Determining Soil Contamination Profiles from Intensities of Capture-Gamma Rays Using Above-Surface Neutron Sources

by

J. Kenneth Shultis, F. Khan, B. Letellier†, and Richard E. Faw

Department of Mechanical and Nuclear Engineering,
Kansas State University, Manhattan, KS 66506-2503

†*Present address: Los Alamos National Laboratory, Los Alamos, NM*

submitted Nov. 20, 1997 to

Applied Radiation and Isotopes

revised Dec. 19, 1997

(37 pages, including 1 Tables and 19 Figures)

Abbreviated title: **Contaminant Profiles by Capture-Gamma Analysis**

Send correspondence to: Prof. J. Kenneth Shultis
Ward Hall
Kansas State University
Manhattan, KS 66506-2503

VOICE: (785) 532-5626
FAX: (785) 532-6952
e-mail jks@ksu.edu

Determining Soil Contamination Profiles from Intensities of Capture-Gamma Rays Using Above-Surface Neutron Sources

J. Kenneth Shultis, F. Khan, B. Letellier[†], and Richard E. Faw,

Department of Mechanical and Nuclear Engineering,
Kansas State University, Manhattan, KS 66506-2503

[†]*Present address: Los Alamos National Laboratory, Los Alamos, NM*

Abstract

Four methods are presented to estimate contaminant concentration profiles in soil from the intensities of neutron-induced capture-gamma photon intensities measured at the soil surface. In particular, the method of linear regularization with and without an iterative positivity constraint, the Backus-Gilbert method, and the maximum entropy method are applied to the soil contamination problem. Example results obtained with the four methods are given for photon intensities calculated for idealized test contaminant profiles in soil irradiated by neutron sources above the surface.

1 Introduction

In the technique of prompt-gamma neutron-activation analysis (PGNAA), soil is irradiated with fast neutrons, usually from above the surface, and the subsequent capture gamma photons produced when these neutrons are absorbed by elemental contaminants in the soil are measured at, or slightly above, the surface. For a given contaminant, capture photons with typically many distinct characteristic energies are emitted. The uncollided capture photons that are emitted by each element and that reach the surface are measured. With these measured intensities and a PGNAA model that expresses photon intensities in terms of the concentration profile, the model can, in principle, be "inverted" to estimate the concentration of a particular contaminant as a function of depth.

To estimate soil contamination profiles from the intensities of uncollided gamma photons measured above the soil surface, it is first necessary to develop a mathematical model that relates contaminant concentration profiles to the detector measurements. In the next section idealized soil irradiation geometries are considered and the fundamental relationship between the detector measurement and the soil contaminant profile is derived. In later sections four modern approaches for inverting the PGNAA model to obtain the contaminant profile from measurements of uncollided capture gamma intensities at the soil surface are summarized and example results given.

2 PGNAA Soil Contamination Model

We assume that the soil surface is a horizontal plane and that the bulk soil composition is uniform. The contaminant concentration varies only with depth z into the soil, i.e., no lateral variations are considered. Further, the contaminant is assumed to be sufficiently dilute that it has negligible influence on the neutron field in the soil. The capture-gamma detector is assumed to be a point isotropic detector which may be arbitrarily collimated (i.e., only radiation incident from certain directions can reach the detector). In this section a quite general PGNAA model is developed and, for two special irradiation geometries, the model is shown to simplify considerably.

2.1 Capture-Gamma Photon Source Strengths

Central to the development of any PGNAA mathematical model that relates a contaminant profile to detector measurements is the determination of the number of capture-gamma photons emitted throughout the soil exposed to neutrons.

Soil positions are defined in a cylindrical coordinate system with the vertical axis through the detector, which is located at elevation h_d above the soil surface at $z = 0$. At a position in the soil at radial distance r, depth z, and azimuthal angle ψ, the energy-dependent neutron fluence produced by the neutron source is denoted by $\Phi(r, z, \psi, E)$. If the atomic concentration of the contaminant of concern is denoted by $u(z)$ (u for *unknown*), the number of neutrons absorbed by this contaminant, per unit soil volume, at location (r, z, ψ) is

$$N_{abs}(r, z, \psi) = u(z) \int_0^{E_{\max}} dE\, \Phi(r, z, \psi, E)\sigma_c(E), \tag{1}$$

where $\sigma_c(E)$ is the microscopic absorption cross section of the contaminant element for neutrons with energy E, and $E_{\max}$ is the maximum energy of the incident neutrons.

To avoid the necessity of using the energy-dependent fluence in the PGNAA model, this neutron absorption can be expressed in terms of the thermal neutron fluence $\Phi_{th}(r,z,\psi)$, obtained by integrating $\Phi(r,z,\psi,E)$ over all thermal energies. Toward this end, define an effective capture cross section $\overline{\sigma}_c$ as

$$\overline{\sigma}_c \equiv \frac{1}{\Phi_{th}(r,z,\psi)} \int_0^{E_{\max}} dE\, \Phi(r,z,\psi,E)\sigma_c(E). \tag{2}$$

For most contaminants the effective capture cross section is remarkably close to the actual thermal absorption cross section, a good indication that most neutron absorption occurs only after thermalization [Shue and Faw 1996; Shue et al. 1996]. Moreover, it has been found that the shape of the epithermal neutron spectrum is almost independent of depth so that $\overline{\sigma}_c$ also is independent of depth [Shue and Faw 1996; Shue et al. 1996]. With this effective capture cross section, the neutron absorption N_{abs} can then be computed knowing only the thermal fluence, namely $N_{abs}(r,z,\psi) = \overline{\sigma}_c u(z)\Phi_{th}(r,z,\phi)$.

For each thermal neutron absorbed by the contaminant element, f_i capture photons of energy E_i are emitted. Capture-photon yields for nonthermal neutrons are usually not known, and in this study are assumed to equal those for thermal neutron absorption. Moreover, since most neutrons are absorbed only after thermalization, the use of thermal-capture yields for all neutron energies is a reasonable approximation. Finally, the number $S_i(r,z,\psi)$ of capture photons with energy E_i that are emitted per unit soil volume at position (r,z,ψ) as a result of neutron capture in the contaminant is give by

$$S_i(r,z,\psi) = f_i\, \overline{\sigma}_c\, u(z)\Phi_{th}(r,z,\psi). \tag{3}$$

Before the source strength of capture-gamma photons can be determined from Eq. (3), the thermal fluence $\Phi_{th}(r,z,\psi)$ must first be determined. In a companion paper, empirical expressions and tabular data, based on detailed neutron transport calculations, are presented for the thermal fluence in five representative soils with several practical neutron sources [Shue and Faw 1996; Shue et al. 1997]. Example fluence profiles for uniform and perpendicular illumination of the soil surface by 14-MeV neutrons are shown in Fig. 1.

2.2 A General PGNAA Model

Consider a differential soil volume $dV = r\,dr\,dz\,d\psi$ in which the density of emitted photons of energy E_i is S_i. Let the *attenuation kernel* $\mathcal{G}_i(r,z,\psi)$ be the probability a photon of energy E_i emitted at (r,z,ψ) results in a measured full-energy (photopeak) interaction in the gamma-ray spectrometer detector. Since the soil composition is assumed to be uniform and the detector is assumed to respond isotropically, $\mathcal{G}_i$ is independent of ψ. Then the *expected* number of counts c_i measured by the detector caused by photons of energy E_i emitted in dV is simply $\mathcal{G}_i(r,z)S_i(r,z,\psi)\,dV$. The expected total counts c_i from capture-gamma photons emitted everywhere in the soil is thus

$$c_i = \int_0^{\infty} dz \int_0^{2\pi} d\psi \int_0^{r_{\max}(z,\psi)} dr\, r\, \mathcal{G}_i(r,z) S_i(r,z,\psi), \tag{4}$$

where the radial limit $r_{\max}(z,\psi)$ depends on the position and collimation of the gamma-ray detector. Finally, substitution of Eq. (3) into Eq. (4) and accounting for noise and statistical uncertainty in the detection process shows that the *observed* counts c_i^{obs} for the ith capture-gamma photon is related to the contaminant concentration profile $u(z)$ by the following Fredholm integral equation of the first kind

$$c_i = c_i^{obs} \pm n_i = \int_0^{\infty} dz\, u(z)\mathcal{R}_i(z), \tag{5}$$

where $n_i \geq \sqrt{c_i^{obs}}$ is the contribution of noise to the ith gamma-photon measurement and the *emission kernel* $\mathcal{R}_i(z)$ is

$$\mathcal{R}_i(z) \equiv f_i\,\overline{\sigma}_c \int_0^{2\pi} d\psi \int_0^{r_{\max}(z,\psi)} dr\, r\, \mathcal{G}_i(r,z)\Phi_{th}(r,z,\psi). \tag{6}$$

The key problem of PGNAA soil contamination analysis is to solve or invert Eq. (5) for $u(z)$ given the counts $c_i,\ i = 1,\ldots,N$. For simplicity, in this paper the noise in the observed counts is assumed negligible so that $c_i \simeq c_i^{obs}$. However, the methods presented here can be readily combined with appropriate noise models.

