62

**Sampling Errors Associated With
Soil Composites Used to Estimate
Mean Ra-226 Concentrations at
an UMTRA Remedial Action Site**

**R. O. Gilbert
K. R. Baker
R. A. Nelson**

**R. H. Miller
M. L. Miller**

July 1987

**Prepared for the U.S. Department of Energy
under Contract DE-AC06-76RLO 1830**

**Pacific Northwest Laboratory
Operated for the U.S. Department of Energy
by Battelle Memorial Institute**

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government of any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or Battelle Memorial Institute.

PACIFIC NORTHWEST LABORATORY
operated by
BATTELLE MEMORIAL INSTITUTE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC06-76RLO 1830

Printed in the United States of America
Available from
National Technical Information Service
United States Department of Commerce
5285 Port Royal Road
Springfield, Virginia 22161

NTIS Price Codes
Microfiche A01

Printed Copy

Pages	Price Codes
001-025	A02
026-050	A03
051-075	A04
076-100	A05
101-125	A06
126-150	A07
151-175	A08
176-200	A09
201-225	A010
226-250	A011
251-275	A012
276-300	A013

PNL-6262
UC-70A

SAMPLING ERRORS ASSOCIATED WITH SOIL
COMPOSITES USED TO ESTIMATE MEAN Ra-226
CONCENTRATIONS AT AN UMTRA REMEDIAL
ACTION SITE

R. O. Gilbert (a)
K. R. Baker (b)
R. A. Nelson (b)
R. H. Miller (b)
M. L. Miller (a)

July 1987

Prepared for
U.S. Department of Energy
under Contract DE-AC06-76RLO 1830

Pacific Northwest Laboratory
Richland, Washington 99352

(a) Roy F. Weston, Inc., Albuquerque, NM.
(b) Jacobs Engineering Group, Inc., Albuquerque, NM.

SUMMARY

At the present time the decision whether to take additional remedial action (removal of soil) from regions contaminated by uranium mill tailings involves collecting 20 plugs of soil from each 10-m by 10-m plot in the region and analyzing a 500-g portion of the mixed soil for ^{226}Ra . Due to the high cost of collecting 20 soil plugs in each plot, a soil sampling study was conducted in the windblown mill-tailings flood plain area at Shiprock, New Mexico, to evaluate whether reducing the number of soil plugs to 9 would have any appreciable impact on remedial-action decisions. The results of the Shiprock study are described and used in this paper to develop a simple model of the standard deviation of ^{226}Ra measurements on composite samples formed from 21 or fewer plugs. This model is then used to predict as a function of the number of soil plugs per composite: (1) the percent accuracy with which the mean ^{226}Ra concentration in surface soil can be estimated, and (2) the probability of making incorrect remedial action decisions on the basis of statistical tests.

The developed model indicates that if the true plot mean (including background contributions) is 5 pCi/g, this mean can be estimated with about 30% accuracy at the 80% confidence level if the mean is estimated using one 21-plug composite sample measurement. This compares with about 50% and 65% accuracy if one 9- or 5-plug composite sample, respectively, is used to estimate the mean. More generally, regardless of the number of composite samples collected from the plot, and for any given confidence level, the model indicates that the percent accuracy of the estimated mean using 21-plug composite samples is multiplied by the factor $(21/p_1)^{1/2}$ to obtain the percent accuracy of the mean if p_1 -plug composite samples are used, where $p_1 < 21$. This factor equals 1.5 and 2.0 for 9- and 5-plug samples, respectively.

The magnitude of changes in the probability of making incorrect remedial action decisions due to changing the number of soil plugs per composite sample from 21 to a lesser number depends on the particular statistical test used to make the decision. Suppose additional remedial action will be done whenever the estimated plot mean above background is greater than the Environmental Protection Agency (EPA) limit of 5 pCi/g above background. Also, assume that the standard deviation of composite-sample ^{226}Ra concentrations is a known

constant as modeled using the Shiprock data. Then using one or more 9-plug rather than 21-plug composite samples increases the probability of making decision errors (incorrectly deciding additional remedial action is or is not needed) by no more than 10 probability points. If 5-plug samples are used, the increase is no more than about 17 probability points. These maximum increases are over relatively narrow bands of true plot means above background; between 2.5 and 4.5 pCi/g and between 6 and 13 pCi/g. These bands become smaller if more than one composite sample per plot is used to estimate the plot mean. If the plot mean is estimated using one or more 21- or 9-plug samples, the probability of incorrectly deciding additional remedial action is not needed is small (≤ 0.05) when the true plot mean above background exceeds about 15 pCi/g.

Suppose that in the interest of increased protection for the public, remedial action is taken whenever the upper 95% confidence limit on the mean concentration above background exceeds the EPA limit, which implies remedial action is taken even though the estimated plot mean is less than the limit. Then, assuming the standard deviation is known, using one or two 9-plug rather than 21-plug samples increases the probability of incorrectly deciding additional remedial action is needed from between about 15 to 30 probability points for plots with true means between about 1.5 and 4 pCi/g above background. If 5-plug samples are used, the increase is between about 20 and 50 probability points for true plot means in that same range. Regardless of whether 21-, 9-, or 5-plug samples are used, the probability of incorrectly deciding additional remedial action is not needed is small (≤ 0.05) when the true mean above background equals or exceeds the EPA limit of 5 pCi/g above background.

Suppose that in the interest of reducing remedial action sampling costs, remedial action is taken only when the lower 95% confidence limit on the mean concentration above background exceeds the EPA limit, which implies remedial action is taken only when the estimated mean is somewhat greater than the limit. Then, assuming the standard deviation is known, using one or more 9-plug rather than 21-plug samples increases the probability of incorrectly deciding additional remedial action is not needed by 10 to 30 probability points when the true plot mean above background is between about 6 and 10 (or

more) pCi/g. If 5-plug samples are used, the increase is between about 10 and 50 probability points. The probability of incorrectly deciding additional remedial action is needed is small (≤ 0.05) when the true plot mean above background is less than or equal to the EPA limit. If the lower 95% confidence limit test is used, the plot mean above background must exceed 25 pCi/g before that test using one 9-plug sample will have a probability exceeding 0.70 of indicating that additional remedial action is needed.

The probabilities of making remedial action decision errors reported in this paper are applicable when the standard deviation of composite samples is known with assurance. If the standard deviations are unknown (the usual case) then the probabilities will be greater than reported here, but the conclusions given above about the increase in the probabilities as the number of soil plugs per composite sample decreases remain valid on the whole.

ACKNOWLEDGMENTS

The authors are pleased to acknowledge the assistance of Dave Andrews and David Miller for their field work in collecting samples, of Robert Fencil for sample collection and preparation, of Jean-Claude Dehmel for performing preliminary statistical analyses of the data, and of Rick Bates and Lynn Williams for their technical review and helpful comments on a draft of this paper. We also thank Susan Cammann for typing the manuscript.

CONTENTS

SUMMARY	iii
ACKNOWLEDGMENTS	vii
1.0 INTRODUCTION	1
2.0 METHODS	5
2.1 Field Sampling Design	5
2.2 Rationale for Estimating Standard Deviations	7
3.0 RESULTS	11
3.1 Description of the Data	11
3.2 Estimating Changes in Standard Deviations	15
3.3 Modeling Changes in Standard Deviations	20
3.4 Percent Accuracy of Estimated Mean ²²⁶ Ra Concentrations	22
3.5 Probabilities of Remedial Action Decision Errors	26
4.0 STANDARD DEVIATIONS OF REPEAT MEASUREMENTS AND OF REPLICATE COMPOSITE SAMPLES	35
5.0 ACCEPTANCE SAMPLING FOR REDUCING THE NUMBER OF PLOTS REQUIRING SOIL SAMPLES	41
6.0 CONCLUSIONS	43
7.0 REFERENCES	45
APPENDIX A - ²²⁶ Ra Measurements for 5-, 8-, 9-, 16-, and 21-plug Composite Soil Samples Collected in Areas 1 Through 10 in the Windblown Mill-tailings Flood Plain Area at Shiprock, New Mexico, in January and March of 1986	A.1
APPENDIX B - "On the Application of Acceptance Sampling to the Decommissioning of Nuclear Facilities," by R. O. Gilbert et al. 1981	B.1

FIGURES

1	Location of the Ten 30-m by 30-m Areas in the Windblown Mill-tailings Flood Plain Region at Shiprock, New Mexico, Within Which Multiple-composite Soil Samples were Collected Following Initial Removal of Surface Soil	6
2	Sampling Patterns for 5-, 8-, 9-, 16-, and 21-plug Composite Soil Samples Collected from Ten 30-m by 30-m Areas in the Wind-blown Mill-tailings Flood Plain at Shiprock, New Mexico	8
3	Hypothetical Normal Distributions of ^{226}Ra Measurements of 20-plug Composite Samples (Distribution A) and 9-plug Composite Samples (Distribution B) when the True Mean for the Plot is 6 pCi/g Above Background	10
4	Hypothetical Normal Distributions of ^{226}Ra Measurements for 20-plug Composite Samples (Distribution A) and 9-plug Composite Samples (Distribution B) when the True Mean for the Plot is 4 pCi/g Above Background	10
5	^{226}Ra Measurements (pCi/g) of 5-, 8-, 9-, 16-, and 21-plug Composite Soil Samples Taken from Areas 8, 9, and 10 in the Windblown Mill-tailings Flood Plain at Shiprock, New Mexico. $\bar{x}$ and s are the Arithmetic Mean and the Standard Deviation of the n Measurements for each Data Set	12
6	^{226}Ra Measurements (pCi/g) of 5-, 8-, 9-, 16-, and 21-plug Composite Soil Samples Taken from Areas 2, 3, 5, and 6 in the Windblown Mill-tailings Flood Plain at Shiprock, New Mexico. $\bar{x}$ and s are the Arithmetic Mean and the Standard Deviation of the n Measurements for each Data Set	13
7	^{226}Ra Measurements (pCi/g) of 5-, 8-, 9-, 16-, and 21-plug Composite Soil Samples Taken from Areas 1, 4, and 7 in the Windblown Mill-tailings Flood Plain at Shiprock, New Mexico. $\bar{x}$ and s are the Arithmetic Mean and the Standard Deviation of the n Measurements for each Data Set	14
8	Standard Deviations of Multiple Composite Samples from Areas 1 through 10 at the Windblown Mill-tailings Flood Plain at Shiprock, New Mexico. Mean Ra-^{226} Concentrations for each Area are given to Illustrate that Areas with Lower Average Concentrations Tend to have Smaller and more Stable Standard Deviations	16
9	Plot of the Ratio s_5/s_{21} Versus the Mean ^{226}Ra Measurement (pCi/g) for each of the Ten 10-m by 10-m Study Plots at Shiprock, New Mexico. GM = Geometric Mean, GSE = Geometric Standard Error, AM = Arithmetic Mean, SE = Standard Error	18
10	Plot the Ratio s_9/s_{21} Versus the Mean ^{226}Ra Measurement (pCi/g) for each of the Ten 10-m by 10-m Study Plots at Shiprock, New Mexico	18

