Gemperline, Mark C. 1994. “Surface Sampling to Detect Hot Spots Presenting Unacceptable Human Health Risk”. Proceedings of the First International Congress on Environmental Geotechnics. BiTech Publishers Ltd., 173 – 11860 Hammershith Way, Richmond, British Columbia, Canado, V7A 5G1.ISBN 0 921095 32 5. pp.233-239.

Surface Sampling to Detect Unacceptable Human Health Risk

by
Mark Gemperline
U.S. Bureau of Reclamation
Denver, Colorado

Abstract

An equation is derived to estimate the minimum number of samples required to detect surface hot spot contamination which presents an unacceptable human health risk. It combines common statistical analysis and simplified human health risk analysis. Multiple contaminants having varying degrees of toxicity are considered. The number of samples required is shown to be dependent on a preselected contaminant concentration which may represent the detection limit of an analytical technique.

The equation focuses the user on the problem variables resulting in a better understanding of the uncertainties involved in identifying conditions representing unacceptable human health risk. It can be used directly to calculate the required surface sampling density for a prescribed level of acceptable human health risk or used with an existing sampling plan to help estimate the human health risk associated with a hot spot which may be left undetected.

Introduction

Identifying conditions which present an unacceptable risk to human health is paramount to environmental site investigations. The required number of surface soil samples collected is frequently determined by arbitrarily selecting the "hot spot" size which must be found and applying basic statistical concepts to assure detection with adequate confidence.

The method of selecting hot spot size varies from site to site and often is selected to represent the reasonable size of a contaminant spill. It is based on historical records or observations. This approach does not attempt to assure detection of a condition which represents a human health risk.

An equation is derived to estimate the minimum number of samples required to detect surface hot spot contamination which may present an unacceptable human health risk. It combines common statistical analysis and simplified human health risk analysis. A smallest hot spot which potentially poses an unacceptable threat to human health is hypothesized. Its size depends on a simplified model for contaminant distribution within the hot spot, site specific variables which describe potential for human ingestion and dermal contact, and the toxicity of the contaminant involved.

The equation is restricted to surface soils which may represent direct ingestion or dermal contact hazards and does not address exposure pathways associated with subsurface soil, air, or water contamination. The derivation considers only a single contaminant, however it can be used with multiple contaminants by normalizing contaminant concentrations with respect to relative toxicity. This procedure is discussed.

The number of samples required is shown to be dependent on a preselected contaminant concentration which may represent the detection limit of an analytical technique; the minimum acceptable average concentration of an exposure unit which would result in a threat to human health; and the maximum

Gemperline, Mark C. 1994. “Surface Sampling to Detect Hot Spots Presenting Unacceptable Human Health Risk”. Proceedings of the First International Congress on Environmental Geotechnics. BiTech Publishers Ltd., 173 – 11860 Hammershith Way, Richmond, British Columbia, Canado, V7A 5G1.ISBN 0 921095 32 5. pp.233-239.

expected contaminant concentration.

Spacial variation of sample density and the use of successive stages having decreasing sampling density and contaminant detection limit are discussed.

Many site and exposure conditions are simplified for mathematical convenience. These simplifications are discussed so the reader may develop an understanding of the uncertainties involved in the calculation.

Simplified Risk Assessment

An exposure unit is defined herein as the contiguous area containing hot spots of contamination to which a person may be exposed. No exposure is expected from outside this area. A person exposed to a contaminant has a mass M, ingests soil at approximately SI mass-units/day and experiences DE mass-units/day soil dermal exposure. A long exposure period is assumed, consequently the average exposure unit concentration represents the average human exposure concentration. SIA and DEA are terms which represent the fraction of the ingested or dermal contacted contaminant which is absorbed into the body respectively.

The analysis period is the time considered significant to the analysis. For example, a lifetime of 70 years may be considered appropriate for assessing the risk associated with carcinogens. The variable, t_f, is the fraction of the analysis period which a person actually spends within the exposure unit.