2.3 PGNAA Models for Uncollided Capture Photons

In most gamma-ray activation analyses, only the uncollided photons reaching the detector are considered. Although contributions to c_i from scattered photons have been proposed [Letellier 1997] only uncollided contributions are considered here. For uncollided photons,

the attenuation kernel is

$$\mathcal{G}_i(r,z) = \frac{\eta_i}{4\pi[r^2+(z+h_d)^2]} \exp(-\mu_i z \sec\theta), \tag{7}$$

where η_i is the expected number of counts contributing to c_i per unit fluence of photons with energy E_i at the detector, $\mu_i \equiv \mu(E_i)$ is the total attenuation coefficient (less the coherent scattering component) in soil for a photon of energy E_i, and θ is the angle between the vertical and the line from the detector to the emission point (r,z,ψ), given by

$$\sec\theta = \frac{\sqrt{r^2+(z+h_d)^2}}{(z+h_d)}. \tag{8}$$

For certain sources, the neutron field is axially symmetric, i.e. independent of ψ. For example, a point isotropic source at any location on the detector axis or uniform perpendicular illumination of the surface produces such axial symmetry in $\Phi_{th}(r,z)$. If in addition, the collimated detector has an aperture half-angle of θ_d with respect to the vertical axis, the upper limit for the integral of r in Eq. (6) becomes $r_{\max} = (z+h_d)\tan\theta_d$. To show explicitly the dependence of the fluence on the strength of the irradiating neutron source Q_o, the thermal fluence is written as $\Phi_{th}(r,z,\psi) \equiv Q_o \widehat{\Phi}_{th}(r,z,\psi)$. For a point source $Q_o = S_o$, the total number of neutrons emitted by the source. If the source is deeply buried, then $\widehat{\Phi}_{th}$ becomes spherically symmetric about the source (i.e., the air-ground interface has negligible effect on the fluence profile), and $\widehat{\Phi}_{th}$ is a function of only the distance from the source $\ell \equiv \sqrt{r^2+|z+h_s|^2}$ where h_s is the source elevation above the surface (negative for a buried source). With these simplifications, the emission kernel of Eq. (6) can be written as

$$\mathcal{R}_i(z) = f_i \eta_i Q_o \overline{\sigma}_c \widehat{\mathcal{R}}_i(z), \tag{9}$$

where the dimensionless, normalized, emission kernel is

$$\widehat{\mathcal{R}}_i(z) \equiv \int_0^{(z+h_d)\tan\theta_d} dr \, \frac{r\,\widehat{\Phi}_{th}(r,z)}{2[r^2+(z+h_d)^2]} \exp(-\mu_i z \sec\theta). \tag{10}$$

For uniform and normal illumination of the surface by a beam of neutrons $Q_o = J_o$, the beam intensity (number incident per unit surface area). With this source, the fluence is a function only of soil depth and the integral in Eq. (10) can be performed analytically to give

$$\widehat{\mathcal{R}}_i(z) = \tfrac{1}{2}[E_1(\mu_i z) - E_1(\mu_i z \sec\theta_d)]\widehat{\Phi}_{th}(z), \tag{11}$$

where E_1 is the exponential integral function of order 1.

To evaluate the emission kernel $\widehat{\mathcal{R}}_i$ from Eq. (10) or Eq. (11), the thermal fluence profile must first be determined. Recently, detailed calculations of $\widehat{\Phi}_{th}$ profiles in five representative soils have been performed for 14-MeV deuterium-tritium accelerator neutrons and for neutrons emitted by ^{241}Am/Be and ^{252}Cf sources [Shue et al. 1997]. Empirical approximations have also been obtained for $\widehat{\Phi}_{th}(z)$ for wide-beam normal illumination of the soil, so that Eq. (11) may be evaluated at any z. For point sources, detailed tabulations of $\widehat{\Phi}_{th}(r,z)$ are available, which, together with appropriate interpolation schemes, allow the integration in Eq. (10) to be performed numerically [Shue and Faw 1996].

2.4 Discretization of the PGNAA Model

The solution of Eq. (5) for the contaminant concentration profile is an inversion problem encountered in many diverse fields. For example, this same inversion problem, but with different kernels, is encountered in oil-well logging, neutron scattering, geophysical data analysis, atmospheric remote sensing, astrophysics, medical tomography, and many other data analysis applications. The principal difficulty with solving this equation for the unknown profile $u(z)$ in the PGNAA soil contamination problem is that the number N of different photopeak counts c_i is generally many fewer than the number of depths at which one would like to determine $u(z)$. Although heavy-atom contaminants typically emit dozens or even hundreds of capture-gamma photons with different energies, most are emitted with such low probabilities or yields that accurate measurement of their intensities is not possible. Typically only 2 to 8 capture-gamma photons have sufficiently high yields and unique energies so that they can be measured reliably against the background of multiply scattered gamma photons produced in the normal soil constituents. Consequently, in practical applications of the PGNAA method, N is almost always a small integer.

In many solution schemes, the Fredholm integral equation of Eq. (5) is approximated first by a set of linear algebraic equations for the concentration $u_j \equiv u(z_j)$ at M specified depths z_j by use of some appropriate numerical quadrature scheme to approximate the integral. The resulting linear equations can be written as

$$c_i = \sum_{j=1}^{M} R_{ij} u_j, \quad i = 1, ..., N, \tag{12}$$

or in matrix notation $\mathbf{c} = \mathbf{R}\mathbf{u}$. This set represents N equations in the M unknowns $\{u_j\}$. Here the $N \times M$ matrix $\mathbf{R}$ depends on the numerical quadrature approximation selected.

Two possible schemes are presented below.

2.4.1 A Piece-Wise Linear Approximation

Suppose the contaminant concentration is negligible beyond a certain depth $z_{\max} = z_M$ so that Eq. (5) can be approximated by

$$c_i = \int_0^{\infty} dz\, \mathcal{R}_i(z) u(z) \simeq \int_0^{z_{\max}} dz\, \mathcal{R}_i(z) u(z) = \sum_{j=1}^{M-1} \int_{z_j}^{z_{j+1}} dz\, \mathcal{R}_i(z) u(z), \tag{13}$$

where $z_{j+1} > z_j$ and $z_1 = 0$. Further, assume $u(z)$ varies linearly between its values at the endpoints of each subinterval, i.e.,

$$u(z) \simeq (z_{j+1} - z) u_j / \Delta_j + (z - z_j) u_{j+1} / \Delta_j, \quad z_j \le z \le z_{j+1}, \tag{14}$$

where $\Delta_j \equiv (z_{j+1} - z_j)$. Substitution of this linear approximation into Eq. (13) gives Eq. (12) where the R_{ij} are given by

$$R_{ij} = \begin{cases} \displaystyle\int_{z_1}^{z_2} dz\, f_1(z) \mathcal{R}_i(z), & j = 1 \\ \displaystyle\int_{z_{j-1}}^{z_{j+1}} dz\, f_j(z) \mathcal{R}_i(z), & j = 2, \ldots, M-1 \\ \displaystyle\int_{z_{M-1}}^{z_M} dz\, f_M(z) \mathcal{R}_i(z), & j = M \end{cases} \tag{15}$$

where the weighting functions $f_j(z)$ are defined as

$$f_j(z) = \begin{cases} (z - z_{j-1})/\Delta_{j-1}, & z_{j-1} \le z < z_j \\ (z_{j+1} - z)/\Delta_j, & z_j \le z < z_{j+1} \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

The integrals in Eq. (15) generally must be evaluated using numerical integration.