11	Plot of the Ratio s_5/s_9 Versus the Mean ^{226}Ra Measurement (pCi/g) for each of the Ten 10-m by 10-m Study Plots at Shiprock, New Mexico. GM = Geometric Mean, GSE = Geometric Standard Error, AM = Arithmetic Mean, and SE = Standard Error	19
12	Plot of the Ratio s_8/s_{16} Versus The Mean ^{226}Ra Measurement (pCi/g) for each of the Ten 30-m by 30-m Study Plots at Shiprock, New Mexico	19
13	Least-squares Linear Regression Lines Relating the Standard Deviation of Replicate Composite Samples from a Plot to the Estimated Mean Concentration of Ra-226 for the Plot	23
14	Probabilities of Taking Additional Remedial Action in a Plot for Three Decision Rules when one 500-g Sample from a Composite Sample Composed of Either 21, 9, or 5 Soil Plugs from the Plot is Measured for Ra-226. Probabilities are from Tables 3, 4, and 5	32
15	Probabilities that the 50% Confidence Rule will Indicate Additional Remedial Action is Needed if One, Two, or Three 500-g Samples from a Composite Sample Composed of 9 Soil Plugs are Measured for Ra-226. Probabilities are from Table 4	32

TABLES

1	Comparing Estimated and Predicted Ratios of Standard Deviations for Composite Samples Formed from Different Numbers of Soil Plugs	21
2	Estimated Percent Accuracy With Which the Mean ^{226}Ra Concentration, μ , in a 10-m by 10-m Plot can be Estimated with 95% or 80% Confidence Using 21-, 9-, or 5-plug Composite Samples	25
3	Probabilities that Decision Rule 1 (the Upper 95% Confidence Limit Rule) will Indicate Additional Remedial Action is Needed in a Plot	29
4	Probabilities that Decision Rule 2 (Comparing $\bar{x}'$ with 5 pCi/g Above Background) will Indicate Additional Remedial Action is Needed in a Plot	30
5	Probabilities that Decision Rule 3 (the Lower 95% Confidence limit rule) will Indicate Additional Remedial Action is Needed in a Plot	31
6	Repeat Measurements of ^{226}Ra Concentrations in Composite Surface Soil Samples with Measurements Greater than 2 pCi/g	36
7	Repeat Measurements of ^{226}Ra Concentrations in Composite Surface Samples with Measurements Less than 2 pCi/g	37
8	Measurements of ^{226}Ra Concentrations (pCi/g) on Replicate 500-g Composite Samples Taken from the Same Soil Mass	39
A1	^{226}Ra Measurements for 5-, 8-, 9-, 16-, and 21-plug Soil Composite Samples Collected in Areas 1 through 10 in the Windblown Mill-tailings Flood Plain Area at Shiprock, New Mexico	A.1

SAMPLING ERRORS ASSOCIATED WITH SOIL COMPOSITES USED TO ESTIMATE
MEAN Ra-226 CONCENTRATIONS AT AN UMTRA REMEDIAL-ACTION SITE

1.0 INTRODUCTION

The United States federal government is required under the Uranium Mill Tailings Radiation Control Act (U.S. Congress Public Law 95-604, 1978) to perform remedial actions on inactive uranium mill-tailings sites that had been federally supported and on properties that had been contaminated by the tailings. Remedial actions are required if the standards of the U.S. Environmental Protection Agency (EPA) for contamination of surface soil by ^{226}Ra are not met. The current EPA standard (EPA, 1983) requires that remedial action be taken if the average concentration of ^{226}Ra in soil over any area of 100 square meters exceeds the background level by more than 5 pCi/g for the first 15 cm of soil below the surface, or if the average exceeds 15 pCi/g for subsequent 15-cm thick layers of soil more than 15 cm below the surface. Since there are many thousands of 100 square-meter areas that must be evaluated, the soil sampling plan should be as economical as possible while still meeting the intent of the regulations.

After initial remedial action at a site has been conducted, it is necessary to determine whether the EPA standard has been met. The field sampling procedure presently being used for this verification effort is to first grid the entire site into 10-m by 10-m plots. Then, in each plot, 20 plugs of soil (each to 15-cm depth) are collected and physically mixed together from which a single 500-g composite sample is withdrawn and assayed for ^{226}Ra . This measurement, which is an estimate of the true mean ^{226}Ra concentration in surface soil for the plot, is compared with the EPA standard to determine if additional remedial action is required.

In this paper we are concerned with evaluating the effect of using fewer than 20 plugs per composite sample on (1) the percent accuracy with which the mean ^{226}Ra concentration for the plot can be estimated, and (2) the probabilities of making remedial action decision errors, i.e., of incorrectly deciding remedial action is required or of incorrectly deciding remedial action is not required. This paper also provides guidance on the degree that

the percent accuracy of estimated means is improved and the probabilities of making decision errors are decreased when more than one composite sample per plot is assayed. In addition, we compare the probabilities of doing additional remedial action that result from using the present decision rule (comparing the estimated plot mean above background with the EPA limit) with those when either an upper or lower 95% confidence limit is compared with the 5 pCi/g limit. We also briefly consider how acceptance sampling, which would involve collecting soil samples in only a subset of the 10-m by 10-m plots at a site, might be used in conjunction with in-situ gamma measurements to reduce costs associated with verifying compliance with the EPA limit.

The percent accuracy of estimated means, which is defined here as $100(\text{estimated mean} - \text{true mean})/\text{true mean}$, and the probabilities of making decision errors are directly related to the variability of ^{226}Ra measurements. Hence, our first task is to estimate and then model this variability. To do this, a soil sampling study was conducted in the windblown mill-tailings flood plain area at Shiprock, New Mexico after the site had undergone an initial remedial action operation. Measurements for ^{226}Ra concentrations were made for multiple surface-soil composite samples collected from 30-m by 30-m areas and from within the central 10-m by 10-m plot within each of these areas. Measurements for ^{226}Ra were made on composite samples of different types, namely, those formed by mixing either 5, 8, 9, 16, or 21 plugs of soil.

The field sampling design used for this study is described in Section 2.1 and the ^{226}Ra data are listed in Appendix A. Section 2.2 illustrates how the variability between composite-sample measurements affects the probabilities of making remedial-action decision errors. In Section 3.1 the ^{226}Ra measurements for the composite samples obtained at the Shiprock site are graphically displayed and summarized. The data in Section 3.1 are used in Section 3.2 to estimate the average increase in standard deviation that occurs as the number of plugs per composite sample decreases. These estimates are then compared with predicted increases obtained using a simple statistical model.

In Section 3.3 a model is developed for predicting the standard deviation of ^{226}Ra measurements on composite samples formed from fewer than 21 plugs. This model is used in Sections 3.4 and 3.5 to predict percent accuracies with which plots are estimated and the probabilities of making remedial-action

decision errors, respectively. In Section 3.4 the results obtained from the Shiprock data are also compared with those reported by Williams et al. (1986) for 10- and 20-plug soil composite samples taken from five properties in Grand Junction, Colorado. In Section 4 we compute the variability of replicate composite samples and of repeat measurements on the same composite sample to obtain a lower bound for the variability between replicate composite soil samples. Acceptance sampling is introduced in Section 5 and Appendix B, and concluding remarks are given in Section 6. References are given in Section 7.

2.0 METHODS

2.1 Field Sampling Design

Figure 1 shows the windblown mill-tailings flood plain region and the location of the ten 30-m by 30-m study areas from which soil samples were collected. Areas 1, 2, 3, 4, and 5 were selected and sampled in January 1986 after initial remedial action (removal of soil) in the region had occurred. One or more of the 10-m by 10-m plots in each of these five selected areas except Area 2 were believed on the basis of previously collected 20-plug composite surface (0 to 15-cm) soil samples to have ^{226}Ra concentrations exceeding 6 pCi/g. These higher-concentration plots are indicated in Fig. 1. The remaining 5 areas (those numbered 6, 7, 8, 9, and 10) were selected and sampled for this study in March 1986 after additional remedial action in the flood plain region had occurred. Hence, these areas were expected to have ^{226}Ra concentrations below 5 pCi/g.

Eight-plug and 16-plug composite samples were formed by pooling soil plugs that were collected over the ten 30-m by 30-m areas according to the three sampling patterns shown in the lower half of Fig. 2. The 5-, 9-, and 21-plug composite samples were formed by pooling soil plugs collected from only the central 10-m by 10-m plot in each 30-m by 30-m area using the three patterns shown in the upper half of Fig. 2. In this paper, emphasis is placed on the results for the 5-, 9-, and 21-plug composite samples since they were formed from soil plugs removed from plots of the size specified by the EPA (EPA, 1983).

Up to nine composite samples of each type were formed in each of the ten areas. Each composite sample of a given type used the same pattern that had been shifted slightly in location. For example, referring to Fig. 2, the 21-plug composite sample number 1 in a given 10-m by 10-m plot was formed by pooling soil plugs collected at the 21 positions numbered 1 in the plot. Similarly, the 21-plug composite number 3 was formed by pooling soil plugs at the 21 positions numbered 3 in the plot. This sampling design allowed replicate composite samples of a given type to be collected without altering the basic pattern that would be used in practice.

FIGURE 1. Location of the Ten 30-m by 30-m Areas in the Windblown Mill-tailings Flood Plain Region at Shiprock, New Mexico, Within Which Multiple-composite Soil Samples were Collected Following Initial Removal of Surface Soil.

We note that due to a possible positive correlation between ^{226}Ra measurements in replicate composite soil samples based on plugs in close proximity, the variance between replicate composite samples may be smaller than if each composite sample had been formed from soil plugs collected at random locations in the plot. Nevertheless the variance of interest here is that which applies to the sampling pattern used in practice.

Each soil plug was collected to a depth of 15 cm using a garden trowel, and all plugs were about equal in volume. The plugs collected for a given composite sample were placed in a bucket and mixed vigorously by stirring and shaking. The composite sample analyzed for ^{226}Ra consisted of about 500 g of the mixed soil that had been removed from the bucket and placed in a sample can. The larger pebbles, rocks, and stones were removed as the soil was transferred from the bucket to the can. The composite samples in the cans were dried at 100°C for 24 h and then sealed air-tight. The ^{226}Ra concentration (pCi/g) for the sample was then obtained by the Remedial Action Contractor (RAC) after a 20-to 30-day waiting period. These data are listed in Appendix A.

For a subset of the 500-g composite samples, a repeat measurement of the ^{226}Ra concentration was obtained by the RAC. Each repeat measurement was made by the same individual on the same counting system, but not necessarily on the same day as the original measurement. Most of these samples were also measured once at a later time by the Technical Assistance Contractor (TAC) to assess the magnitude of differences in measurements that occur due to a different person using a different counting system on a different day. In some cases, replicate 500-g composite samples were obtained by filling more than one sample can with the mixed soil. The repeat-measurement and replicate-composite-sample data are listed in Tables 6, 7, and 8 and are discussed in Section 4.

2.2 Rationale For Estimating Standard Deviations

To determine the effect on the probability of making decision errors by using fewer than 20 plugs of soil per composite sample or by using more than 1 composite sample per plot, it is necessary to estimate the standard deviation, σ , of ^{226}Ra measurements made on composite samples from fewer than 20 plugs.

FIGURE 2. Sampling Patterns for 5-, 8-, 9-, 16-, and 21-plug Composite Soil Samples Collected From Ten 30-m by 30-m Areas in the Windblown Mill-tailings Flood Plain at Shiprock, New Mexico.

The standard deviation must also be known to assess the percent accuracy with which the mean ^{226}Ra concentration for a plot can be estimated.