Maximum allowable daily dose of contaminant per unit body mass, Rfd*, are published Rfd values for non-carcinogens or calculated by dividing the acceptable carcinogenic risk, R, by a slope factor, SF, for carcinogens. Rfd and slope factors values are found in the USEPA IRIS database as well as other databases and publications. Hazard index is defined as the ratio of the average daily adsorbed mass of contaminant during the analysis period to Rfd* [1].

$$C_{index} = \frac{Rfd^* x M}{(SI x SIA + DE x DEA) x t} \qquad \textbf{(1)}$$

The average exposure unit surface soil concentration which would result in a hazard index of 1, or a carcinogenic risk R, is approximated by the equation:
A hot spot, or cumulative sum of hot spots, which causes the average exposure unit concentration to be greater than C_{index} must be located.

Hot Spot Size

Presume for a moment that there is a single hot spot in the exposure unit which causes the average concentration of the exposure unit to be unacceptable. This is subsequently referred to as a significant hot spot. The size of the smallest hot spot which could result in this condition is sought.

It is assumed that the average hot spot contaminant concentration, C_{ave}, is 1/3 of the maximum exposure unit contaminant concentration, C_{max}. This is equivalent to assuming that a single circular hot spot exists which has C_{max} at its center with linearly decreasing concentration toward the perimeter as indicated by figure 1. C_{max} must be estimated based on site history, scoping information and expected fate and transport of site contaminants.

The contribution to the average exposure unit concentration from a single hot spot is:

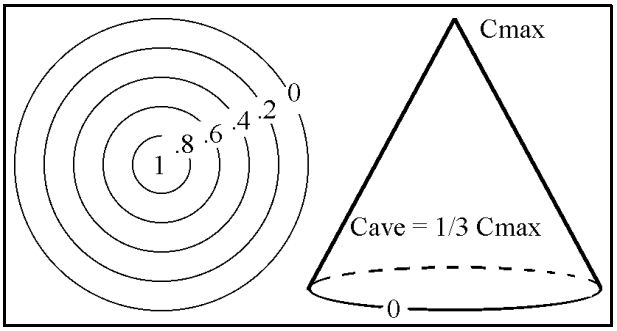

Figure 1. Conceptual Hot Spot Model. Contours of Concentration Normalized with respect to C_{max}.

Gemperline, Mark C. 1994. “Surface Sampling to Detect Hot Spots Presenting Unacceptable Human Health Risk”. Proceedings of the First International Congress on Environmental Geotechnics. BiTech Publishers Ltd., 173 – 11860 Hammershith Way, Richmond, British Columbia, Canado, V7A 5G1.ISBN 0 921095 32 5. pp.233-239.

$$C = C_{ave} \times \frac{A_{hot\,spot}}{A_{unit}} \qquad (2)$$

where:
$A_{hot\ spot}$ = area of the hot spot
A_{unit} = area of the exposure unit
C = average exposure unit concentration.

The smallest hot spot that would cause the average exposure unit concentration to be greater than or equal to C_{index} is sought. The area of this hot spot is obtained by setting C equal to C_{index} in Eq. 2 and rearranging.

$$A_{hot\,spot} = A_{unit} \times \frac{C_{index}}{C_{ave}} \qquad (3)$$

Minimum Number of Samples Required

The number of randomly located surface soil samples is sought which assures that at least one sample will be obtained from within a significant hot spot at a detectable concentration, C_d, or higher. Figure 1 shows the concentration contours within the model hot spot with C_{max} equal to 1. Assume that one of these contours represents the detection limit concentration above which it is desired to detect the hot spot. This value may be equal to, or greater than, the detection limit of a chemical analysis procedure. The circle representing this contour surrounds the detectable area,

$$A1 = A_{hot\,spot} \times \left(1 - \frac{C_d}{C_{\max}}\right)^2 \qquad (4)$$

A1. It is calculated by:
The ratio of this area to the total exposure unit area is the probability that a single sampling event at the site will be selected within the detectable region of the hot spot. The probability of having no successes in N trials, α, is given by:

$$\alpha = \left(1 - \frac{A1}{A_{unit}}\right)^N \qquad (5)$$

Rearranging yields:

$$N = \frac{\log(\alpha)}{\log\left(1 - \frac{A1}{A_{unit}}\right)} \qquad (6)$$

The number of randomly selected samples, N, required to assure with 1-α confidence that at least one sample will be from within the detectable area of a hot spot having unacceptable risk is given by Eq. 6.

Although the analysis assumes a single circular hot spot as the worst case, it also applies if it is desired to assume there are multiple hot spots that are smaller in size, however demonstrate the same cumulative concentration distribution as the model hot spot.