2.4.2 A Piece-Wise Quadratic Approximation

Equation (5) can also be approximated by

$$c_i \simeq \int_0^{z_{\max}} dz\, \mathcal{R}_i(z) u(z) = \sum_{j=1}^{M-2}{}' \int_{z_j}^{z_{j+2}} dz\, \mathcal{R}_i(z) u(z), \tag{17}$$

where the prime on the summation indicates that the summation is over only odd values of j. For the approximation developed here, M is assumed odd. Now approximate $u(z)$ in

each pair of adjacent subintervals by a quadratic function. For equally spaced nodes with $\Delta = z_{j+1} - z_j$, $u(z)$ in interval (z_j, z_{j+2}) is approximated by

$$u(z) \simeq \frac{(z - z_{j+1})(z - z_{j+2})}{2\Delta^2} u_j - \frac{(z - z_j)(z - z_{j+2})}{\Delta^2} u_{j+1} + \frac{(z - z_j)(z - z_{j+1})}{2\Delta^2} u_{j+2}. \quad (18)$$

Substitution of this result into Eq. (17) gives Eqs. (12) where the R_{ij} are now given by

$$R_{ij} = \frac{1}{2\Delta^2} \begin{cases} \displaystyle\int_{z_1}^{z_3} dz\,(z - z_2)(z - z_3)\mathcal{R}_i(z), & j = 1 \\ \displaystyle\int_{z_{j-2}}^{z_j} dz\,(z - z_{j-2})(z - z_{j-1})\mathcal{R}_i(z) & \\ \qquad + \displaystyle\int_{z_j}^{z_{j+2}} dz\,(z - z_{j+1})(z - z_{j+2})\mathcal{R}_i(z), & j \text{ odd},\ j \neq 1, M \\ -2\displaystyle\int_{z_{j-1}}^{z_{j+1}} dz\,(z - z_{j-1})(z - z_{j+1})\mathcal{R}_i(z), & j \text{ even} \\ \displaystyle\int_{z_{M-2}}^{z_M} dz\,(z - z_{M-2})(z - z_{M-1})\mathcal{R}_i(z), & j = M \end{cases} \quad (19)$$

Other quadrature schemes yield slightly different approximations for R_{ij} [Shultis et al. 1996].

3 Inversion Methods

The inversion of Eqs. (12) is an ill-posed problem since the number of unknowns M (the u_j) is generally greater than the number N of data (the c_i). This means there are an infinite number of solutions because the solution space (of dimension M) has an $(M - N)$ dimensional degeneracy, i.e., any $(M - N)$ components of $\mathbf{u}$ can be specified arbitrarily and still have Eq. (12) satisfied.

One might be tempted to solve Eq. (12) by minimizing the difference between some model $\widehat{\mathbf{u}}(z)$ and the measured data. This difference between a model and measured data is often quantified by the χ^2 statistic, namely,

$$\chi^2 = \sum_{i=1}^{N}\sum_{j=1}^{M} \left[c_i - \sum_{k=1}^{M} R_{ik}\widehat{u}_k\right] S_{ij}^{-1} \left[c_j - \sum_{k=1}^{M} R_{jk}\widehat{u}_k\right] \quad (20)$$

$$\simeq \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \left[c_i - \sum_{k=1}^{M} R_{ik}\widehat{u}_k\right]^2 = |\mathbf{A}\,\widehat{\mathbf{u}} - \mathbf{b}|^2. \quad (21)$$

Here $S_{ij} = \text{Covar}[n_i, n_j]$ are the elements of the covariance matrix. The approximate equality in the above result holds if we can neglect the off-diagonal covariance terms, with

$\sigma_i^2 = \text{Covar}[n_i, n_i]$. The matrix $\mathbf{A}$ has elements $A_{ij} = R_{ij}/\sigma_i$ and the vector $\mathbf{b}$ has elements $b_i = c_i/\sigma_i$. For the counting data of uncollided capture-gamma photons used in the soil contamination problem, the estimate of σ_i is $\sqrt{c_i}$ provided c_i is sufficiently large (namely, $c_i \gtrsim 20$).

However, the minimization of the positive functional $\mathcal{A}[\hat{\mathbf{u}}] \equiv \chi^2 = |\mathbf{A}\hat{\mathbf{u}} - \mathbf{b}|^2$ for a matrix $\mathbf{A}$ that is degenerate, i.e., has fewer rows than columns, will not give a unique solution for $\mathbf{u}$. To obtain a unique solution, additional constraints must be imposed on the minimization problem. For example, if any non-degenerate strictly convex functional $\mathcal{B}[\hat{\mathbf{u}}]$, for example $\hat{\mathbf{u}}^T H \hat{\mathbf{u}}$, is added, then the minimization of $\mathcal{A}[\hat{\mathbf{u}}] + \lambda\mathcal{B}[\hat{\mathbf{u}}]$ *will* produce a unique solution $\hat{\mathbf{u}}$ [Press et al. 1992]. The addition of the term $\lambda\mathcal{B}[\hat{\mathbf{u}}]$ is said to "regularize" the minimization problem, i.e., to produce a unique solution.

Thus in the inverse problem, to obtain a unique solution for $\mathbf{u}$, one solves the following minimization problem

$$\text{minimize:} \quad \mathcal{A}[\hat{\mathbf{u}}] + \lambda\mathcal{B}[\hat{\mathbf{u}}]. \tag{22}$$

This is the central principle of inversion theory. As the Lagrange multiplier λ varies from 0 to ∞, the unique solution $\hat{\mathbf{u}}$ varies from one minimizing $\mathcal{A}[\hat{\mathbf{u}}]$ to one minimizing $\mathcal{B}[\hat{\mathbf{u}}]$. To obtain the "best" solution (corresponding to a particular value of λ) one must choose a particular criterion. For example, one might pick λ so that $\chi^2 = N$ to agree with the expected value of χ^2. Alternatively, one might pick λ purely subjectively so as to produce, for example, a "smooth" solution or a solution sensitive to abrupt changes in the profile $u(z)$. Finally, for simulated count data obtained by accurate numerical integration of Eq. (12), the most accurate inversion will be obtained with λ made as small as possible, but still large enough to avoid numerical instabilities in the minimization algorithm.

The many apparently different approaches used for inversion problems by the regularization technique all involve minimizing the functional of Eq. (22) with the choice for $\mathcal{A}[\hat{\mathbf{u}}]$ and $\mathcal{B}[\hat{\mathbf{u}}]$ dependent on the problem and the inversion philosophy. We next summarize four widely used inversion methods, all of which use the above regularization approach.

3.1 The Linear Regularization (LR) Method

The linear regularization method goes by many names, for example, Tikhonov-Miller regularization [Tikhonov 1964; Tikhonov and Arsenin 1977; Miller 1970; Biemond et al. 1990], the Phillips-Twomey method [Phillips 1962; Twomey 1963], the constrained linear inversion

method [Twomey 1977], and the method of regularization [Craig and Brown 1986].). As with any method that has evolved from many different disciplines, the notation and ideas in the many seminal works are often quite different. In the summary of this and the other inversion methods discussed in this paper, we adhere closely to the notation of Press et al. [1992].

In the linear regularization approach, the functional $\mathcal{A}[\hat{\mathbf{u}}]$ of Eq. (22) is taken as the χ^2 of Eq. (21), i.e., $\mathcal{A}[\hat{\mathbf{u}}] = |\mathbf{A}\hat{\mathbf{u}} - \mathbf{b}|^2$, and the functional $\mathcal{B}[\hat{\mathbf{u}}]$ is chosen as some measure of the smoothness of $u(z)$, which is derived from first or higher derivatives of $u(z)$. In particular, the linear regularization method requires that $\mathcal{B}[\hat{\mathbf{u}}] = \hat{\mathbf{u}}^T\mathbf{H}\hat{\mathbf{u}}$ where $\mathbf{H}$ is some appropriate symmetric smoothing matrix. The inversion solution is thus determined by the following minimization problem:

$$\text{minimize:} \quad \mathcal{A}[\hat{\mathbf{u}}] + \lambda\mathcal{B}[\hat{\mathbf{u}}] = |\mathbf{A}\hat{\mathbf{u}} - \mathbf{b}|^2 + \lambda\hat{\mathbf{u}}^T\mathbf{H}\hat{\mathbf{u}}. \tag{23}$$

The matrix $\mathbf{H}$ is obtained by making some *a priori* assumption about the nature of the profile $u(z)$. An example is given in Section 3.1.1.