If σ increases as the number of plugs per composite sample decreases, then the uncertainty in the estimated mean concentration for the plot also increases, as does the probability of making decision errors. This is illustrated in Fig. 3 where two hypothetical normal (Gaussian) distributions of ^{226}Ra measurements from a plot are shown, where both distributions have a mean ^{226}Ra concentration of 6 pCi/g above background. The distribution labeled A represents the distribution of measurements that might be obtained if a great many 20-plug composite samples were collected from the plot and a single measurement was made on each. Distribution B represents the hypothetical distribution of measurements from a great many 9-plug composite samples collected from the plot. We have made Distribution B wider than Distribution A because the standard deviation of ^{226}Ra measurements should be greater for 9-plug than for 20-plug samples. Figure 3 shows that the proportion of measurements less than 5 pCi/g above background is greater for 9-plug than for 20-plug samples. That is, the probability of incorrectly deciding that remedial action is not needed for the plot is greater for 9- than for 20-plug samples.

Figure 4 shows the same two distributions as in Fig. 3, but this time their common mean is 4 pCi/g above background, which is below the EPA standard of 5 pCi/g above background. The proportion of measurements greater than 5 pCi/g is now greater for the 9-plug than for the 20-plug samples. Hence, the probability of incorrectly deciding to take remedial action on this plot is greater for 9-plug than for 20-plug samples.

Figures 3 and 4 illustrate the importance of determining the extent that σ is affected by the number of soil plugs used to form composite samples. These figures also illustrate that the size of σ will have no effect on remedial action decisions if both distributions of measurements from the plot are located entirely below or entirely above the EPA limit, since in that case, neither distribution will overlap the limit. In other words, the number of plugs per composite is not a crucial concern if ^{226}Ra concentrations above background are much larger or smaller than 5 pCi/g above background.

FIGURE 3. Hypothetical Normal Distributions of ^{226}Ra Measurements of 20-plug Composite Samples (Distribution A) and 9-plug Composite Samples (Distribution B) When the True Mean for the Plot is 6 pCi/g Above Background.

FIGURE 4. Hypothetical Normal Distributions of ^{226}Ra Measurements for 20-plug Composite Samples (Distribution A) and 9-plug Composite Samples (Distribution B) When the True Mean for the Plot is 4 pCi/g Above Background.

3.0 RESULTS

3.1 Description of the Data

The measurements of ^{226}Ra obtained for the 5-, 8-, 9-, 16-, and 21-plug 500-g composite samples from the Shiprock site are listed in Appendix A and plotted in Figs. 5, 6, and 7. These figures show the variation between replicate composite-sample measurements for each type of composite. They also show how the variability (scatter) of the data differs between the 10 areas that were sampled. Also given in the figures are the arithmetic mean, $\bar{x}$, and the standard deviation, s , for each data set, where $\bar{x}$ and s were computed as follows:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

$$s = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{1/2}, \quad (2)$$

where n is the number of measurements in the data set, and x_i is the ^{226}Ra measurement (including background) for the i^{th} 500-g sample.

As explained in Section 2.2, we wish to determine the extent to which the true standard deviation, σ , increases when fewer than 20 plugs are used to form a composite sample. To avoid confusion in making that determination, we digress momentarily to point out a feature of the data in Figs. 6 and 7 that appears at first glance to be anomalous. Specifically, Figs. 6 and 7 show that the ^{226}Ra measurements for most 5-, 9-, and 21-plug composite samples from Areas 1, 3, and 4 are larger than the ^{226}Ra measurements for the 8- and 16-plug composite samples in those areas. This is believed to have occurred because the soil in the central 10-m by 10-m plot (from which 5-, 9-, and 21-plug composite samples were formed) had higher concentrations of ^{226}Ra than the soil in the 30-m by 30-m areas from which the 8- and 16-plug composite samples were formed (see Fig. 1). Hence, the lower measured concentrations for 8- and 16-plug composite samples do not appear to be due to laboratory, transcription, or other types of errors.

FIGURE 5. ^{226}Ra Measurements (pCi/g) of 5-, 8-, 9-, 16-, and 21-plug Composite Soil Samples Taken from Areas 8, 9, and 10 in the Windblown Mill-tailings Flood Plain at Shiprock, New Mexico. $\bar{x}$ and s are the Arithmetic Mean and Standard Deviation of the n Measurements for each Data Set.

FIGURE 6. ^{226}Ra Measurements (pCi/g) of 5-, 8-, 9-, 16-, and 21-plug Composite Soil Samples Taken from Areas 2, 3, 5, and 6 in the Windblown Mill-tailings Flood Plain at Shiprock, New Mexico. $\bar{x}$ and s are the Arithmetic Mean and Standard Deviation of the n Measurements for each Data Set.

FIGURE 7. ^{226}Ra Measurements (pCi/g) of 5-, 8-, 9-, 16-, and 21-plug Composite Soil Samples Taken from Areas 1, 4, and 7 in the Windblown Mill-tailings Flood Plain at Shiprock, New Mexico. $\bar{x}$ and s are the Arithmetic Mean and Standard Deviation of the n Measurements for each Data Set.

We return now to the main question of interest; the relationship between the standard deviation and the number of soil plugs per composite sample. All measurements for Areas 8, 9, and 10 (Fig. 5) were below 5 pCi/g and the estimated standard deviations, s , ranged from 0.2 to 0.8 pCi/g, with no apparent systematic changes in s with changes in the number of plugs per composite sample. In Fig. 6 we see that for Areas 3, 5, and 6, s was larger for 5-plug than for 21-plug composites, and for Areas 2, 3, and 5, s was larger for 9-plug than for 21-plug composites. The ^{226}Ra measurements for Areas 1, 4, and 7 (Fig. 7) were higher and more variable than those for the areas in Figs. 5 and 6. For Areas 4 and 7, s was larger for both 5- and 9-plug composite samples than for 21-plug samples. On the other hand, for Area 1, s for 5- and 9-plug composite samples was smaller than for the 21-plug samples.

When making these types of comparisons, one should keep in mind that the computed value of s for a plot is only an estimate of σ , the true standard deviation for the plot, and that relatively few samples (at most 9) are available here for computing s . Hence, some of the variability in the values of s given in the figures is due to this estimation process.

Figure 8 plots the values of s for each of the ten areas to show more clearly the changes in s that occurred as the number of plugs per composite were changed. This plot shows that, at least for the higher concentration plots, the values of s tend to decline as the number of plugs per composite sample increase. For four of the six areas with average ^{226}Ra measurements less than 5 pCi/g, the standard deviations were equal or smaller for 9-plug than for 21-plug composite samples. For three of the four areas with average ^{226}Ra measurements exceeding 5 pCi/g, the standard deviations for 9-plug composite samples were higher than for 21-plug samples. Similar statements apply to the comparison of 5- and 21- and of 8- and 16-plug composite samples.

3.2 Estimating Changes in Standard Deviations

In this section we first estimate the changes in σ that occur as the number of plugs per composite sample decreases from 21 to a smaller number. Then a model for these changes is developed for use in later sections.

FIGURE 8. Standard Deviations of Multiple Composite Samples from Areas 1 Through 10 at the Windblown Mill-tailings Flood Plain at Shiprock, New Mexico. Mean Ra-226 Concentrations for each Area are Given to Illustrate that Areas with Lower Average Concentrations tend to have Smaller and More Stable Standard Deviations.

Plots of the ratios s_5/s_{21} , s_9/s_{21} , s_5/s_9 , and s_8/s_{16} versus average soil measurements are given in Figs. 9 through 12, respectively. (s_5 denotes the estimated standard deviation of 5-plug composite samples, with similar definitions for s_8 , s_9 , s_{16} , and s_{21} .) The ratios themselves are listed next to the plots in the figures. To estimate the average increase in σ as the number of plugs decreased, we computed the geometric mean (GM) and the arithmetic mean (AM) of the ratios listed in Figs. 9 through 12. The geometric standard error (GSE) and standard error (SE) of the ratios are also given in Figs. 9 through 12 to indicate the uncertainty in the GM and AM, respectively, due to having data from only a limited number of areas. The AM was computed using Eq. (1); the SE is $s/\sqrt{n}$, where s (the standard deviation of the n ratios) was computed using Eq. (2); the GM and GSE were computed as $\exp(\bar{y})$ and $\exp(s_y/\sqrt{n})$, respectively, where $\bar{y}$ and s_y are the arithmetic mean and standard deviation of the natural logarithms of the ratios. The GM was computed since ratios tend to be lognormally distributed. In that situation the GM is a good measure of the mid-point (median) of the distribution, while the AM, which is always larger than the GM for the right-skewed data obtained here, estimates the true mean of the distribution.

The AMs in Figs. 9 through 12 suggest that σ may increase by about 40 to 50% if the number of plugs is reduced from 21 to 9 (Fig. 10), from 9 to 5 (Fig. 11), or from 16 to 8 (Fig. 12). Also, the AM in Fig. 9 suggests that σ increases about 90% when the number of plugs is reduced from 21 to 5. The GMs and AMs of the ratios were also computed after excluding the ratios for Areas 9 and 10. These averages are shown in columns 2 and 4 of Table 1. Areas 9 and 10 had very low ^{226}Ra measurements, and for that reason, changes in σ for these plots may be less than what occurs for plots with concentrations closer to the EPA limit. The GMs in Table 1 suggest that σ increases by 30 or 40% if the number of plugs per composite sample is reduced by about half. The averages in Table 1 are somewhat higher than those in Figs. 9 through 12 since the ratios for Areas 9 and 10 were usually less than one.

FIGURE 9. Plot of the Ratio s_5/s_{21} Versus the Mean ^{226}Ra Measurement (pCi/g) for each of the Ten 10-m by 10-m Study Plots at Shiprock, New Mexico. GM = Geometric Mean, GSE = Geometric Standard Error, AM = Arithmetic Mean, SE = Standard Error.

FIGURE 10. Plot of the Ratio s_9/s_{21} Versus the Mean ^{226}Ra Measurement (pCi/g) for each of the ten 10-m by 10-m Study Plots at Shiprock, New Mexico.

FIGURE 11. Plot of the Ratio s_5/s_9 Versus the Mean ^{226}Ra Measurement (pCi/g) for each of the ten 10-m by 10-m Study Plots at Shiprock, New Mexico. GM = Geometric Mean, GSE = Geometric Standard Error, AM = Arithmetic Mean, and SE = Standard Error.

FIGURE 12. Plot of the Ratio s_8/s_{16} Versus the Mean ^{226}Ra Measurement (pCi/g) for each of the ten 30-m by 30-m Study Plots at Shiprock, New Mexico.

3.3 Modeling Changes in Standard Deviations

In this section we develop a simple model for predicting the standard deviation of composite samples formed from fewer than 21 plugs. This model is used in Section 3.4 to approximate the percent accuracy with which mean ^{226}Ra concentrations can be estimated. The model is also used in Section 3.5 to compute the probabilities that additional remedial action is required in 10-m by 10-m plots when different statistical decision rules (tests) are used.

Let us assume that measurements of ^{226}Ra in individual soil plugs are uncorrelated and that the soil plugs are thoroughly mixed together in the bucket before the composite sample is removed. Then, the ratio of true standard deviations for composite samples formed from p_1 and p_2 plugs ($p_1 < p_2$) is

$$\sigma_{p_1}/\sigma_{p_2} = \frac{\sigma'}{\sqrt{p_1}} / \frac{\sigma'}{\sqrt{p_2}} = (p_2/p_1)^{1/2}, \quad (3)$$

where σ' is the standard deviation of ^{226}Ra measurements for individual soil plugs. (σ' cancels out of the equation and we make no use of it here.)