Equations 3, 4, and 6 are combined to yield an equation which gives the general solution for N.

$$\frac{\log(\alpha)}{\text{.og}\left(1 - 3\frac{C_{index}}{C_{\max}}\left(1 - z\frac{C_{inde}}{C_{\max}}\right.\right.} \qquad (7)$$

where:

$$z = \frac{C_d}{C_{index}}$$

Gemperline, Mark C. 1994. “Surface Sampling to Detect Hot Spots Presenting Unacceptable Human Health Risk”. Proceedings of the First International Congress on Environmental Geotechnics. BiTech Publishers Ltd., 173 – 11860 Hammershith Way, Richmond, British Columbia, Canado, V7A 5G1.ISBN 0 921095 32 5. pp.233-239.

Equation 7 is valid when N decreases as the ratio C_{index}/C_{max} approaches 1 and the detectable area of the hot spot is less than or equal to the exposure unit area. It is shown in the appendix that the first condition is satisfied for values of z between 0 and 4/9. The second condition is satisfied when

$$\frac{A1}{A_{unit}} = 3\frac{C_{index}}{C_{\max}}\left(1 - z\frac{C_{index}}{C_{\max}}\right)^2 \quad \textbf{(8)}$$

If $A1/A_{unit}$ is greater than 1 then, by definition, any sample taken should detect contamination.

Values of z greater than 4/9 may be used, however, a staged sampling plan is required.

Staged Sampling Plan

It is often cost effective to use screening technologies to locate hot spots when the hot spot size which must be found is small and the number of sampling locations large. Screening technologies often have higher detection limits than standard laboratory tests. Consequently z may be greater than 4/9. Equation 7 is still applicable, however, it must be realized that a significant hot spot, one that poses an unacceptable risk, may be missed. Such a hot spot would be much larger than the hypothesized smallest significant hot spot however, would have a smaller detectable area and a lower maximum concentration. A second sampling stage, using a lower detection concentration, is required to find this hot spot if it exists.

It is shown in the appendix that the number of samples required in the second sampling stage can be calculated by Eq. 7 letting C_{max2} equal $3C_{d1}$. The numeric term added to the subscript refers to the sample stage. A hot spot having a maximum concentration greater than $3C_{d1}$ and less then C_{max1} will be sampled in the first stage with no less than 1-α confidence.

If z_2 is greater than 4/9 then a third stage is required. Stages are added until z is less than 4/9. Two stages are expected to be sufficient in most cases. Subsequent stages will always require substantially fewer samples.

Decision Units

Often it is desired to sample sub-regions within an exposure unit at different sampling densities. This results from the desire to make independent decisions for these sub-regions, consequently, they are identified as decision units.

Equation 7 can be used to determine the number of samples required in each decision unit. However, C_{index} must be calculated appropriately. C_{indexI} is calculated using Eq. 1 for the Ith decision unit using fractional Rfd^*_I's. Rfd^*_I's must sum to the Rfd* appropriate for the entire exposure unit. The Rfd^*_I values are selected to represent the portion of the total human exposure to contamination which is permitted from a specific decision unit. The portion of total human exposure which may be tolerated in a small decision unit is expected to be greater than the portion tolerated in a larger exposure unit. The addition of decision units always results in more sampling required in the exposure unit.

Normalized Contaminant Concentration

The concept of normalized concentration is introduced to simplify calculations when multiple contaminants are involved. The idea is to convert real contaminant concentration to the equivalent concentration of a user selected standard contaminant so that it reflects an equivalent toxicity or carcinogenic risk. Consequently, calculations are made with a user selected standard contaminant and its corresponding toxicity or carcinogenicity instead of multiple contaminants with varying toxicity.

The following equation is used to calculate Hazard Index for multiple contaminants [1].

$$HI = \frac{M_1}{Rfd_1} + \frac{M_2}{Rfd_2} + \ldots + \frac{}{R}$$

Gemperline, Mark C. 1994. “Surface Sampling to Detect Hot Spots Presenting Unacceptable Human Health Risk”. Proceedings of the First International Congress on Environmental Geotechnics. BiTech Publishers Ltd., 173 – 11860 Hammershith Way, Richmond, British Columbia, Canado, V7A 5G1.ISBN 0 921095 32 5. pp.233-239.

where: Rfd_i = reference dose for contaminant i, and M_i = daily mass of absorbed contaminant i.