To obtain the minimum of the functional of Eq. (23) and find $\hat{\mathbf{u}}$, we write Eq. (23) in its component form as

$$\mathcal{F}[\hat{\mathbf{u}}] \equiv \mathcal{A}[\hat{\mathbf{u}}] + \lambda\mathcal{B}[\hat{\mathbf{u}}] = \sum_{i=1}^{N}\left[\sum_{j=1}^{M} A_{ij}\hat{u}_j - b_i\right]^2 + \lambda\sum_{i=1}^{M}\hat{u}_i\sum_{j=1}^{M} H_{ij}\hat{u}_j. \tag{24}$$

The values of $\hat{u}_j$ that minimize this functional are the solutions of the M *normal equations* obtain by setting the derivative of $\mathcal{F}[\hat{\mathbf{u}}]$ with respect to $\hat{u}_j$ to zero. Differentiation of Eq. (24) with respect to $\hat{u}_j$, setting the result to zero, and use of the symmetry property of $\mathbf{H}$ gives

$$\sum_{j=1}^{M}\left\{\left(\sum_{i=1}^{N} A_{ik}A_{ij}\right) + \lambda H_{kj}\right\}\hat{u}_j = \sum_{i=1}^{M} A_{ik}b_i, \quad k = 1, ..., M, \tag{25}$$

or, in matrix form,

$$(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{H})\hat{\mathbf{u}} = \mathbf{A}^T\mathbf{b}. \tag{26}$$

This set of M linear algebraic equations is readily solved for $\hat{\mathbf{u}}$ using standard techniques such as the Lower-Upper (LU) decomposition method or the Singular Value Decompostion (SVD) method [Press et al. 1992].

3.1.1 Smoothing Matrices

The construction of the symmetric $M \times M$ matrix $\mathbf{H}$ depends on the smoothness criterion chosen. For example, if one believes that $u(z)$ is approximately quadratic, then a reasonable functional to minimize so as to enforce this belief is (assuming equi-spaced values of z_j and using forward finite differences)

$$\mathcal{B}[\hat{\mathbf{u}}] \propto \int_0^\infty [\hat{u}'''(z)]^2 \, dz \propto \sum_{j=1}^{M-3} [-\hat{u}_j + 3\hat{u}_{j+1} - 3\hat{u}_{j+2} + \hat{u}_{j+3}]^2. \tag{27}$$

Note that this functional is nonnegative and vanishes only when $\hat{u}(z)$ is a polynomial of degree less than three. The constant of proportionality can be absorbed into the parameter λ so that the discretized form of $\mathcal{B}[\hat{\mathbf{u}}]$ can be written as

$$\mathcal{B}[\hat{\mathbf{u}}] = |\mathbf{B}\hat{\mathbf{u}}|^2 = \hat{\mathbf{u}}^T(\mathbf{B}^T\mathbf{B})\hat{\mathbf{u}} \equiv \hat{\mathbf{u}}^T\mathbf{H}\hat{\mathbf{u}}, \tag{28}$$

where $\mathbf{B}$ is the $(M-3) \times M$ first-order, forward finite-difference matrix

$$\mathbf{B} = \begin{pmatrix} -1 & 3 & -3 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & -1 & 3 & -3 & 1 & 0 & 0 & \dots & 0 \\ \vdots & & & & \ddots & & & & \vdots \\ 0 & \dots & 0 & 0 & -1 & 3 & -3 & 1 & 0 \\ 0 & \dots & 0 & 0 & 0 & -1 & 3 & -3 & 1 \end{pmatrix}. \tag{29}$$

The same procedure can be used to construct $\mathbf{B}$ (and hence $\mathbf{H}$) for any finite difference representation of the function $\hat{u}(z)$, such as a discretized differential equation or polynomial of arbitrary degree. In this study smoothing functionals based on low-order (< 5) polynomial approximations to $\hat{u}(z)$ are used. Higher order approximations, while capable of describing complex profile shapes, tend to produce unrealistic spurious oscillations in simple profiles.

3.2 Constrained Linear Regularization (CLR) Method

Often there are physical constraints on the unknown $\hat{u}(z)$ which should also be incorporated into the inversion process. For example, one may want $\hat{u}(z) \geq 0$ or $u_L(z) \leq \hat{u}(z) \leq u_U(z)$ for specified bounding functions u_L and u_U. In the PGNAA problem, the concentration profile $\hat{u}(z)$ clearly must be non-negative. The method of *projections onto convex sets* [Biemond et al. 1990; Press et al. 1992] easily imposes such deterministic constraints if an iterative solution of the functional minimization problem is used.

Many iterative schemes can be used to find the $\hat{u}(z)$ that minimizes the functional $\mathcal{A}[\hat{\mathbf{u}}] + \lambda\mathcal{B}[\hat{\mathbf{u}}]$. An unsophisticated approach is to use the method of steepest descent, whereby the minimum of $\mathcal{A}[\hat{\mathbf{u}}] + \lambda\mathcal{B}[\hat{\mathbf{u}}]$ is approached by proceeding from some arbitrary starting point in $\hat{\mathbf{u}}$ space by taking small steps always in the direction opposite the gradient of $\mathcal{A}[\hat{\mathbf{u}}] + \lambda\mathcal{B}[\hat{\mathbf{u}}]$, i.e., downhill. Mathematically, the iteration search is

$$\hat{\mathbf{u}}^{(k+1)} = \hat{\mathbf{u}}^{(k)} - \epsilon\nabla(\mathcal{A}[\hat{\mathbf{u}}] + \lambda\mathcal{B}[\hat{\mathbf{u}}]), \tag{30}$$

where ϵ is a parameter that determines how far to move downhill in each step. For the linear regularization method, based on minimizing the functional $\mathcal{A}[\hat{\mathbf{u}}] + \lambda\mathcal{B}[\hat{\mathbf{u}}] = |\mathbf{A}\hat{\mathbf{u}} - \mathbf{b}|^2 + \lambda\hat{\mathbf{u}}^T\mathbf{H}\hat{\mathbf{u}}$ (see Eq. (23)), the minimization iteration scheme becomes

$$\begin{aligned}
\hat{\mathbf{u}}^{(k+1)} &= \hat{\mathbf{u}}^{(k)} - \epsilon\nabla(|\mathbf{A}\hat{\mathbf{u}} - \mathbf{b}|^2 + \lambda\hat{\mathbf{u}}^T\mathbf{H}\hat{\mathbf{u}}) \\
&= \hat{\mathbf{u}}^{(k)} - 2\epsilon[(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{H})\hat{\mathbf{u}} - \mathbf{A}^T\mathbf{b}] \\
&= [\mathbf{1} - \epsilon(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{H})]\hat{\mathbf{u}}^{(k)} + \epsilon\mathbf{A}^T\mathbf{b},
\end{aligned} \tag{31}$$

where, to guarantee convergence [Press et al. 1992],

$$0 < \epsilon < \frac{1}{\text{max eigenvalue}\,(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{H})}. \tag{32}$$

The converged solution $\lim_{k\to\infty}\hat{\mathbf{u}}^{(k)}$ will be the same as the LR solution obtained from Eq. (26).

To impose a non-negativity constraint on this iterative solution, define $\mathcal{P}$ as the projection operator that sets to zero all negative components of $\hat{\mathbf{u}}$. Then modify Eq. (31) to

$$\hat{\mathbf{u}}^{(k+1)} = \mathcal{P}\{[\mathbf{1} - \epsilon(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{H})]\hat{\mathbf{u}}^{(k)} + \epsilon\mathbf{A}^T\mathbf{b}\}, \tag{33}$$

ensuring that, after each iteration, any negative components of $\hat{\mathbf{u}}$ are set to zero. The converged solution of this equation is the vector with nonnegative components that minimizes $\mathcal{A}[\hat{\mathbf{u}}] + \lambda\mathcal{B}[\hat{\mathbf{u}}]$ using the linear regularization method.

3.3 The Backus-Gilbert (BG) Method

The Backus-Gilbert method [Backus and Gilbert 1967, 1968; Parker 1977; Loreda and Epstein 1989] is different from the linear regularization method in its special selection of the functionals $\mathcal{A}$ and $\mathcal{B}$. For $\mathcal{B}$ this method chooses a functional to ensure *stability* of the

solution $\widehat{u}(z)$ rather than its smoothness. Specifically, the BG method uses a regularization functional that seeks to minimize the sensitivity of the solution $\widehat{u}$ to errors in the data. As a measure of how much the solution $\widehat{u}(z)$ varies as the data vary within their measurement errors, the variance of the estimated profile $\widehat{u}(z)$ is used, i.e., $\mathcal{B} \equiv \mathrm{var}[\widehat{u}(z)]$.

The Backus-Gilbert method also attempts to make $\widehat{u}(z)$ reproduce any sudden changes in $u(z)$ as faithfully as possible. Since the method is linear, the estimated profile $\widehat{u}(z)$ is related to the true profile $u(z)$ by a linear mapping

$$\widehat{u}(z) = \int_0^\infty dz' \, \widehat{\delta}(z', z) u(z'), \tag{34}$$

where $\widehat{\delta}(z', z)$ is an *averaging kernel*. The BG method also attempts, in the limit of negligible measurement errors, to make $\widehat{u}(z)$ and $u(z)$ as close to each other as possible. This is done by seeking to minimize the spread of $\widehat{\delta}$. Consequently, the functional $\mathcal{A}$ is chosen to be some positive measure of the spread of the averaging kernel. By minimizing the spread of $\widehat{\delta}$ the BG method thus seeks a high resolution solution that is most sensitive to abrupt discontinuities in $u(z)$.