Using Eq. (3) when $p_1 = 9$ and $p_2 = 21$ gives $(21/9)^{1/2} = 1.53$, a 53% increase in σ if 9 rather than 21 plugs are used. This and other values of Eq. (3) are given in column 6 of Table 1 for values of p_1 and p_2 of interest in this study. We see that the estimated geometric means and arithmetic means of the ratios (columns 2 and 4 of Table 1) are in reasonably good agreement with the values obtained using Eq. (3). The model predicts that the standard deviation will increase by a factor of 2 (100% increase) if 5 rather than 21 plugs are used. This prediction is in good agreement with the arithmetic-mean ratios 1.9 and 2.2 in Fig. 9 and Table 1, respectively. Equation (3) predicts that reducing the number of plugs from 20 to 9 results in a 49% increase in the standard deviation.

Solving Eq. (3) for σ_{p_1} gives

$$\sigma_{p_1} = \sigma_{p_2} (p_2/p_1)^{1/2}. \quad (4)$$

TABLE 1. Comparing Estimated and Predicted Ratios of Standard Deviations for Composite Samples Formed From Different Numbers of Soil Plugs.

Ratio of Standard Deviations	Estimated Ratios ⁺ Computed Using Data from Areas 1 through 8				Predicted** Ratios Computed Using Equation 3
	Geometric Mean (GM)	Geometric Standard Error (GSE)	Arithmetic Mean (AM)	Standard Error (SE)	
σ_9/σ_{21} *	1.3	1.3	1.6	0.3	1.53
σ_5/σ_{21}	1.7	1.3	2.2	0.7	2.05
σ_5/σ_9	1.3	1.2	1.5	0.3	1.34
σ_8/σ_{16}	1.4	1.3	1.7	0.5	1.41

* σ_j = true standard deviation of j-plug composite samples.

** Computed as $(p_2/p_1)^{1/2}$, where p_1 and p_2 are the smaller and larger number of soil plugs per composite sample, respectively.

+ Ratios for Areas 1 through 8 at Shiprock were used. Areas 9 and 10 were excluded because of their very low ²²⁶Ra measurements.

This equation is used here to predict the standard deviation for p_1 -plug composite samples using the standard deviation for p_2 -plug composite samples (σ_{p_2}), where $p_2 = 21$ and $p_1 < 21$.

The model used for σ_{p_2} was

$$\sigma_{p_2} = 0.10 + 0.23\mu, \quad (5)$$

where μ is the true mean ^{226}Ra concentration (including background) for the plot. This model implies that the standard deviation of 21-plug composite samples increases as the mean ^{226}Ra concentration increases. Evidence for this is given in Fig. 13 where we have plotted, for each of the 10 areas at Shiprock, the value of s versus the mean ^{226}Ra measurement for composite samples formed from 5, 9, and 21 plugs of soil. Least-squares linear regression lines were fit to the three sets of data. These lines and the associated correlations between s and $\bar{x}$ are shown in Fig. 13. Note that the line for 21-plug composite samples lies below the 9-plug line, which in turn lies below the line for 5-plug samples. These results are expected if σ increases as the number of plugs per composite decreases. The least-squares line for the 21-plug composite samples given in Fig. 13 is the basis for the model in Eq. (5).

Substituting Eq. (5) in Eq. (4) gives

$$\sigma_{p_1} = (0.10 + 0.23\mu)(p_2/p_1)^{1/2}, \quad (6)$$

which is the model used here to predict the standard deviation of p_1 -plug composite samples, where $p_1 < 21$.

3.4. Percent Accuracy of Estimated Mean ^{226}Ra Concentrations

Using Eq. (6) and assuming that ^{226}Ra measurements of composite samples are normally distributed, the following formula was used to estimate the

r = estimated correlation between s and $\bar{x}$

FIGURE 13. Least-squares Linear Regression Lines Relating the Standard Deviation of Replicate Composite Samples from a Plot to the Estimated Mean Concentration of Ra-226 for the Plot.

percent accuracy with which the post-remedial-action mean ^{226}Ra concentration for a plot at Shiprock would be estimated with specified confidence:

$$\text{Percent Accuracy} = 100 Z (0.10 + 0.23\mu)(p_2/p_1)^{1/2}/(\mu\sqrt{n}) , \quad (7)$$

where Z equals 1.96 or 1.28 if 95% or 80% confidence, respectively, is required, n is the number of p_1 -plug composite samples collected in the plot and averaged together to estimate the plot mean, and μ is the true plot mean. Equation 7 is based on the usual formula for estimating the number of samples required to estimate a mean with prespecified relative accuracy and confidence; see, e.g., Gilbert (1987, p. 33). Note that Eq. (7) gives the same result whether there are n p_1 -plug composites or one np_1 -plug composite.

By dividing Eq. (7) when $p_2 = 21$ and $p_1 < 21$ by Eq. (7) when $p_2 = p_1 = 21$ we obtain $(21/p_1)^{1/2}$, which is the factor by which the percent accuracy of 21-plug composite samples is multiplied to get the percent accuracy of p_1 -plug samples. This formula gives 1.5 and 2.0 when $p_1 = 9$ and 5, respectively. Notice that this factor is not site dependent since it does not depend on μ or σ . Table 2 gives values of Eq. (7) for 95% and 80% confidence, $p_1 = 21, 9$, and 5 plugs, $n = 1$ and 2 composite samples per plot, and for μ ranging from 1 to 10 pCi/g.

The results in Table 2 for 80% confidence when one composite sample per plot is used may be compared with the results of a study reported by Williams et al. (1986). That study included collecting multiple 20- and 10-plug composite samples on five properties in Grand Junction, Colorado. These properties were known to be contaminated above the EPA limit, but they had not as yet undergone remedial action. Their data indicated that using either 20- or 10-plug composite samples gave about 80% confidence that the estimated mean for a property would fall within plus or minus 30% of the true mean for the property (on the average over the five properties). This result agrees with the results in column 9 of Table 2 for 21-plug composite samples. However, the data in column 10 of Table 2 suggest that the percent accuracy associated with 9-plug composite samples collected at Shiprock was about 50% rather than the 30% obtained by Williams et al. (1986) for the Grand Junction properties.

Table 2. Estimated percent accuracy with which the mean ^{226}Ra concentration, μ , in a 10-m by 10-m plot can be estimated with 95% or 80% confidence using 21-, 9-, or 5-plug composite samples.

Mean Concentration Including Background μ (pCi/g)	σ_{21}^+ (pCi/g)	95% Confidence						80% Confidence					
		Number of Composite Samples per Plot (n)						Number of Plugs per Composite Sample					
		1			2			1			2		
		21	9	5	21	9	5	21	9	5	21	9	5
1	0.33	65	99	133	46	70	94	42	65	87	30	46	61
1.5	0.44	58	89	119	41	63	84	38	58	78	27	41	55
2	0.56	55	84	112	39	60	80	36	55	74	25	39	52
3	0.79	52	79	106	36	56	75	34	51	69	24	36	49
4	1.02	50	76	102	35	54	72	33	50	67	23	35	47
5	1.25	49	75	100	35	53	71	32	49	66	23	35	46
8	1.94	48	73	97	34	51++	69	31	47	64	22	34	45
10	2.40	47	72	96	33	51	68	31	47	63	22	33	45

+ Computed using Eq. 5 and μ in column 1.

++ Interpreted as follows: If two 9-plug composite samples are obtained from a 10-m by 10-m plot that has a true mean concentration of 8 pCi/g including background, then we can be 95% confident that the arithmetic mean of the two measurements will fall within 51% of the true mean, i.e., between about 4 and 12 pCi/g.

This discrepancy may be due in part to the fact that for two of the five properties at Grand Junction, the estimates of variability for the 10-plug composite samples may have been biased low because a substantial portion of those two properties were not sampled by the 10-plug method (Williams et al. 1986). Also, at Grand Junction, a smaller size drill bit was used to collect 10-plug composite samples than the 20-plug samples. Use of the smaller drill bit may have introduced additional variability into the 10-plug sample data. Another possible explanation is that Eq. (5) does not apply to the sampled properties at Grand Junction. However, the means and standard deviations in Table 1 of Williams et al. (1986) suggest Eq. (5) is not an unreasonable model for their data. A fourth possible influencing factor is that either the Shiprock or Grand Junction ^{226}Ra measurements on composite samples are not normally distributed. Williams et al. (1986) reported that Chi-square tests did not indicate their data were not normally distributed. However, Chi-square tests are not very powerful for detecting non-normality. We have conducted the W test (discussed by Gilbert, 1987) for normality and lognormality on the data sets in Figs. 5, 6, and 7, and found that the 21-plug composite samples were more likely to be normally distributed than the 9- or 5- plug composite samples. But the data are inadequate to accurately quantify the effects of possible non-normality on the percent accuracies in Table 2 or those in Table 1 of Williams et al. (1986). Finally, we note that the amount of data from either study is not large, and there is large variability in the ratios of standard deviations (Figs. 9, 10, 11, and 12) and about the regression lines of s on $\bar{x}$ (Fig. 13). Hence, the model given by Eq. (7) is not free from uncertainty.

3.5 Probabilities of Remedial Action Decision Errors

In Table 2 we saw how the percent accuracy of the estimated mean ^{226}Ra concentration declines when the number of plugs per composite sample decreases and when one rather than two composite samples per plot are collected. A consequence of declining accuracy in the estimated mean is an increase in the occurrence of remedial-action decision errors when the true mean concentration for the plot is close to (above or below) the EPA limit. In this section this increase in decision errors is quantified by computing the probability of

making these decision errors. This information is useful when evaluating costs and benefits of using different numbers of plugs per composite samples and composite samples per plot.

The results in this section are obtained under the following two assumptions: (1) that Eq. (6) is an appropriate model for the variance of p_1 -plug composite samples ($p_1 < 21$), and (2) the estimated ^{226}Ra mean concentration for a plot based on p_1 -plug composite samples is normally distributed.

The probabilities of making remedial action decision errors are computed for three different decision rules:

Decision Rule 1

Take additional remedial action if $\bar{x}' + 1.645 \sigma_{p_1} / \sqrt{n}$ (the upper 95% confidence limit on the true plot mean) exceeds 5 pCi/g above background, where $\bar{x}'$ is the estimated mean concentration (above background) for the plot based on n p_1 -plug composite samples.

Decision Rule 2

Take additional remedial action if $\bar{x}'$ exceeds 5 pCi/g above background.

Decision Rule 3

Take additional remedial action if $\bar{x}' - 1.645 \sigma_{p_1} / \sqrt{n}$ (the lower 95% confidence limit on the true plot mean) exceeds 5 pCi/g above background.

It will be seen below that rule 1 offers the greatest protection to the public because the probabilities of taking additional remedial action are higher than for rules 2 or 3. Rule 3 will result in fewer decisions to take remedial action than rules 1 or 2 for plots with true mean ^{226}Ra concentrations near 5 pCi/g above background. Hence, rule 3 will tend to reduce costs of remedial action. Rule 2 is a compromise strategy in that the probabilities of taking remedial action fall between those for rules 1 and 3.

Let us define β to be the probability that a statistical test will indicate additional remedial action is needed. When decision rule 1 is used the probability β is obtained in two steps:

STEP 1. Compute the quantity Z_β as follows:

$$Z_\beta = \frac{(5 - \mu')(np_1/21)^{1/2}}{\sigma_{21}} - 1.645 \quad , \quad (8)$$

where 5 is the EPA limit, μ' is the true plot mean above background, σ_{21} is the standard deviation of 21-plug composite samples given by Eq. (6), p_1 is the number of soil plugs used to form each of the n composite samples from the plot, and 1.645 is the value from a standard normal distribution that cuts off the upper 5% of the distribution. Using the value 1.645 implies that the probability of incorrectly deciding additional remedial action is not needed is never greater than 0.05.