This can be rewritten as:

$$HI = \frac{K}{Rfd_1}\left(C_1 + C_2\frac{Rfd_1}{Rfd_2} + \right.$$

where K is a constant associated with exposure and contaminant adsorption. For simplicity K is assumed to be the same for all contaminants. If desired, K can be related to the absorption characteristics of each contaminant resulting in a K_i value associate with each term in parentheses. C_i is the average concentration of contaminant i in the exposure unit. The expression in parentheses is the sum of the normalized concentration of all contaminants with respect to contaminant 1. The Rfd ratios are normalization factors. If all contaminant concentrations are normalized by multiplication by the appropriate normalization factor then they may be treated as the standard contaminant.

The calculation of normalized concentrations for carcinogens is the same as for non-carcinogens except that the Rfd values are replaced with the value of the risk divided by the appropriate slope factor.

Average Exposure Unit Concentration

The presented approach attempts to assure that at least one sampling event occurs in the detectable region of the hot spot. Consequently it is unreasonable to assume that the average of all sample concentrations would guarantee adequate representation of exposure unit average concentration. To assure adequate representation of a hot spot in the determination of average concentration it is recommended that a composite sampling program be conducted in addition to a hot spot detection sampling program.

Composite specimen collection density should guarantee several specimens are taken within a significant hot spot. This will increase the certainty with which the composite concentration, or the average concentration of a group of composites, represents the average exposure unit concentration. The required number of composite sample specimens, M, can be related to N from Eq. 7 by:

$$= \frac{\log\sum\left(\frac{M!}{(i!(M-i)!)}(p)^i(1-(p))^{M-}\right.}{\log(1-(p))} \quad \textbf{(9)}$$

where the summation is with respect to i; i is the desired number of specimens to be collected from within the hot spot minus one; p equals $3C_{index}/C_{max}$. The confidence, $(1-\alpha)$, used to obtain N from Eq. 7 is equal to the confidence used to obtain M from Eq. 9. A short computer program which solves Eq. 9 is presented in a previous paper [2].

Example

The following hypothetical situation exemplifies the use of equations 1 and 2.

An old transformer storage site in the midst of a national park has potential PCBs hot spot contamination. The site is approximately 25 acres and it has been 15 years since transformers were stored there. Historical record review, site inspection, and cursory analyses estimating the extent of the PCBs volatilization and leaching are conducted. It is concluded that the maximum expected value of PCBs, C_{max}, that could exist at this time is approximately 4000 mg/kg. The following values are used with Eq. 1 to approximate C_{index} as 10 mg/kg. An average site concentration greater than this is expected to represent an unacceptable risk.

M ≈ 16 kg
SI ≈ 200 mg/day
DE ≈ 10000 mg/day
SIA ≈ 0.3
DEA ≈ 0.1
t_f ≈ 0.02
Rfd* ≈ 1.25×10^{-5} mg/kg/day

Rfd* was calculated as the acceptable risk divided by the slope factor. These are estimated as:

Acceptable risk = 1×10^{-4}
Slope factor ≈ 8 day/(mg/kg)

A three stage program is used to assure that adequate information is obtained for

Gemperline, Mark C. 1994. “Surface Sampling to Detect Hot Spots Presenting Unacceptable Human Health Risk”. Proceedings of the First International Congress on Environmental Geotechnics. BiTech Publishers Ltd., 173 – 11860 Hammershith Way, Richmond, British Columbia, Canado, V7A 5G1.ISBN 0 921095 32 5. pp.233-239.

risk assessment. Stage 1 will detect small hot spots having high PCBs concentrations whereas stage 2 will find large hot spots having low concentrations. A third stage will consist of a composite sampling program to be used to determine the average site concentration. It is desired to have a minimum of 5 specimens (i=4) obtained from the smallest significant hot spot. Ninety-five percent confidence is desired.

A PCBs surface soil screening using an analytical method with a 50 mg/kg detection limit, C_{d1}, is used for stage 1. A laboratory analytical method with a detection limit of 1 mg/kg, C_d, is used for stages 2 and 3.