Now for the details. A set of *inverse response kernels* $q_i(z)$ is sought so that the estimated profile is related to the measurements by

$$\widehat{u}(z) = \sum_i q_i(z) c_i = \mathbf{q}^T(z)\mathbf{c}. \tag{35}$$

where $\mathbf{q}(z)$ and $\mathbf{c}$ are vectors both of length N, the number of photopeak measurements. Toward this end, define the integral of the response kernel for each photon energy as

$$r_i \equiv \int_0^\infty dz \, \mathcal{R}_i(z). \tag{36}$$

Then, substitution of Eq. (35) into Eq. (5) and comparison of the result with Eq. (34) shows that the averaging kernel is given by

$$\widehat{\delta}(z', z) = \sum_{i=1}^N q_i(z) \mathcal{R}_i(z'). \tag{37}$$

To conserve contaminant atoms in Eq. (34), this kernel should be normalized to unity for every z, so that

$$1 = \int_0^\infty dz' \, \widehat{\delta}(z', z) = \sum_{i=1}^N q_i(z) \int_0^\infty dz' \, \mathcal{R}_i(z') = \sum_{i=1}^N q_i(z) r_i = \mathbf{q}^T(z)\mathbf{r}. \tag{38}$$

Many choices can be used to define a measure of the spread of $\widehat{\delta}(z', z)$ at each z. The Backus-Gilbert method uses the second moment of the square of the averaging kernel, i.e., the functional $\mathcal{A}$ is selected as

$$\mathcal{A} = \int_0^\infty dz' \, (z' - z)^2 [\widehat{\delta}(z', z)]^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} q_i(z) W_{ij}(z) q_j(z) = \mathbf{q}^T(z) \mathbf{W}(z) \mathbf{q}(z), \tag{39}$$

where $\mathbf{W}(z)$ is the $N \times N$ *spread matrix* whose elements are defined as

$$W_{ij} \equiv \int_0^\infty dz' \, (z' - z)^2 \mathcal{R}_i(z') \mathcal{R}_j(z'). \tag{40}$$

The functional $\mathcal{B}$ can also be expressed in terms of $\mathbf{q}(z)$. Through Eq. (35), standard propagation of errors gives

$$\mathcal{B} \equiv \text{var}\,[\widehat{u}(z)] = \sum_{i=1}^{N} \sum_{j=1}^{N} q_i(z) S_{ij} q_j(z) = \mathbf{q}^T(z) \mathbf{S} \mathbf{q}(z), \tag{41}$$

where $\mathbf{S}$ is the covariance matrix. If the counts c_i are assumed to be independent (as is usually assumed), the off-diagonal covariances terms in $\mathbf{S}$ can be neglected. Thus the elements of this matrix are simply $S_{ij} = \sigma_i^2 \delta_{ij}$ where δ_{ij} is the Kronecker delta function.

The column vector function $\mathbf{q}(z)$ is now selected as the function that minimizes the functional $\mathcal{A} + \lambda \mathcal{B}$ at every depth z, namely,

$$\text{minimize:} \quad \mathcal{A} + \lambda \mathcal{B} = \mathbf{q}^T(z) \, [\mathbf{W}(z) + \lambda \mathbf{S}] \, \mathbf{q}(z), \tag{42}$$

subject to the constraint of Eq. (38) that requires $\mathbf{q}^T(z)\mathbf{r}$ to equal unity. The minimizing solution is [Press et al. 1992; Shultis et al. 1996]

$$\mathbf{q}(z) = \frac{[\mathbf{W}(z) + \lambda \mathbf{S}]^{-1} \, \mathbf{r}}{\mathbf{r}^T \, [\mathbf{W}(z) + \lambda \mathbf{S}]^{-1} \, \mathbf{r}}. \tag{43}$$

For any particular set of data $\mathbf{c}$, whose elements are the measurements c_i, the solution $\widehat{u}(z)$ is thus formally given by

$$\widehat{u}(z) = \mathbf{q}^T(z)\mathbf{c} = \mathbf{c}^T \mathbf{q}(z) = \frac{\mathbf{c}^T \, [\mathbf{W}(z) + \lambda \mathbf{S}]^{-1} \, \mathbf{r}}{\mathbf{r}^T \, [\mathbf{W}(z) + \lambda \mathbf{S}]^{-1} \, \mathbf{r}}. \tag{44}$$

In the BG method, the choice of λ determines the tradeoff between resolution and numerical stability. If a more stable solution is needed, then a higher value of λ should be chosen, and, if a more resolved solution is desired, a lower value of λ should be chosen.

Unlike the LR solution, the BG solution for the soil contamination problem is usually non-negative. The matrix $[\mathbf{W}(z) + \lambda \mathbf{S}]$ is symmetric and positive definite with positive real

elements. Necesarily $[\mathbf{W}(z)+\lambda\mathbf{S}]^{-1}$ must have some negative elements, so that Eq. (43) may not yield a $\mathbf{q}(z)$ with non-negative elements. Indeed, for c_i distorted by sufficient noise, the estimated profile $\widehat{u}(z)$ can become negative for some depths. Nevertheless, for the noiseless data considered in this study, Eq. (44) was found always to produce non-negative profiles $\widehat{u}(z)$.

In contrast to the other methods, the closed form solution of Eq. (44) involves no discretization of the spatial variable. At each depth z at which $u(z)$ is to be evaluated from Eq. (44), the $\mathbf{W}(x)$ must be evaluated from Eq. (40) using some accurate numerical integration scheme (here an adaptive Gaussian quadrature procedure). Then, a set of $N \times N$ linear equations must be solved to evaluate the vector $[\mathbf{W}(z) + \lambda\mathbf{S}]^{-1}\mathbf{r}$. As a result the BG method involves considerably more computational effort than the LR method, although far less than the CLR method. One cautionary note: It was found that the determinant of $[\mathbf{W}(z) + \lambda\mathbf{S}]$ usually vanishes at a few values of z, and for z near such critical depths the linear equations become ill-conditioned, leading sometimes to inaccurate values of $\widehat{u}(z)$ being calculated near these critical depths. Such numerical inaccuracies produce small spurious bumps in the profile and can be seen in many of the BG examples presented in Section 4.

3.4 The Maximum Entropy (ME) Method

Maximum entropy methods provide yet another approach for regularizing the inversion of an underdetermined system by replacing the inversion with the minimization of a functional of the form $\mathcal{A}[\widehat{\mathbf{u}}] + \lambda\mathcal{B}[\widehat{\mathbf{u}}]$. In this approach, Bayesian techniques are used to combine both measured c_i data and prior expectations of the contaminant profile. Rigorous development is provided by Jaynes [1985], Gull and Skilling [1985, 1989], and Shultis et al. [1996], and a very readable discussion of the technique is given by Press et al. [1992]. Here, the important results are presented.

In the ME method, the functional $\mathcal{A}[\widehat{\mathbf{u}}]$ defining goodness of fit to the data is taken as the scalar $-\frac{1}{2}\chi^2(\widehat{\mathbf{u}}) = -\frac{1}{2}|\mathbf{A}\widehat{\mathbf{u}}-\mathbf{b}|^2$, whose minimization yields an estimated profile $\widehat{\mathbf{u}}$ that generates simulated data consistent with the observed c_i. The coefficient $-\frac{1}{2}$, which comes from the exponent of an assumed multivariate normal likelihood function, has been absorbed in the definition of the Lagrange multiplier λ for the previously described methods, but it is often shown explicitly in the ME development.

From arguments based on logical consistency, information theory, and statistical me-

chanics, the ME regularizing functional is based on the generalized Shannon-Jaynes entropy,

$$\mathcal{S}[\widehat{u}, m] = \int dz \left\{ \widehat{u}(z) - m(z) - \widehat{u}(z) \log\left[\frac{\widehat{u}(z)}{m(z)}\right] \right\} \leq 0, \tag{45}$$

or in discretized form

$$\mathcal{S}[\widehat{\mathbf{u}}, \mathbf{m}] = \kappa \sum_{j=1}^{M} \left[\widehat{u}_j - m_j - \widehat{u}_j \log\left(\frac{\widehat{u}_j}{m_j}\right) \right] \tag{46}$$

where κ is a positive constant. The maximum ($\mathcal{S} = 0$) of this scalar functional occurs when $\widehat{u}(z)$ equals $m(z)$, and thus, in the absence of any conclusive evidence to support a new profile $\widehat{u}(z)$, $m(z)$ becomes the default profile that represents the analyst's best prior estimate, or belief, about what the unknown concentration profile should be. Equation (45) has been used to define an entropic prior probability distribution of potential profiles that maximizes possible variations with respect to the assumed default model [Gull and Skilling 1985, 1989]. In this regard, entropic regularization of a degenerate inversion problem gives the least possible weight to prior information that is introduced via $m(z)$ to stabilize the solution and to reduce the chance that inaccurate conclusions will be drawn from spurious data.