STEP 2. Refer Z_β to tables of the cumulative normal distribution to determine β .

For Decision Rule 2, the same procedure is used to obtain β except that Eq. (8) is computed with the constant 1.645 replaced by zero.

For Decision Rule 3, Eq. (8) is replaced by

$$Z_{1-\beta} = \frac{(\mu' - 5)(np_1/21)^{1/2}}{\sigma_{21}} - 1.645 \quad . \quad (9)$$

Then β is obtained by referring $Z_{1-\beta}$ to the cumulative normal distribution tables to find $1 - \beta$ and then solving for β . Equations 8 and 9 were derived from the formula given by Burr (1976, p. 328). The general approach used above for finding β is given by Wine (1964, pp. 215-221).

We have computed β for various values of μ' when the background ^{226}Ra concentration was assumed to be 1 pCi/g (the approximate background value for the windblown flood plain at the Shiprock site) when $n = 1, 2$, or 3 , and $p_1 = 5, 9$, or 21 . The results are given in Tables 3, 4, and 5. The interpretation of these probabilities is given in a footnote to Table 3. The results when n equals 1 are plotted in Fig. 14, and the results for one, two, or three 9-plug composite samples are plotted in Fig. 15.

The probabilities in these tables and figures indicate that:

TABLE 3. Probabilities That Decision Rule 1 (the upper 95% confidence limit rule) Will Indicate Additional Remedial Action Is Needed in a Plot

Mean Concentration Above Background (μ^+)	Mean Concentration Including Background (μ)	σ_{21}^+	Number of Composite Samples per Plot (n)								
			1			2			3		
			Number of Plugs per Composite Sample (p_i)								
			21	9	5	21	9	5	21	9	5
0	1	0.33	0	0	0	0	0	0	0	0	0
1	2	0.56	0	0	0.03	0	0	0	0	0	0
2	3	0.79	0.02	0.20	0.41	0	0.03	0.17	0	0	0.06
2.5	3.5	0.90	0.13	0.44	0.62	0.01	0.18	0.40	0	0.07	0.24
3	4	1.02	0.37	0.64	0.75	0.13	0.43	0.61	0.04	0.28	0.55
3.5	4.5	1.14	0.63	0.78	0.84	0.41	0.67	0.77	0.26	0.56	0.70
4	5	1.25	0.80	0.87 ⁺⁺	0.89	0.69	0.82	0.86	0.60	0.77	0.83
4.5	5.5	1.36	0.90	0.92	0.93	0.87	0.90	0.92	0.84	0.89	0.91
4.75	5.75	1.42	0.93	0.94	0.94	0.92	0.93	0.94	0.91	0.93	0.93
5	6	1.48	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
6	7	1.71	0.99	0.98	0.97	0.99	0.99	0.98	0.99	0.99	0.98
7	8	1.94	1.0	0.99	0.98	0.99	0.99	0.99	0.99	0.99	0.99
8	9	2.17	1.0	0.99	0.99	1.0	1.0	0.99	1.0	1.0	0.99

+Computed using Eq. (5) and μ from Column 2.

++Interpreted as follows: If the true mean above background is 4 pCi/g, if the standard deviation of 9-plug composite samples from the plot is $1.25 (21/9)^{1/2} = 1.91$ pCi/g, and if the ^{226}Ra measurement of one 9-plug composite sample is used to estimate the mean of the plot, then the probability is 0.87 that the upper 95% confidence limit rule will indicate additional remedial action is needed.

TABLE 4. Probabilities That Decision Rule 2 (comparing $\bar{x}'$ with 5 pCi/g above background) Will Indicate Additional Remedial Action Is Needed in a Plot

Mean Concentration Above Background (μ')	Mean Concentration Including Background (μ)	σ_{21}^+	Number of Composite Samples per Plot (n)								
			1			2			3		
			Number of Plugs per Composite Sample (p_1)			Number of Plugs per Composite Sample (p_1)			Number of Plugs per Composite Sample (p_1)		
			21	9	5	21	9	5	21	9	5
0	1	0.33	0	0	0	0	0	0	0	0	0
1	2	0.56	0	0	0	0	0	0	0	0	0
2	3	0.79	0	0.01	0.03	0	0	0.004	0	0	0
2.5	3.5	0.90	0	0.04	0.09	0	0.005	0.03	0	0	0.01
3	4	1.02	0.02	0.10	0.17	0.003	0.03	0.09	0	0.01	0.05
3.5	4.5	1.14	0.10	0.19	0.26	0.03	0.11	0.18	0.01	0.07	0.13
4	5	1.25	0.21	0.30	0.35	0.13	0.23	0.29	0.08	0.18	0.25
4.5	5.5	1.36	0.36	0.40	0.43	0.30	0.37	0.40	0.26	0.34	0.38
4.75	5.75	1.42	0.43	0.45	0.47	0.40	0.44	0.45	0.38	0.42	0.44
5	6	1.48	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
6	7	1.71	0.72	0.65	0.61	0.80	0.70	0.66	0.84	0.75	0.69
7	8	1.94	0.85	0.75	0.69	0.93	0.83	0.76	0.96	0.88	0.81
8	9	2.17	0.92	0.82	0.75	0.98	0.90	0.83	0.99	0.94	0.88
9	10	2.40	0.95	0.86	0.79	0.99	0.94	0.87	1.0	0.97	0.92
10	11	2.63	0.97	0.89	0.82	1.0	0.96	0.90	1.0	0.98	0.95
11	12	2.86	0.98	0.91	0.85	1.0	0.97	0.93	1.0	0.99	0.96
12	13	3.09	0.99	0.93	0.87	1.0	0.98	0.94	1.0	1.0	0.97
13	14	3.32	0.99	0.94	0.88	1.0	0.99	0.95	1.0	1.0	0.98
14	15	3.55	1.0	0.95	0.89	1.0	0.99	0.96	1.0	1.0	0.98
15	16	3.78	1.0	0.96	0.90	1.0	1.0	0.97	1.0	1.0	0.99
16	17	4.01	1.0	0.96	0.91	1.0	1.0	0.98	1.0	1.0	0.99

+Computed using Eq. (5) and μ from Column 2.

TABLE 5. Probabilities That Decision Rule 3 (the lower 95% confidence limit rule) Will Indicate Additional Remedial Action Is Needed in a Plot

Mean Concentration Above Background (μ')	Mean Concentration Including Background (μ)	σ_{21}^+	Number of Composite Samples per Plot (n)								
			1			2			3		
			Number of Plugs per Composite			Sample (p_1)					
			21	9	5	21	9	5	21	9	5
<5	<6	<1.48	-----Less than 0.05-----								
5	6	1.48	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
5.25	6.25	1.54	0.07	0.06	0.06	0.08	0.07	0.06	0.09	0.07	0.07
5.5	6.5	1.59	0.09	0.08	0.07	0.12	0.09	0.08	0.16	0.10	0.08
6	7	1.71	0.14	0.10	0.09	0.21	0.14	0.11	0.26	0.16	0.12
6.5	7.5	1.82	0.21	0.13	0.11	0.32	0.19	0.14	0.41	0.24	0.17
7	8	1.94	0.27	0.17	0.13	0.43	0.24	0.18	0.56	0.32	0.22
8	9	2.17	0.40	0.23	0.17	0.62	0.36	0.24	0.77	0.47	0.32
9	10	2.40	0.51	0.29	0.20	0.76	0.46	0.31	0.89	0.60	0.40
10	11	2.63	0.60	0.34	0.24	0.85	0.55	0.37	0.95	0.69	0.49
11	12	2.86	0.67	0.39	0.27	0.91	0.62	0.42	0.98	0.77	0.55
12	13	3.09	0.73	0.44	0.29	0.94	0.67	0.47	0.99	0.82	0.61
13	14	3.32	0.78	0.48	0.32	0.96	0.72	0.51	0.99	0.86	0.65
14	15	3.55	0.81	0.50	0.34	0.97	0.76	0.54	1.0	0.89	0.69
18	19	4.47	0.90	0.60	0.41	0.99	0.85	0.64	1.0	0.95	0.79
25	26	6.08	0.95	0.69	0.48	0.99	0.92	0.73	1.0	0.98	0.87

+Computed using Eq. (5) and μ from Column 2.

Probabilities of Taking Additional Remedial Action for
3 Decision Rules; One Composite Sample per Plot

FIGURE 14. Probabilities of Taking Additional Remedial Action in a Plot for Three Decision Rules When One 500-g Sample from a Composite Sample Composed of Either 21, 9, or 5 Soil Plugs from the Plot is Measured for Ra-226. Probabilities are from Tables 3, 4, and 5.

Probabilities of Taking Additional Remedial Action

FIGURE 15. Probabilities that the 50% Confidence Rule will Indicate Additional Remedial Action is Needed if One, Two, or Three 500-g Samples from a Composite Sample Composed of 9 Soil Plugs are Measured for Ra-226. Probabilities are from Table 4.

1. Decreasing the number of plugs per composite sample increases the probability of incorrectly deciding additional remedial action is needed. For example, if the upper confidence limit rule is used (rule 1), if one composite sample is collected, if the true mean for the plot is 3 pCi/g above background, and if background is 1 pCi/g, then the probability the rule will indicate additional remedial action is needed increases from 0.37 to 0.64 if a 9-plug rather than a 21-plug composite sample is used to estimate the plot mean (see Table 3 and Fig. 14).
2. Decreasing the number of plugs per composite sample increases the probability of incorrectly deciding additional remedial action is not needed. For example, if the lower confidence limit rule is used (rule 3), if one composite sample is collected, if the true plot mean is 10 pCi/g above background, and if background is 1 pCi/g, then the probability that rule 3 will correctly indicate additional remedial action is needed decreases from 0.60 to 0.34 if a 9-plug rather than a 21-plug sample is used (see Table 5 and Fig. 14).
3. Taking more than one composite sample per plot reduces the probability of incorrectly deciding additional remedial action is needed. For the example in number 1 immediately above, the probability decreases from 0.64 to 0.43 if two composite samples rather than one are collected to estimate the mean (see Table 3).
4. For plots with mean concentrations near 5 pCi/g above background, the probabilities of taking additional remedial action are highly dependent on which decision rule is used. For example, if the upper confidence limit rule is used, the probability is greater than 0.95 that the test will indicate additional remedial action is needed when the plot has a mean ²²⁶Ra concentration greater than 5 pCi/g above background. But if the lower confidence limit rule is used, even if 21-plug composite samples are collected, the probability that the test will indicate additional remedial action is needed does not reach 0.95 until the true plot mean is about 20 pCi/g above background. Decision rule 2 falls between these two extremes. It achieves a 0.95 probability (for one or more 21-plug samples) when the true mean above background is about 9 or 10 pCi/g (see Table 4 and Fig. 14).

The three decision rules may find application at different times in the remedial action process. The upper confidence limit rule seems most

appropriate at initial stages when it may be prudent to assume that the plot is contaminated until proven otherwise. The "price" of using this rule is increased remedial action costs for plots that have true mean concentrations just under 5 pCi/g above background. The lower confidence limit rule seems more appropriate for plots that are strongly believed to have already been cleaned to below the EPA limit. Using this rule, the probability of taking additional remedial action is less than 0.05 when the true plot mean is 5 pCi/g above background or less.

The choice of a decision rule and of the number of plugs and samples will have an impact on remedial action decisions only for plots with true means less than about 20 or 25 pCi/g above background (see Fig. 14). For more highly contaminated plots, all three test rules, even when only one 9-plug composite sample is used to estimate the plot mean, are highly likely to indicate additional remedial action is needed. Decision rule 2 is obviously a compromise between rules 1 and 3. It also has the advantage that the test is very simple. One simply checks to see if the estimated plot mean is greater than 5 pCi/g above background.