Equation 7 is used to calculate that 408 sample locations are required to be screened in stage 1. Stage 2 sampling requirements are calculated using Eq. 7 by setting $C_{max} = 3C_{d1} = 150$ mg/kg and $C_d=C_{d2}$. Stage 2 requires 14 samples be tested using the laboratory analytical procedure.

Equation 9 is used to calculate the stage 3 composite samples specimen requirement. The number of composite specimens required, M, is determined by trial and error. Composite samples should incorporate 1234 specimens.

It is decided to perform stage 1 screening on a 50 ft. grid having a randomly selected origin. This will result in approximately 440 stage 1 point samples. Grid sampling was selected instead of random sampling for simplicity. Samples will be obtained at 14 randomly selected locations to determine PCBs concentration by the laboratory method. This will accommodate stage 2 data needs.

Twenty-five composite samples will be collected each consisting of fifty specimens, i.e. two randomly selected specimens from each of the sites 25 subdivided acres. This amounts to for 1250 stage 3 specimens. The selection of 25 composites samples was arbitrary for this example.

The average of the composite sample specimens is expected to provide a reasonable estimate of the site average concentration even for the condition of a smallest single significant hot spot. If no significant hot spots exist on the site then the calculated mean is expected to be less than C_{index}. The variance of the composites will provide an estimate of the accuracy with which the mean may be estimated. It is considered unreasonable for composite sampling to indicate a mean concentration greater than C_{index} without point samples identifying significant hot spots. If this happens, the conceptual model is probably in error and the sampling program needs would require reevaluation.

The need to characterize individual hot spots for risk assessment purposes will be determined after the proposed sampling and analyses are complete.

Conclusion

Equations are presented which permit the calculation of the number of surface samples required to assure detection of a significant hot spot i.e., a surface hot spot which may pose a threat to human health. The minimum number of samples required for hot spot detection is also related to the number of composite specimen sampling locations required to assure adequate representation of a significant hot spot in the calculation of mean exposure unit concentration. These equations may be modified and adjusted to meet the specific needs of the user.

It is expected that the numerous simplifying assumptions used in equation development will provide a basis for discussion and future improvement. Understanding how variables influence sampling density will assist the user in developing an adequate sampling program.

Equations 7 and 9 may be solved to estimate minimum sampling requirements or used to help understand the representation of a proposed sampling plan.

The equation is expected to be used in a trial-and-error process in which variables and solutions are optimized for site specific needs.

References

[1] United States Environmental Protection Agency (USEPA). 1989. Risk

Gemperline, Mark C. 1994. “Surface Sampling to Detect Hot Spots Presenting Unacceptable Human Health Risk”. Proceedings of the First International Congress on Environmental Geotechnics. BiTech Publishers Ltd., 173 – 11860 Hammershith Way, Richmond, British Columbia, Canado, V7A 5G1.ISBN 0 921095 32 5. pp.233-239.

Assessment Guidance for Superfund, Volume 1, Human Health Evaluation Manual (Part A) Interim Final. Office of Emergency and Remedial Response, Washington D.C. EPA/540/1-89/002.

[2] Gemperline, Mark C. 1993. Surface Sampling to Detect Hot Spots Presenting Unacceptable Human Health Risk, Proceedings of the Superfund XIV Conference, November 30- December 2, 1993, Washington, DC.

Appendix: Equation 7 Limitations

Equation 7 is derived for a hot spot scenario. A hot spot detectable area, A1, greater than the exposure unit area A_{unit} no longer constitutes a hot spot scenario and invalidates Eq. 7. When a hot spot has a detectable area, A1, equal to the total area, A_{unit}, then the probability of detection, $A1/A_{unit}$, becomes 100 percent. The value $A1/A_{unit}$ is given by equation 8. Observe that Eq. 6 is undefined for $A1/A_{unit}$ greater than 1. Equation 7 variables, z and C_{index}/C_{max}, which result in $A1/A_{unit}$ greater than or equal to 1 are graphically illustrated on figure 2.