At first glance, the explicit introduction of prior information may seem an unfair advantage for the ME algorithm. (It may even appear dangerously close to forcing a desired conclusion from limited data!) However, an honest inspection of any inversion method will reveal prior information in the form of implicit assumptions like the nature of continuity between adjacent concentrations, the choice of a Lagrange multiplier that yields reconstructions resembling results that have been seen before, and the construction of a response matrix that captures what is known about the physics of the problem before the data are collected. The sole motivation of regularization is, after all, to limit the solution space to a set of physically realistic profiles. The Bayesian philosophy supporting the use of an entropic prior places the burden of proof on the data, forcing them to provide convincing evidence of deviations from the default, while at the same time maximizing random possibilities with respect to the default that are available as potential solutions.

For the special case that $m(z)$ is constant for all z and that $\int dz\, u(z)$ is a fixed constant for all possible trial profiles, $\widehat{u}(z)$, the entropy functional becomes [Press 1992; Shultis et al. 1996]

$$\mathcal{S}[\widehat{u}] = -\int dz\, \widehat{u}(z) \log[\widehat{u}(z)], \tag{47}$$

where constant terms have been omitted since they do not affect the optimization process. It is this form of $\mathcal{S}$ that is often used in the ME method. However, in the PGNAA soil contamination problem, there is no way to set the normalization of $u(z)$ *a priori* for an unknown profile, so the more general entropy functional of Eq. (45) is used.

The Bayesian posterior probability distribution is proportional to the likelihood probability of observing the c_i given a profile $u(z)$ multiplied by the prior probability of actually having a given profile $u(z)$. By maximizing this posterior distribution, the best estimator of the contaminant profile $\widehat{u}(z)$ (or in discretized form $\widehat{\mathbf{u}}$) is obtained. This is achieved by maximizing the functional $\mathcal{Y}[\widehat{\mathbf{u}}] = \alpha\mathcal{S}[\widehat{\mathbf{u}}] - \frac{1}{2}\chi^2[\widehat{\mathbf{u}}]$, where α is a Lagrange multiplier. Equivalently, the estimated contaminant profile is found from the following minimization problem (after factoring a constant of $-\frac{1}{2}$, letting $\lambda = 2\alpha > 0$ and defining $\mathcal{H} \equiv -\mathcal{S} \geq 0$, the negative entropy or *negentropy*):

$$\text{minimize:} \quad \chi^2(\widehat{\mathbf{u}}) + \lambda\mathcal{H}[\widehat{\mathbf{u}}, \mathbf{m}]. \tag{48}$$

From a comparison of this result to that of the general regularization method of Eq. (22), it is seen that the nonlinear negentropy is now the regularizing functional and that χ^2 measures the misfit between the data and any assumed concentration profile. In fact, the optimization of Eq. (48) can be interpreted as maximizing the entropy subject to a constraint on χ^2.

There are several important distinctions of the entropy functional compared to the other regularization methods [Press et al. 1992]. First, $\mathcal{S}[\widehat{\mathbf{u}}, \mathbf{m}]$ is defined using only *local* values of $\widehat{u}_i$, i.e., $\mathcal{S}[\widehat{\mathbf{u}}, \mathbf{m}]$ sums the function $\widehat{u} - m - \widehat{u}\log(\widehat{u}/m)$ over all depths, and a random rearrangement of the $\widehat{u}/m$ values does not change $\mathcal{S}$. By contrast the LR method uses difference of $\widehat{u}$ between neighboring depths and, unlike the ME method, abrupt changes in $\widehat{u}(z)$ will greatly affect the regularizing functional. Second, as any $\widehat{u}_i$ goes to zero the slope of $\mathcal{H}$ becomes infinite forcing the minimization problem to maintain a positive solution, i.e. the positivity of the estimated profile is assured without the need to explicitly impose a positivity constraint.

Because of the nonlinear negentropy term in Eq. (48), the process of finding the optimum profile cannot be performed analytically as in the BG method, nor can it be reduced to a set of linear equations as in the LR method. Rather, numerical non-linear optimization algorithms must be employed. Skilling and co-workers have developed an effective algorithm

[Skilling and Bryan 1984; Gull and Skilling 1985] that has been used in the present work. This algorithm is summarized in the Appendix.

4 Example Results

To determine the capabilities of the four regularization methods to predict soil contamination profiles, several idealized profiles were used. With these test profiles for $u(z)$, simulated values of c_i were evaluated from Eq. (5) using an adaptive Gaussian numerical integration algorithm which is capable of achieving high accuracy. For the examples presented here, the soil was assumed to be uniformly illuminated by normally incident 14-MeV neutrons. The resulting thermal fluence profiles are those of Fig. 1 and the appropriate emission kernel is that of Eqs. (9) and (11). Results very similar to those presented here are obtained for neutrons from a point source above the soil as well as for neutrons emitted by ^{241}Am/Be or ^{252}Cf sources.

For the examples presented here, chromium is the assumed contaminant and, mostly, only the five capture-gamma photons ($N = 5$) with yields greater than 10% are used in the simulation. An example set of simulated c_i data generated in this manner is given in Table 1. These values of c_i, which are used in most of the examples, are for an idealized detector that produces no statistical uncertainty or interference from scattered photons in the measured c_i. Thus, the relative error σ_i/c_i is assumed to be negligibly small. Except for the ME results which set σ_i/c_i to a very small value, the covariance matrix $\mathbf{S}$ was taken as the identity matrix. Finally, unless otherwise stated, problem parameters are as stated in Table 1, the number of discrete depths is $M = 51$, and the PGNAA model discretization is based on Eq. (19).

With such simulated data for c_i, test contaminant profiles were obtained by inversion of Eq. (5) and compared to the actual profiles shown by the heavy dashed lines in the figures. Additional examples are presented by Shultis et al. [1996a, 1996b].

4.1 Results for Linear Regularization

The estimated contaminant profile depends strongly on the parameter λ that balances the trade-off between smoothness and accuracy of the profile as shown in Fig. 2. The smaller λ the better is the fit to the data, i.e., the smaller $\chi^2 \equiv \mathcal{A}[\hat{\mathbf{u}}] \equiv |\mathbf{A}\hat{\mathbf{u}} - \mathbf{b}|^2$ becomes. By

contrast, as λ becomes large, the inverted profile becomes smoother and the fit becomes less accurate, i.e, $\hat{\mathbf{u}}^T\mathbf{H}\hat{\mathbf{u}}$ becomes smaller while χ^2 becomes larger. This behavior is illustrated in Fig. 3 which shows how χ^2 and $\hat{\mathbf{u}}^T\mathbf{H}\hat{\mathbf{u}}$ vary with λ. In these examples, for $\lambda \geq 10$ the inverted profile assumes a strict linear shape as prescribed by the linear smoothing constraint used. At the other extreme when $\lambda = 10^{-13}$ the inverted profiles agrees very closely with the actual profile. However, if λ is made even smaller, the effect of the smoothing regularization term is lost because of the finite precision of the computer and numerical instabilities begin to develop as shown in the profile in Fig. 2 obtained with $\lambda = 3 \times 10^{-15}$.

Since the simulated c_i values used in these examples are exact, i.e., have no associated statistical uncertainties, the optimal (most accurate) computed profile is obtained with the smallest value of λ that can be used before numerical instabilities appear. These best-fit profiles, i.e., those with the smallest χ^2, represent the maximum capability of the inversion method since any real data for the c_i will have uncertainties and thus contain less information for the inversion process.

In Figs. 4 and 5 optimally resolved inverted profiles for step and Gaussian profiles are shown for the three smoothing regularization methods considered in the study. As can be seen, the estimated profiles are quite reasonable, especially for the Gaussian profile of Fig. 5. The three regularization smoothing methods give nearly the same results, differing primarily at large soil depths where spurious negative and positive variations are characteristically produced.