If rules 1 and 3 are to yield the probabilities given in Tables 3 and 5, the true standard deviation for the plot must be given by Eq. (6). At contaminated sites where this model does not apply, special soil sampling studies could be conducted to determine whether Eq. (6) or some other model is applicable. Alternatively, if two or more composites were collected from each plot then the standard deviation could be estimated directly for each plot using those data. Then upper or lower confidence limits would be computed using the t distribution rather than the normal distribution [see Exner et al. (1985) for the upper confidence limit test]. Use of the t distribution will generally give more decision errors, which is the price paid when the standard deviation must be estimated.

However, as concerns the comparison of 21-, 9-, and 5-plug samples, the increase in probabilities of decision errors as the number of plugs per composite sample is reduced is, on the whole, about the same as obtained here in Tables 3-5 when the standard deviation, σ_{p_1} , was assumed known. This conclusion is based on probabilities of decision errors we obtained using the noncentral t distribution and the methods in Wine (1964, pp. 254-260).

4.0 STANDARD DEVIATIONS OF REPEAT MEASUREMENTS AND OF REPLICATE COMPOSITE SAMPLES

In Sections 3.1 and 3.2 we examined the Shiprock data to assess how the standard deviation of ^{226}Ra measurements on multiple composite samples from a plot is related to the number of soil plugs used to form each composite sample. To establish a lower bound on these standard deviations and to assess the quality of these data, we estimate in this section the standard deviation of (1) repeat ^{226}Ra measurements on 500-g composite samples, and of (2) replicate 500-g composite samples withdrawn from the same bucket of mixed soil.

Table 6 lists the individual repeat measurements greater than 2 pCi/g that were made on composite samples collected at Shiprock. Table 7 lists the repeat measurements that were less than 2 pCi/g. The measurements in Column B in Tables 6 and 7 were made by the same person (employed by the Remedial Action Contractor, RAC) using the same count system on the same sample as the measurement in Column A. The measurement in Column C of the tables was made on the same composite sample as measured in Columns A and B, but by a different person (employed by the Technical Assistance Contractor, TAC) on a different count system on a different day. Tables 6 and 7 give the range (maximum minus minimum repeat measurement) and the standard deviation, s , of the measurements made on each composite sample, where s was computed using Eq. (2).

The differences between repeat measurements in Table 6 were less than 1 pCi/g except for one sample for which the maximum difference was 3 pCi/g. The standard deviations for the repeat measurements ranged from 0 to 0.71 pCi/g except for one s equal to 1.68, which was atypical. The arithmetic mean of the 35 values of s was 0.31 pCi/g. The maximum observed difference between repeat measurements in Table 7 was 0.50, which is smaller than for the measurements in Table 6 that had higher ^{226}Ra measurements. The values of s in Table 7 ranged from 0.10 to 0.29, and the average was 0.19.

These data suggest that for plots with concentrations less than 5 pCi/g (including background) it seems reasonable to assume repeat measurements will not usually differ by more than about 0.5 pCi/g, and the standard deviation of these repeat measurements will range from about 0.1 to 0.3 pCi/g. But plots

Table 6. Repeat Measurements of ^{226}Ra concentrations in Composite Surface Soil Samples with Measurements Greater than 2 pCi/g.

Number of Plugs Per Composite	Area	Replicate Measurements of the Same Composite Sample (pCi/g)			Range (max.-min.) (pCi/g)	Standard Deviation (pCi/g)
		A*	B*	C**		
5	4	13.3	14.3	13.5	1.0	0.53
5	7	9.2	9.4	10.0	0.8	0.42
5	1	9.3	8.5	8.7	0.8	0.42
5	3	6.8	7.2	7.1	0.4	0.21
5	6	6.2	5.6	6.3	0.7	0.38
5	8	3.5	4.1	3.5	0.6	0.35
8	7	11.0	11.2	10.9	0.3	0.15
8	1	5.0	4.3	4.8	0.7	0.36
8	5	3.2	2.6	2.5	0.7	0.38
8	4	2.4	2.7	2.5	0.3	0.15
8	8	2.4	2.4	2.2	0.2	0.12
9	9	2.2	2.0	1.9	0.3	0.15
9	7	15.8	15.6	18.6	3.0	1.68
9	1	10.9	11.4	10.8	0.6	0.32
9	4	5.7	5.5	5.4	0.3	0.15
9	2	4.9	4.9	4.8	0.1	0.058
9	3	4.6	4.3	4.3	0.3	0.17
9	8	2.7	3.1	2.6	0.5	0.26
9	2	2.0	2.3	2.0	0.3	0.17
16	7	6.7	6.7	6.7	0.0	0.0
16	1	4.1	5.1	---	1.0	0.71
16	4	3.6	3.9	3.8	0.3	0.15
16	10	3.1	2.5	3.0	0.6	0.32
16	6	2.6	2.3	2.1	0.5	0.25
16	8	2.4	2.6	2.1	0.5	0.25
16	9	2.3	2.0	2.2	0.3	0.15
21	1	11.2	11.2	11.6	0.4	0.23
21	4	9.7	10.1	9.9	0.4	0.20
21	3	5.6	5.8	5.6	0.2	0.12
21	7	4.5	5.3	5.8	1.3	0.66
21	3	4.8	4.2	4.2	0.6	0.35
21	8	3.9	3.2	3.0	0.9	0.47
21	2	2.4	2.4	2.8	0.4	0.23
21	5	2.4	2.5	2.4	0.1	0.058
21	9	2.4	1.8	1.6	0.8	0.42

Arithmetic Mean = $\overline{0.31}$ pCi/g

- * Duplicate measurements made by the Remedial Action Contractor (RAC). Measurements were made by the same person using the same count system, but not necessarily on the same day.
- ** Measurement made by the Technical Assistance Contractor (TAC) on a different count system on a different day than the duplicate measurements made by the RAC.

Table 7. Repeat Measurements of ^{226}Ra Concentrations in Composite Surface Samples with Measurements less than 2 pCi/g.

Number of Plugs Per Composite	Area	Replicate Measurements of the Same Composite Sample (pCi/g)			Range (max.-min.) (pCi/g)	Standard Deviation (pCi/g)
		A*	B*	C**		
5	2	1.5	1.2	1.7	0.5	0.25
5	9	1.1	1.4	1.2	0.3	0.15
5	10	1.1	1.0	0.7	0.4	0.21
8	10	1.2	1.6	1.1	0.5	0.26
8	5	1.6	1.1	1.1	0.5	0.29
8	2	1.4	0.9	1.4	0.5	0.29
8	6	1.1	1.1	1.4	0.3	0.17
9	6	1.6	1.2	1.1	0.5	0.26
9	5	1.0	1.2	1.0	0.2	0.12
9	9	1.2	1.3	1.0	0.3	0.15
9	10	1.1	1.3	0.9	0.4	0.20
16	5	1.5	1.6	1.4	0.2	0.10
21	6	1.3	1.1	1.3	0.2	0.12
21	10	1.2	1.3	1.4	0.2	<u>0.10</u>

Arithmetic Mean = 0.19 pCi/g

-
- * Duplicate measurements made by the Remedial Action Contractor (RAC). Measurements were made by the same person using the same count system, but not necessarily on the same day.
- ** Measurement made by the Technical Assistance Contractor (TAC) on a different count system on a different day than the duplicate measurements made by the RAC.

with concentrations exceeding 5 pCi/g could be expected in many cases to have repeat measurements differing by more than 0.5 pCi/g, and differences of more than 1 pCi/g could occur on occasion.

In addition to taking repeat measurements on the same 500-g composite sample, replicate 500-g composite samples from each of 10 buckets of mixed soil were also obtained and assayed for ^{226}Ra . These measurements, one per 500-g sample, are given in Table 8. The variability in the measurements for multiple composite samples from a given bucket reflects measurement error in addition to bucket sampling error.

In Table 8 we find that for the 6 buckets with measurements exceeding 2 pCi/g, the standard deviations ranged from 0.15 to 0.92 pCi/g, with an arithmetic mean of 0.50 pCi/g. For the 4 buckets with measurements less than 2 pCi/g, the standard deviations ranged from 0.13 to 0.85 pCi/g, and the arithmetic mean was 0.34 pCi/g. The grand arithmetic mean for the 10 values of s from the two groups was 0.44 pCi/g, which may be compared with the arithmetic means of 0.31 and 0.19 pCi/g for repeat measurements (Tables 6 and 7).

To summarize Tables 6, 7, and 8, the data suggest that the standard deviations for repeat measurements will average about 0.3 pCi/g, and the standard deviation of replicate composite samples from the same bucket will average about 0.4 pCi/g. This suggests that about 75% of the measurement plus bucket sampling error was due to measurement error ($0.3/0.4 = 0.75$). In Section 3.1 above we saw that the standard deviations of multiple composite samples from decontaminated field plots (plots 9 and 10) were also about 0.3 and 0.4 in most cases. This suggests that rather uniform levels of ^{226}Ra remaining after these plots had been decontaminated.

Table 8. Measurements of ^{226}Ra Concentrations (pCi/g) on Replicate 500-g Composite Samples taken from the Same Soil Mass.

<u>Number of Plugs Per Composite</u>	<u>Area</u>	<u>Measurements on Replicate Composite Samples (pCi/g)</u>	<u>Range (max.-min.) (pCi/g)</u>	<u>Standard Deviation (pCi/g)</u>
9	2	3.6, 4.9	1.3	0.92
21	4	9.7, 9.7, 9.7, 10.1	0.4	0.20
21	7	8.1, 8.9, 9.4, 10.1	2.0	0.84
21	3	4.3, 4.9, 5.0, 5.6, 5.7	1.4	0.57
21	8	2.5, 2.6, 2.8, 2.8	0.3	0.15
21	5	2.1, 2.3, 2.6, 2.7, 2.9	0.8	<u>0.32</u>
Arithmetic Mean =				0.50
8	9	1.0, 2.2	1.2	0.85
16	2	1.5, 1.5, 1.5, 1.5, 2.0	0.5	0.22
21	9	1.3, 1.3, 1.4, 1.6, 1.7	0.4	0.18
21	10	0.7, 0.8, 0.8, 1.0	0.3	<u>0.13</u>
Arithmetic Mean =				0.34
Grand Arithmetic Mean =				0.44

5.0 ACCEPTANCE SAMPLING FOR REDUCING THE NUMBER OF PLOTS REQUIRING SOIL SAMPLES

There are a very large number of 10-m by 10-m plots in the United States that must be verified to have mean ^{226}Ra concentrations less than 5 pCi/g above background after remedial action has been performed. The present practice is to take at least one composite soil sample in each and every plot. Since the cost of collecting and analyzing these samples will be very large, it is important to consider whether a sampling plan could be developed so that soil samples need not be collected in all plots. In this section we briefly introduce an approach called acceptance sampling (Schilling, 1982) that, in conjunction with in-situ gamma scans of all plots, might be appropriate for this objective.

Acceptance sampling would involve selecting a subset of the plots from which soil samples would be collected. For each of these plots, a decision rule such as one of the three rules discussed in Section 3.5 would be used to test whether additional remedial action was needed for the plot. If, on the basis of these tests, one or more of these plots required additional remedial action, then the plan might require that soil samples be collected in some or all of the remaining plots in the region. As noted above, gamma scans would be made in all plots, whether or not soil samples were collected.

