T=30 ug/g for the arsenic concentration in a single sample. If we use the within-composite standard deviation that we calculated earlier, s_{within} = 2.71 ug/g, then the composite compliance threshold for this situation is:

$$\text{Composite compliance threshold} = 30 - 3 \times 2.71 \times \left[1 + \frac{1}{\sqrt{5}}\right] = 18.2 \text{ ug/g}$$

If a composite has an analytical value less than 18.2 ug/g it is very unlikely that any of its individual discrete sample values would exceed 30 ug/g. All six of the composites listed in Table 2

would count as compliant samples under this rule; under the T ÷ N rule, all of them would be viewed as non-compliant. From the actual discrete values listed in Table 2, we can see that none of the discrete samples is, in fact, above 30 ug/g. Although the rule based on s_{within} is less stringent than the T ÷ N rule, its false negative rate is still very low.

The key to this approach is the use of the actual within-composite standard deviation from composites whose discretes have also been analyzed as an estimate of the within-composite standard deviation for composites whose discretes have not been individually analyzed. This assumes that the composites from which the standard deviation is borrowed belong to the same population as the composites to which the standard deviation is being applied. We need to make sure that we are not mixing apples and oranges when we use a statistic calculated from one set of data as an estimate of a critical parameter for a different set of data.

We should analyze all of the discrete samples for one in every ten composites and use this information to monitor fluctuations in the statistics of the samples. If the composite mean or standard deviation changes unexpectedly, we should consider whether the within-composite standard deviation based on historical information remains an accurate estimate of the within-composite standard deviation of the composites we are currently creating. As new discrete samples and their corresponding composite analyses become available, we should also use this additional information to continuously update and improve our estimate of the within-composite standard deviation.

RECOMMENDED PRACTICE

1. If compositing is likely to be used on a project, use the available discrete samples to establish the degree of *in situ* spatial variability.

2. Use composite samples only after *in situ* characterization has established that all of the material within a particular area belongs to the same regulatory category. If an *in situ* characterization has not been done, then collect enough samples to document that the material has a coefficient of variation less than 1.

3. When compositing samples:

   (a) maintain a clear record of the samples that contribute to each composite;

   (b) homogenize each discrete sample before drawing the sub-sample that will contribute to the composite;

   (c) ensure that the individual discrete samples each contribute the same amount of material to the composite; and

   (d) archive a sufficient quantity of each sample to permit the discrete samples to be analyzed in the event that the composite is non-compliant.

4. From the first several composite samples, select a group that collectively contain at least 30 discretes, analyze the individual discrete samples as well as the composite sample and calculate the standard deviation of the discrete sample values about their respective composite values.

5. In every group of ten composites, randomly select one and, in addition to the analysis of the composite sample, also perform analyses on the individual discrete samples. Use this information to monitor the homogeneity of the composites and to improve the estimate of the within-composite standard deviation.

6. To decide if it is reasonable to suppose that the N discrete samples within a composite all have a concentration less than the threshold T, compare the composite's analytical value to the following quantity:

$$\text{Composite compliance threshold} = T - 3 \cdot s_{within} \cdot \left[1 + \frac{1}{\sqrt{N}}\right]$$

   where s_{within} is the standard deviation of the discrete samples about their corresponding composite analysis and is based on at least 30 actual analyses of discrete samples and the corresponding composites.

7. If there are less than 30 analyses of discrete samples and the corresponding composites, then compare the the composite's analytical value to the following quantity to decide if it is reasonable to suppose that the N discrete samples within a composite all have a concentration less than the threshold T:

$$\text{Composite compliance threshold} = \frac{T}{N}$$

REFERENCES AND FURTHER READING

In addition to the other guidance documents in this series, the following reference provides useful supplementary material on how the EPA views the issue of compositing:

Boomer, B.A., *Verification of PCB Spill Cleanup by Sampling and Analysis*, EPA-560/5-85-026, United States Environmental Protection Agency, 1985.