Fig. 5 also demonstrates the effect of using different numbers of capture-gamma photons. Generally, as more c_i are used the better are the computed results. However the increased accuracy in going from 5 chromium photons (minimum capture photon yield 10%) to 10 photons (minimum yield 6%) is surprisingly slight. The three smoothing methods produce nearly identical results for the Gaussian peak. Only for the spurious oscillations at deep depths, where $u(z)$ is negligibly small, do differences become apparent for the three smoothing schemes and for the number of photons used.

As seen in these examples, the linear regularization method is capable of extracting quite reasonable profiles given only a small number of c_i values (here 5). However, the inverted profiles for all three regularization methods typically exhibit large negative and even positive concentrations at large (optical) depths in the soil where $u(z)$ is usually extremely small. Also, when $u(z)$ is small near the surface, the computed profiles often yield negative

concentration estimates near the surface. These spurious values are characteristic artifacts of the inversion process since linear regularization imposes no positivity constraint on the $\widehat{u}(z)$.

One simplistic approach is simply to ignore such negative concentration estimates and to set the offending value to zero. Similarly, large positive values at great depth may be false and should likewise be ignored, particularly if preceded by a region of negative estimates. However, such an ad hoc procedure is not very satisfying.

4.2 Results for Constrained Linear Regularization

In Fig. 6 the converged CLR iterative solution based on Eq. (33) for a chromium contaminant bilinear profile is shown along with the LR solution. From this figure, it is seen that the CLR method produces a much better estimate of the contaminant profile than does the unconstrained LR method. The great disadvantage of the CLR method is the large number of iterations typically required to achieved a converged solution. This need for a large number of iterations is illustrated in Fig. 7 for a step profile. Millions of iterations may be required for convergence; however, quite usable profiles are realized before convergence. Although the converged profile is independent of the initial profile used to start the iterations, experience has shown that it is usually best to start with the null solution, $\widehat{u}^{(0)}(z) = 0$, since spurious positive concentrations at great depth (see Fig. 4) require an extreme number of iterations to suppress.

Constrained iterative results for a Gaussian test profile are shown in Fig. 8. It is seen again that the CLR method produces computed profiles that agree much more closely with the actual profiles than do results of the LR method. One important feature of the CLR solution is that the inverted profile is less sensitive to the assumed value of λ than is the LR solution. This effect is seen in Fig. 8 where the iterative solutions for $\lambda = 10^{-2}$ and $\lambda = 10^{-6}$ are seen to be very similar to the LR solutions. Since λ, which specifies the balance between smoothness and accuracy is difficult to determine *a priori*, the CLR method is very attractive, not only because it eliminates unrealistic negative concentrations, but also because it is relatively insensitive to λ. The number of iterations required for convergence is found to increases as λ decreases, i.e., the stronger the smoothing component the faster is the convergence. Moreover, as λ becomes very small the computed profile typically displays a saw-tooth pattern after only a few iterations (see Fig. 9). Fortunately, these saw-tooth

patterns smooth out as the number of iterations increases, as is seen in Fig. 9. However, this effect becomes more severe as λ decreases, and eventually numerical instabilities prevail and no convergence is achieved. Fortunately, the converged profile is rather insensitive to the value of λ used, and since convergence is faster and less biased toward the surface for larger values of λ, the CLR method should generally be employed with a value of λ that is several orders of magnitude greater than that used in the linear regularization method.

By far the most severe test of inversion techniques is provided by discontinuities in the contaminant profile, particularly if they occur at deep depths. In Figs. 10 and 11 the ability of the CLR and the LR methods, respectively, is demonstrated for resolving a uniformly contaminated stratum at different depths. From these figures, it is clear that the CLR method is far superior at resolving contaminant layers, although at the expense of many millions of times the computing effort.

4.3 Backus-Gilbert Results

The usual positivity of the inverted profile for the soil contamination problem with minimal noise is an obvious advantage of the BG method over the LR method. Moreover, the BG attempt to maximize the resolution of sudden changes in the concentration profile is very appealing for this particular application.

In Fig. 12 BG results for a bilinear test profile are shown for an inversion based on $N = 5$ photons. Similar results are shown in Fig. 13 for 10 photons. As can be seen the inverted profiles depend on λ. As expected, with large λ, e.g., $\lambda = 1$, a very smooth stable profile is obtained, which resolves the profile peak poorly. As λ decreases more importance is placed on resolving profile changes and better agreement with the exact profile (dashed lines) is obtained. From these figures it is seen also that as more photon energies are used, the BG method is able to resolve the profile to deeper depths.

Similar BG results are shown for step profiles in Figs. 14 and 15. As can be seen, the BG method can resolve contaminated layers if they are sufficiently close to the surface; however, it resolves deeply buried layers very poorly. Of all the methods considered in this study, the BG method was the poorest at finding deeply buried structure in the contaminant profiles.

Offsetting this deficiency of the BG method is its ability to yield, for the present soil contamination problem with noiseless data, non-negative estimates of the contaminant profile. More important, of all the inversion methods considered in this study, the BG profiles

were found to be least sensitive to the value of λ. As long as a value of λ is chosen between extremely large values (when the regularization dominates the inversion) and extremely small values (when the effect of regularization is weak), the inverted profile is almost independent of λ. It was found that even if exceptionally small values of λ are used so that the effect of regularization is lost, the BG method still usually yields a reasonable result even though severe numerical instabilities are produced in all the other methods. Because the *a priori* selection of an appropriate value for λ for any of the inversion methods is generally not obvious, the insensitivity of the BG technique to the λ value is a very useful feature of this method.

4.4 Maximum Entropy Results

In this study, the generalized Shannon-Jaynes entropy function of Eq. (45) was used in the maximum entropy implementation. This choice allows a default contaminant distribution $m(z)$ to be specified to which the inversion process reverts when the PGNAA data c_i are not capable of supporting any significant difference from the default profile. The set of algorithms and subroutines distributed commercially as MEMSys3 [Gull and Skilling 1989] was used to perform the nonlinear optimization necessary to find the $\widehat{u}(z)$. Also, unlike the previous examples, which were based on quadratic discretization of Section 2.4.2, the ME examples presented here are based on the piece-wise linear approximation of Section 2.4.1.

Figure 16 shows the effect of assuming a constant default that is equal to the peak of the bilinear test profile. As observed for the LR method, the reconstruction is sensitive to the value chosen for the Lagrange multiplier. In this section, however, curves have been labeled with associated values of the fit parameter χ^2/N (where $N = 5$) rather than with the values of λ directly. It is interesting to note transitions as the solutions pass from strongly regularized, very "smooth" fits ($\chi^2/N = 10^6$) to very "tight" fits ($\chi^2/N = 10^{-9}$) along the ME trajectory. Highly regularized solutions attribute the data to a "bathtub" concentration profile with a thin layer at the surface and a thick layer of minimal contribution beyond 20 cm. As tighter agreement with the data is enforced, the shallow peak "migrates" to the depth of the actual maximum while the concentration in the adjacent tails is suppressed. Very tight fits to the data introduce a second "mode" near the surface rather than broadening the existing, well located peak. This behavior suggests a sensitivity to, or an exaggeration of, the exponential integral response kernel, Eq. (11), which declines

rapidly with depth. Note that all reconstructions return to the default of $m(z) = 1$ when deviations are no longer supported by the forward model within the current value of χ^2.

Reconstructions for the bilinear test profile are illustrated in Fig. 17 for an assumed default of $m(z) = 0.1$. Here, the reconstructed peaks are much broader and demonstrate some degree of insensitivity for values of $\chi^2/N \lesssim 10^{-3}$. It was observed that any constant default profile within a factor of 10 higher or lower than the actual peak provided a reasonable set of reconstructions over the range of χ^2. Furthermore, subtle clues like profile suppression and growth in the tails, attempted bi-modal fits, and insensitivity to the Lagrange multiplier are available to an experienced user for the purpose of manually iterating $m(z)$ to achieve even better agreement with an unknown concentration profile. For constant default models below $m(z) = 0.01$, reconstructions were produced with narrow peaks and exaggerated concentrations that could not be considered acceptable. However, consistent locations and behavior of the profile modes provide strong evidence for a more reasonable choice of $m(z)$.

ME results for a single broad step profile are not illustrated here, but they exhibit much the same behavior illustrated in Figs. 16 and 17. By visual examination, the "best" ME results are comparable to those for CLR presented in Fig. 7, but show more sensitivity to the Lagrange multiplier and do a better job of identifying the location of the leading step concentration change. Very tight fits to the data induce a bi-modal reconstruction that equally overestimates the concentration near the corners of the underlying profile and underestimates the concentration near the center. This strongly resembles a Fourier reconstruction of a square pulse, and immediately suggests averaging the peaks to introduce a better default model. Reasonable profiles were obtained with constant default models as low as $m(z) = 10^{-6}$ without serious side-effects, and in many respects, ME performed more predictably for the step profile than for the bilinear example.