The rationale behind this type of acceptance sampling is a willingness to accept that some small proportion of the plots that have not been sampled may actually have mean ^{226}Ra concentrations greater than 5 pCi/g above background. The number of plots initially selected for soil sampling would be determined by specifying (1) a suitable low probability of taking soil samples in additional plots if the percentage of plots exceeding 5 pCi/g is some acceptable small value Q_1 , and (2) a suitable high probability of taking soil samples in additional plots if the percentage of plots exceeding 5 pCi/g is some specified value Q_2 that is greater than Q_1 . Additional information on acceptance sampling is given in the paper by Gilbert et al. (1981) in Appendix B.

6.0 CONCLUSIONS

The ^{226}Ra measurements made on multiple composite soil samples collected in 10 selected areas in the windblown mill-tailings flood plain region at Shiprock, New Mexico, have been examined to study the effect of the number of soil plugs per composite sample on the probability of making remedial-action decision errors and on the accuracy with which mean ^{226}Ra concentrations for 10-m by 10-m plots can be estimated.

The Shiprock data indicate that for plots that had effective remedial action so that residual contamination levels were less than about 2 pCi/g, the number of plugs per composite sample had no effect on the accuracy of estimated means. However, for plots at Shiprock with residual concentrations greater than about 2 pCi/g, reducing the number of plugs by about half increased the standard deviation of composite samples by about 30 or 40% on the average (see the GMs in Table 1).

Based on these results a simple model (Eq. 3) was used to approximate the increase in standard deviation that occurred as the number of plugs per composite sample decreased from 21 to a smaller number p_1 . This model, in conjunction with an estimate of the standard deviation of 21-plug composite samples collected at Shiprock, was used to obtain an equation (Eq. 7) for computing the percent accuracy of the estimated plot mean. This equation indicated that the accuracy with which the mean is estimated will decrease (get worse) by factors of 1.5 and 2.0 as the number of plugs is decreased from 21 to 9 and from 21 to 5, respectively.

The percent accuracy with which the plot mean can be estimated using one 9-plug composite sample was estimated to be 50% using the model, which compares with 30% obtained by a study (Williams et al. 1986) of five contaminated properties in Grand Junction, Colorado. A possible explanation for this apparent discrepancy is that at Grand Junction a smaller size drill bit was used to collect 20-plug samples. This may have caused an increased variance for the 10-plug samples. Another possibility is that the variability for the 10-plug samples at Grand Junction may have been biased low because a substantial portion of two properties were not sampled by the 10-plug method.

The magnitude of changes in the probability of making incorrect remedial action decisions as the number of soil plugs per composite decreases from 21 depends on the particular statistical test used to make the decision. Relatively small changes (less than 10 probability points) occur if the decision is made by simply comparing the estimated plot mean above background with the EPA limit of 5 pCi/g above background. If the decision is made by comparing the upper 95% confidence limit on the plot mean above background to the EPA limit, then the probability of incorrectly deciding additional remedial action is required increases from 15 to 30 probability points for true plot means between about 1.5 and 4 pCi/g above background if 9-plug rather than 21-plug composite samples are used. Similar increases occur in the probability of incorrectly deciding additional remedial action is not required if the true plot mean above background is between 6 and 10 (or greater) pCi/g and the lower 95% confidence limit test rule is used.

The above differences in probabilities were obtained assuming the standard deviation of composite samples is known to be given by the model for standard deviations developed in this paper. In practice the standard deviation for a given plot is never known with absolute assurance. However, even in this situation, the differences in probabilities of making decision errors as the number of soil plugs per composite sample is reduced are overall about the same as stated above.

7.0 REFERENCES

- Burr, I. W. 1976. Statistical Quality Control Methods, Marcel Dekker, New York.
- Exner, J. H., W. D. Keffer, R. O. Gilbert, and R. R. Kinnison 1985. "A Sampling Strategy for Remedial Action at Hazardous Waste Sites: Clean-up of Soil Contaminated by Tetrachlorodibenzo-p-Dioxin," Hazardous Waste and Hazardous Materials 2:503-521.
- United States Environmental Protection Agency 1983. Standard for Remedial Actions at Inactive Uranium Processing Sites; Final Rule (40 CFR Part 192). Federal Register 48 (3):590-604 (January 5, 1983).
- Gilbert, R. O. 1987. Statistical Methods for Environmental Pollution Monitoring, Van Nostrand Reinhold, Inc., New York.
- Gilbert, R. O., R. R. Kinnison, and M. G. Barnes 1981. "On the Application of Acceptance Sampling to the Decommissioning of Nuclear Facilities", In: Environmetrics 81, Summary of Conference Presentations, SIAM Institute for Mathematics and Society, Philadelphia.
- Schilling, E. G. 1982. Acceptance Sampling in Quality Control, Marcel Dekker, Inc, New York.
- Williams, L. R., R. W. Leggett, M. L. Espegren, C. A. Little 1986. Optimization of Sampling for the Determination of the Mean Radium-226 Concentration in Surface Soil, Oak Ridge National Laboratory, Oak Ridge, TN, ORNL/TM-10255
- Wine, R. L. 1964. Statistics for Scientists and Engineers, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

APPENDIX A

²²⁶Ra Measurements for 5-, 8-, 9-, 16-, and 21-plug
Composite Soil Samples Collected in Areas 1 through
10 in the Windblown Mill-tailings Flood Plain Area at
Shiprock, New Mexico, in January and March of 1986.

TABLE A1. ^{226}Ra Measurements for 5-, 8-, 9-, 16-, and 21-plug Soil Composite Samples Collected in Areas 1 through 10 in the Windblown Mill-tailings Flood Plain Area at Shiprock, New Mexico.

Area	Soil Composite	Number of Plugs Per Composite				
		5	8	9	16	21
1	1	12.8	7.6	9.0	4.0	7.7
	2	9.3	5.0	8.7	5.9	5.12
	3	8.1	*	10.7	*	6.50
	4	8.8	7.7	10.9	4.6	14.0
	5	10.6	*	9.4	*	7.8
	6	10.0	*	10.6	*	12.3
	7	9.4	2.8	10.5	3.7	11.2
	8	9.3	*	11.6	*	6.0
	9	13.6	4.7	7.5	4.1	10.3
2	1	*	1.3	2.0	1.4	2.9
	2	1.5	1.4	2.3	1.5	2.2
	3	1.9	*	1.5	*	1.7
	4	*	1.4	4.9	2.0	2.4
	5	2.3	*	1.5	*	2.0
	6	*	*	1.4	*	3.4
	7	2.9	5.3	2.4	1.3	*
	8	*	1.4	*	*	*
	9	*	*	*	0.99	*
3	1	5.2	2.3	3.9	1.2	4.8
	2	6.2	1.8	7.4	1.5	5.2
	3	4.6	1.3	4.9	2.0	4.6
	4	3.7	2.1	4.9	1.7	5.1
	5	7.3	1.8	7.3	1.9	4.7
	6	6.3	0.9	4.4	0.7	4.0
	7	6.8	3.1	4.6	1.3	5.5
	8	6.9	1.7	5.3	2.3	3.9
	9	7.3	1.7	5.4	1.0	4.3
4	1	8.5	1.4	8.4	7.5	10.1
	2	13.3	4.7	7.6	3.6	6.7
	3	*	*	*	*	*
	4	6.7	4.3	14.4	3.6	10.1
	5	*	*	*	*	*
	6	*	*	*	*	*
	7	10.5	2.4	6.5	6.7	11.1
	8	*	*	*	*	*
	9	13.3	3.5	5.7	2.0	9.3

TABLE A1. (Continued)

Area	Soil Composite	Number of Plugs Per Composite				
		5	8	9	16	21
5	1	2.9	1.5	3.0	1.9	2.6
	2	1.3	2.0	2.9	2.1	2.6
	3	6.5	3.2	1.9	2.9	2.1
	4	2.6	1.1	1.1	1.1	2.4
	5	1.8	1.0	8.0	1.7	2.5
	6	7.3	1.6	1.0	1.5	2.6
	7	5.1	2.4	2.2	2.1	1.2
	8	2.5	3.0	1.8	3.0	2.4
	9	16.6	1.7	1.7	0.5	4.0
6	1	3.5	2.1	1.0	2.6	1.2
	2	2.0	1.3	1.6	1.3	2.4
	3	2.8	1.1	0.9	1.6	1.9
	4	2.3	2.0	0.9	1.4	1.3
	5	4.2	2.1	0.5	1.7	0.9
	6	6.2	1.8	0.8	1.6	2.4
	7	4.4	2.1	1.5	1.5	1.9
	8	4.8	1.5	1.3	1.7	2.6
	9	4.4	1.9	0.7	1.5	3.4
7	1	11.8	9.0	10.3	8.6	9.1
	2	15.4	11.0	5.9	7.2	4.5
	3	13.4	6.7	13.4	7.1	9.1
	4	14.6	7.3	6.7	6.8	7.1
	5	13.1	6.7	15.8	6.7	8.9
	6	22.2	8.3	9.7	8.0	8.9
	7	9.2	7.9	12.7	8.7	9.1
	8	6.9	6.5	10.9	6.7	10.1
	9	11.4	*	10.1	*	5.4
8	1	2.4	1.9	2.7	2.9	2.8
	2	3.6	2.0	3.8	2.2	3.0
	3	3.8	2.8	1.7	2.7	3.9
	4	2.5	2.5	2.5	1.5	3.4
	5	3.5	2.4	3.2	2.4	2.6
	6	2.3	3.5	2.4	3.2	2.1
	7	2.1	1.9	2.1	1.6	1.7
	8	2.0	1.2	1.5	1.6	2.4
	9	1.9	1.8	2.7	2.1	1.9

TABLE A1. (Continued)

Area	Soil Composite	Number of Plugs Per Composite				
		5	8	9	16	21
9	1	0.7	1.0	0.05	1.5	1.5
	2	0.7	1.7	0.08	2.3	1.6
	3	0.5	1.5	0.7	1.4	2.0
	4	0.8	1.6	0.3	M	1.1
	5	1.0	1.1	0.9	1.6	1.4
	6	0.01	1.1	0.5	0.8	1.3
	7	0.6	1.4	0.7	1.5	2.4
	8	0.7	2.0	M	1.7	1.0
	9	1.1	1.2	1.2	1.8	0.9
10	1	0.1	2.0	0.8	0.9	1.0
	2	0.8	1.2	0.8	0.8	0.6
	3	0.6	1.6	0.8	0.7	0.7
	4	M	1.1	0.7	1.0	1.2
	5	1.1	1.8	1.0	3.1	0.6
	6	0.8	1.5	0.4	2.0	0.7
	7	0.1	1.7	0.7	1.5	0.7
	8	0.7	0.6	1.1	2.5	0.2
	9	0.3	0.7	0.6	2.1	1.3

* Composite sample not collected.

M = Missing data

APPENDIX B

"On the Application of Acceptance Sampling to the Decommissioning of Nuclear Facilities," by R. O. Gilbert, R. R. Kinnison, and M. G. Barnes, 1981, In: Summaries of Conference Presentations, Environmetrics 81, SIAM Institute for Mathematics and Society, Philadelphia, pp. 140-141.