Intuitively, an upper limit to the acceptable magnitude of the ratio C_d/C_{max} is expected. It must be assured that the detectable hot spot area for all significant hot spots having a maximum concentration less than C_{max} is larger than the detection area of the hot spot containing C_{max}. This is necessary to assure that the probability of detection is greater for significant hot spots having lower maximum concentrations. This condition is satisfied when the detection concentration is less than, or equal to, the average concentration of the largest significant hot spot that does not have a detectable area greater than the exposure unit area. Mathematically, this is deduced by noting that the derivative of Eq. 8 with respect to the ratio C_d/C_{max} equals zero when $C_d/C_{max} = 1/3$. This is the condition in which $A1/A_{unit}$ reaches a maximum value. The following sequence shows this analysis.

Equation 8 is rewritten as:

$$f(w,z) = 3\frac{w}{z}(1-w)^2$$

where $f(w,z) = A1/A_{unit}$, $w = C_d/C_{max}$ and $z = C_d/C_{index}$.

The derivative of f(w,z) with respect to w is given by:

$$\frac{df}{dw} = \frac{3}{z}(1-w)^2 - 6\frac{w}{z}(1-w)$$

solving for df/dw = 0 yields:

$$\frac{df}{dw} = 0 \quad when \quad w = \frac{1}{3}$$

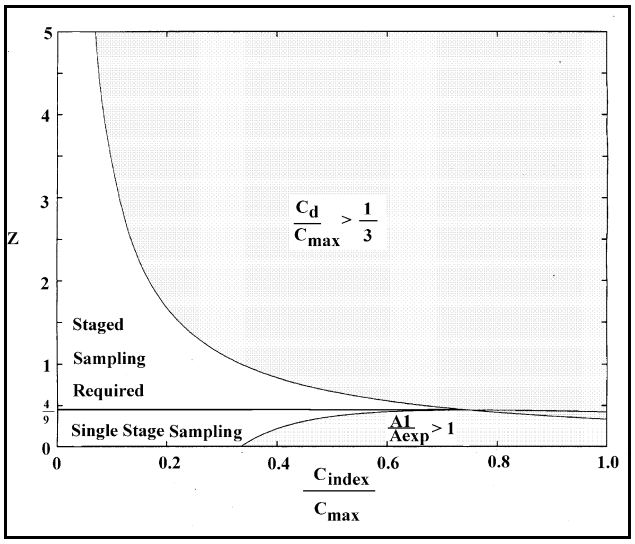

Figure 2. Variable combinations affecting use of Eq. 7.

The region in which C_d/C_{max} is greater than 1/3 is indicated on figure 2.

The intersection of the lines marking the boundary of the regions $C_d/C_{max} > 1/3$ and $A1/A_{unit} > 1$, occurs at z=4/9 and provides the maximum value of z for which equation 7 may be applied with one contingency, i.e., $A1/A_{unit} < 1$. It is assured that any significant hot spot having a maximum concentration less than C_{max} will have a larger detectable area than the significant hot spot containing C_{max}. Consequently the number of samples required to assure detection of a significant hot spot containing C_{max} will assure with greater confidence the

Gemperline, Mark C. 1994. “Surface Sampling to Detect Hot Spots Presenting Unacceptable Human Health Risk”. Proceedings of the First International Congress on Environmental Geotechnics. BiTech Publishers Ltd., 173 – 11860 Hammershith Way, Richmond, British Columbia, Canado, V7A 5G1.ISBN 0 921095 32 5. pp.233-239.

detection of all other possible significant hot spots.

If Eq. 7 is used with z greater than 4/9 then there is inadequate assurance that all possible significant hot spots will be found. Significant hot spots having an average concentration, C_{ave}, below the detection concentration, C_d, experience a shrinking detectable area as the significant hot spot maximum concentration approaches Cd. Consequently, the probability of finding hot spots becomes increasingly lower.

Significant hot spots which are missed by using C_d/C_{index} greater than 4/9 can be found with a second sampling effort. Equation 7 may be used to calculate the number of required samples knowing that the maximum concentration of a significant hot spot which may have been missed in the previous sampling effort is not greater than three times the detection concentration used in previous sampling. A significantly fewer number of samples are required for the second sampling effort. The second effort uses a lower detection concentration with a lower C_{max}. If C_d/C_{index} for the second sampling set is less than 4/9 then it is assured with $1-\alpha$ confidence that all significant hot spots in the exposure unit having a maximum concentration less than C_{max} have been found. If z is greater than 4/9 then a third sampling effort, designed by the same procedure as the second, is required.