ON THE APPLICATION OF ACCEPTANCE SAMPLING TO THE DECOMMISSIONING OF NUCLEAR FACILITIES*

Richard O. Gilbert, Robert R. Kinnison, Madaline G. Barnes

Decommissioning-Decontamination (D&D) Operations

There are many formerly used nuclear sites and facilities that may or will require remedial action (RA) before they can be released for unrestricted public use. These include laboratories, ore processing and fuel fabrication plants, nuclear reactor sites, uranium mill tailings piles, radioactive waste burial areas and storage sites, and nuclear weapons testing sites. Remedial action might consist of dismantlement and proper disposal of contaminated or obsolete facilities, or the decontamination, removal or control of soil or tailings piles. Hine (1) and Ahlquist (2) discuss recently conducted D&D operations. DOE (3) gives an overview of D&D planning by the Department of Energy and outlines environmental health and safety concerns.

There are no universally used statistical designs for obtaining data for judging compliance with guidelines during and following D&D operations. Some of the basic design approaches developed by Leggett et al. (4) seem a step in the right direction. They advocate dividing a site/facility into logical units and smaller subunits or blocks. A unit might be a floor or wall of a room, a loading dock, a tract of land, a piece of equipment, etc. Blocks might be square units of small size on floors or walls. For outdoor areas, blocks would be larger due to the greater areas involved. Each block may be surveyed by in-situ radiation detecting instruments and/or physical samples may be collected for radiochemical analyses. Measurements and samples may be taken at random locations or at the center and corner of each block. Also, the surface of the entire block may be scanned (100% "inspection") to locate the point of maximum radioactivity.

Acceptance Sampling for Compliance Testing

During and following D&D/RA operations there are frequent occasions for comparing radionuclide measurements with regulatory guidelines. In these situations consideration should be given to using acceptance sampling (AC) methods unless 100% inspection can be conducted. Two examples where 100% inspection is not usually practical are (1) subsurface burial or contamination, and (2) very large land areas. In these situations we cannot know with absolute certainty the true average and maximum concentrations that exist in a given unit or block. Hence, there is always some chance of making Type I and Type II decision errors. By using AC methods we can choose the magnitude (α and β , respectively) of these two errors that we are willing or required to accept. In particular, we can construct the operating characteristic (OC) curve of a proposed AC plan and see whether it fits the cost-benefit situation at hand.

The use of AC plans require unambiguous definitions of the "lot" and the "items" that make up the lot. Also, the lot should be homogeneous in distribution

before AC plans are applied. A lot might be a plot of soil to a specified depth, a section of drainage ditch, a room, a piece of equipment, a loading dock, or a "block" as defined above.

The number of items (N) in the lot is the size of the lot, where an item is the area or volume to which a measurement datum refers. For example, N might be the number of potential non-overlapping in-situ soil surface readings in a lot of specified dimensions. If physical samples of soil are being taken for radiochemical analyses, N would be the maximum possible number of such samples of specified volume that could be collected in the lot. Clearly, N may be very large.

Attribute Plans

Both attribute and variables inspection AC plans might be used during D&D/RA operations. Attribute inspection plans are those where each item is classified as defective or not defective. Variables plans apply when a quantitative measurement is obtained for each item. The attribute approach seems appropriate where interest centers on individual measurements being less than some maximum value. Lot quality is measured by the fraction of the N items in the lot that are "defective." Each of the n items in the lot that are examined ($n < N$) is considered "defective" if it exceeds the maximum permitted value (L). If L is exceeded by more than the number c of defectives allowed by the AC plan, then the "lot" is examined more closely and RA may be required. This approach seems compatible with regulations that permit no more than a fixed number of measurements to exceed a specified limiting value over a given time period or in a unit of space. Plans can be set up where $c = 0$, i.e., the lot is accepted only if none of the n examined items are defective (Schilling (5), Barter (6)). However, AC plans that allow one or two defectives can be constructed that provide the same protection to the consumer if n is also increased (Grant (7), p. 339). These latter plans are more likely to result in accepting lots of acceptable quality than if we require $c = 0$.

Variables Plans

Variables AC plans can be constructed for situations where lot quality is expressed in terms of either fraction defective or as acceptable (μ_0) and unacceptable (μ_1) mean lot concentrations (Duncan (8), pp. 195-265). The fraction defective approach is highly dependent on the data being Gaussian distributed (or exactly lognormal if log-transformed data are used). Radionuclide data are frequently skewed towards high values, but the distribution may not be lognormal. If acceptable and unacceptable lot quality is defined in terms of averages the data need only be approximately normal (or lognormal) since only the mean ($\bar{x}$) of the n data obtained from the lot need be normally distributed. This may be a reasonable assumption due to the central limit theorem.

Figure 1 shows one way of setting up a variables AC plan based on mean lot quality, i.e., by specifying μ_0 , μ_1 , and the accompanying $1-\alpha$ and β probabilities of accepting the lot. In D&D/RA situations, these four parameters must be determined on the basis of pertinent

*For further information contact Richard O. Gilbert, Pacific Northwest Laboratory, P.O. Box 999, Richland, WA, 99352. Telephone 509-376-4218.

Prepared for the U.S. Department of Energy under Contract DE-AC06-76RLO 1830.

FIGURE 1. A typical OC curve for variables AC plan.

regulations and guidelines, cost-benefit considerations, environmental background levels, and "as low as reasonable achievable" (ALARA) considerations. The number of measurements n can be determined once u_0 , u_1 , α , β and s (an estimate of the standard deviation for the data in the lot) are specified (Duncan (8), pp. 255-258). The decision to accept or reject the lot is made by computing the one-sided lower confidence limit (LCL). The lot is rejected if $LCL \geq u_0$. In this situation α and β in Fig. 1 are the familiar "producers" and "consumers" risk in AC terminology. Note that we are testing $H_0: \mu < u_0$ versus $H_A: \mu \geq u_0$ and that the lot may be accepted even though $\bar{x} > u_0$.

Greater protection is afforded the consumer (user of the site/facility after D&D) if we test $H_0: \mu \geq u_0$ versus $H_A: \mu < u_0$. For this situation we compute the one-sided upper confidence limit (UCL) and reject the lot if $UCL \geq u_0$, which can happen even when $\bar{x} < u_0$. Statisticians and D&D project planners should understand the practical consequences of choosing the way H_0 and H_A are stated and hence whether the UCL or LCL is used for making decisions. Both approaches are discussed by Leidel et al. (9) when testing for compliance with air quality standards for toxic chemicals. To satisfy ALARA considerations, u_0 as well as u_1 may be set lower than required to meet guidelines based on maximum annual dose to an individual.

In situations where data are available on a fast turnaround basis, it may be possible to use sequential attribute or variables AC plans. This would tend to reduce n over what needed for single sampling plans (as discussed above) or for double sampling plans. Whatever sampling approach is used, the n measurements or samples should be taken from all portions of the lot to assure better coverage.

Kriging

Frequently, environmental radionuclide data are correlated over space. In this situation it is sometimes possible to use Kriging to obtain minimum variance unbiased estimates of average concentrations for lots by estimating the correlation structure (variogram) (Barnes (10), Simpson and Gilbert (11)). If Kriging errors are Gaussian, confidence limits on lot means and confidence bands on the concentration contours can be constructed.

We emphasize that the statistician responsible for advising the D&D/RA manager must be familiar with the practical day to day problems that arise in these operations and must have a basic understanding of the sources of error in radionuclide data. The ivory tower approach won't work. The statistician should become familiar with the site and be a working member of the D&D planning team.

Other Applications

AC methods may be applicable to other environmental compliance testing situations such as the cleanup of toxic chemical waste dumps and storage sites. Presumably, environmental samples will be required to judge whether residual levels of these chemicals are in compliance with guidelines. Also, there may be a need to develop AC techniques specifically for spatially distributed environmental contaminants (including nuclear D&D situations) to better handle problems such as non-independent data and highly skewed data sets.

References

1. Hine, R. E. 1980. Decontamination and Decommissioning of the Organic Moderated Reactor Experiment Facility (OMRE). EG&G Idaho, Inc., Idaho Falls, Idaho, NTIS, EGG-2059.
2. Ahlquist, A. J. 1979. LASL Experience in Decontamination of the Environment. Los Alamos Scientific Laboratory, Los Alamos, NM, LA-UR-79-3275.
3. DOE. 1979. Environmental Development Plan, Decontamination and Decommissioning, U.S. Dept. of Energy, NTIS, DOE/EDP-0055.
4. Leggett, R. W., H. W. Dickson, F. F. Haywood. 1979. "A Statistical Methodology for Radiological Surveying," Advances in Radiation Protection Monitoring, IAEA, Vienna, pp. 541-554.
5. Schilling, E. G. 1978. "A Lot Sensitive Sampling Plan for Compliance Testing and Acceptance Inspection," J. of Quality Technology, 10, No. 2, pp. 47-51.
6. Barter, K. E. December 1967. "Sample Sizes for An Acceptance Number of Zero," Indus. Quality Control, pp. 322-324.
7. Grant, E. L. 1964. Statistical Quality Control, McGraw-Hill Book Co., New York, 3rd Edition.
8. Duncan, A. J. 1959. Quality Control and Industrial Statistics, Richard D. Irwin, Inc., Homewood, IL.
9. Leidel, M. A., K. A. Busch, J. R. Lynch. 1977. Occupational Exposure Sampling Strategy Manual, Dept. Health, Education and Welfare, OHEW (NIOSH), Publication No. 77-173.
10. Barnes, M. G. 1980. "The Use of Kriging for Estimating the Spatial Distribution of Radionuclides and Other Spatial Phenomena," TRAN-STAT, Statistics for Environmental Studies (R.O. Gilbert, ed), Issue No. 13, Pacific Northwest Laboratory, Richland, WA, PNL-SA-9051.
11. Simpson, J. C., R. O. Gilbert. 1980. "The Use of Kriging to Estimate Spatial Patterns and Inventories of Environmental Contaminants," Proceedings of this Environmetrics 81 conference.

DISTRIBUTION

No. of
Copies

OFFSITE

10 J. R. Anderson, Project Manager
Uranium Mill Tailings Project Office
U. S. Department of Energy
5301 Central Avenue, N. E.
Albuquerque, New Mexico 87115

5 T. Brazley
U. S. Department of Energy
19901 Germantown Road
Germantown, Maryland 20767

30 DOE Technical Information Center

K. R. Baker
Roy F. Weston, Inc.
Suite 1700
5301 Central Avenue, N. E.
Albuquerque, New Mexico 87108

B. A. Berven
Oak Ridge National Laboratory
P. O. Box X, Building 7503
MS 382
Oak Ridge, Tennessee 37831

C. A. Little
Oak Ridge National Laboratory
Grand Junction Office
P. O. Box 2567
Grand Junction, Colorado 81502

H. R. Meyer
Chem-Nuclear Systems
P. O. Box 9136
Albuquerque, New Mexico 87119

10 M. L. Miller
Roy F. Weston, Inc.
Suite 1000
5301 Central Avenue, N. E.
Albuquerque, New Mexico 87108

OFFSITE

R. H. Miller
Jacobs Engineering Group, Inc.
Route 2, Highway 94 South
Saint Charles, Missouri 63303

R. A. Nelson
Jacobs Engineering Group, Inc.
Route 2, Highway 94 South
Saint Charles, Missouri 63303

5 P. T. Owen
Remedial Action Program Information
Center
P. O. Box X, Building 2001
Oak Ridge, Tennessee 37831-6050

L. R. Williams
Indiana University at South Bend
1700 Mishawaka Avenue
P. O. Box 7111
South Bend, Indiana 46634

ONSITE

DOE Richland Operations Office

J. J. Sutey

20 Pacific Northwest Laboratory

D. J. Bates
W. M. Bowen
R. O. Gilbert (10)
K. R. Price
Publishing Coordination (2)
Technical Report Files (5)