The apparently innocuous assumption of a constant default value actually introduces a great deal of information about the smoothness of acceptable profiles and it suggests a reasonable normalization or total mass of contaminant that can be present in the profile. Consider, for comparison, the truly non-informative prior that would be represented by an $m(z)$ composed of randomly chosen, uncorrelated concentrations at each discrete depth. In this respect, the default model supplies to ME what smoothing constraints supply to the LR method.

All the ME results presented here share the common features of "top-to-bottom" convergence with tighter data agreement, preferred sensitivity to potential concentrations at shallower depths, return to the assumed default at great depths, and positivity. In addition, the ME algorithm was found to be very fast, converging in a few seconds time with usually far fewer that 1000 iterations. More iterations were required to reach the tightest fits where lack of precision between the simulated data c_i and the matrix reconstruction $\mathbf{R}\hat{\mathbf{u}}$ on the same order as the target parameter χ^2/N seemed to introduce slowly damped oscillations in the ME trajectory. Fits tighter than $\chi^2/N = 10^{-9}$ could not be achieved before the algorithm "stalled" with no clear direction discernible for the next incremental vector step. This precision limit might be improved with more careful attention to the way surrogate data are generated, but it would offer no practical advantage since real data will have much larger uncertainties. In comparison with the LR method, it appears that adherence to the ME trajectory, beginning with a rational choice for the default $m(z)$ effectively prevents the algorithm from "wandering off" towards numerically acceptable but physically absurd solutions.

Figure 18 illustrates the capability of the ME method to resolve a uniformly contaminated stratum located at various depths. All of the results are presented on one figure for convenience, but the six 10-cm source zones were defined for separate calculations. From cursory parametric studies of χ^2/N were chosen fits with the most visual appeal to represent the best possible reconstructions; corresponding values of χ^2/N tended to decrease with increasing depth of the contaminant zone, and a default model of $m(z) = 10^{-6}$ was used throughout. ME appears to perform very well in terms of locating the leading edge and the maximum width of the source zone, even for the deeper strata. Peak concentrations tend to be overestimated with increasing depth, but it is likely that ME can continue to identify deeper sources until the magnitude of the surrogate data approach that of the discrepancy between c_i and $\mathbf{R}\hat{\mathbf{u}}$.

As a severe test of ME's resolving power, a bi-modal example problem for two separated contaminated layers was considered (see Fig. 19). Only the ME method was able to resolve the second more deeply buried contaminated layer. Profiles of this type represent a significant challenge for inversion algorithms because contributions to the data from the second (deeper) stratum are dominated by those of the first (shallow) stratum. Again, with decreasing χ^2 the solution "grows" over the first layer before stretching out to locate the

second. At the tightest available criterion of 10^{-7}, however, ME follows its proclivity for surface layers and introduces a spurious third, very thin, very high concentration peak at the surface, while significantly narrowing the peak over the first source zone. Given only five data points and a simple attenuation physics model, it is amazing that the presence of the second peak is recognized at all. In fact, none of the other methods provide this level of discrimination, and it is doubtful that ME would identify the peak using real data. Since relative contributions of the second source are so small, one must look to the ratios between data values to find information regarding energy-dependent attenuation that directs the algorithm towards a second peak. This fact has important implications for the appropriate selection of available channels from a composite gamma spectrum.

5 Conclusions

From the results of this study, it has been shown that the four regularization methods presented here are capable of estimating concentration profiles from a surprisingly few measured intensities of capture-gamma photons. Although the linear regularization method is the most computationally efficient method, it suffers from the tendency to produce spurious negative or positive values at large soil depths. While the linear regularization method with iterative positivity constraints avoids this difficulty, it is computationally very expensive and useful only when detailed predictions are needed. The Backus-Gilbert method produces nonnegative profiles at modest computational effort and yields good resolution for abrupt changes in the contaminant concentrations if the changes are close to the surface. Finally, the ME method also efficiently produces non-negative profiles which were found in most cases to be superior to those of the other three methods. Moreover, the ability of the ME method to incorporate prior knowledge into the inversion process makes it attractive for performing practical analyses. These conclusions are based on analyses of simulated data without statistical measurement error. As such these results represent upper limits on the capabilities of the four methods for the soil contamination problem. Future work will assess the importance of input data errors on the ability of the different inversion methods to predict contaminant profiles.

Appendix: ME Non-Linear Optimization

The Skilling algorithm[1] is based on maximization of the scalar functional $\mathcal{Y} = \alpha\mathcal{S} - \frac{1}{2}\chi^2$, which is a function of M variables $\widehat{u}_j$. As for the LR and CLR methods, we seek a stationary value of $\mathcal{Y}$ such that the column vector $\partial\mathcal{Y}/\partial\widehat{\mathbf{u}} = \mathbf{0}$. For a stationary value to be a maximum it is necessary that $-\partial(\partial\mathcal{Y}/\partial\widehat{\mathbf{u}})/\partial\widehat{\mathbf{u}}^T$ be a positive-definite matrix for the corresponding value of $\widehat{\mathbf{u}}$. Thus, the desired solution must obey, $\alpha\,\partial\mathcal{S}/\partial\widehat{\mathbf{u}} = \frac{1}{2}\partial\chi^2/\partial\widehat{\mathbf{u}}$. From Eq. (46) it is seen that $\partial\mathcal{S}/\partial\widehat{\mathbf{u}}$ is a column vector with elements $\log m_j - \log\widehat{u}_j$. Equation (21) can be written in matrix form as $\chi^2 = (\mathbf{c} - \mathbf{R}\widehat{\mathbf{u}})^T\mathbf{S}^{-1}(\mathbf{c} - \mathbf{R}\widehat{\mathbf{u}})$, so that differentiation with respect to $\widehat{\mathbf{u}}$, using the substitution $\mathbf{z} = \mathbf{c} - \mathbf{R}\widehat{\mathbf{u}}$, the chain-rule $(\partial y/\partial\widehat{\mathbf{u}}^T)^T = (\partial\mathbf{z}/\partial\widehat{\mathbf{u}}^T)^T(\partial y/\partial\mathbf{z})$, and the relations $\partial(\mathbf{R}\widehat{\mathbf{u}})/\partial\widehat{\mathbf{u}}^T = \mathbf{R}$ and $\partial(\mathbf{z}^T\mathbf{S}^{-1}\mathbf{z})/\partial\mathbf{z} = 2\mathbf{S}^{-1}\mathbf{z}$, gives

$$\frac{\partial\chi^2}{\partial\widehat{\mathbf{u}}} = \left\{\frac{\partial}{\partial\widehat{\mathbf{u}}^T}(\mathbf{c} - \mathbf{R}\widehat{\mathbf{u}})\right\}^T \cdot \frac{\partial}{\partial(\mathbf{c} - \mathbf{R}\widehat{\mathbf{u}})}\left\{(\mathbf{c} - \mathbf{R}\widehat{\mathbf{u}})^T\mathbf{S}^{-1}(\mathbf{c} - \mathbf{R}\widehat{\mathbf{u}})\right\} \tag{49}$$

$$= -2\mathbf{R}^T\mathbf{S}^{-1}(\mathbf{c} - \mathbf{R}\widehat{\mathbf{u}}), \tag{50}$$

This is essentially the same procedure used to derive the normal equations for least squares optimization, which are stated in Eq. (26) in the context of linear regularization.

Now, the non-linear equation to be solved for ME inversion can be stated explicitly as,

$$\partial\mathcal{Y}/\partial\widehat{\mathbf{u}} \equiv \nabla\mathcal{Y}(\widehat{\mathbf{u}}) = \mathbf{0} = \alpha(\log\mathbf{m} - \log\widehat{\mathbf{u}}) + \mathbf{R}^T\mathbf{S}^{-1}(\mathbf{c} - \mathbf{R}\widehat{\mathbf{u}}), \tag{51}$$

and the symmetric, negative Hessian matrix as,

$$-\partial(\partial\mathcal{Y}/\partial\widehat{\mathbf{u}})/\partial\widehat{\mathbf{u}}^T \equiv -\mathbf{Q} = \partial\left\{\alpha\log\widehat{\mathbf{u}} + \mathbf{R}^T\mathbf{S}^{-1}\mathbf{R}\widehat{\mathbf{u}}\right\}/\partial\widehat{\mathbf{u}}^T \tag{52}$$

$$= \alpha\,[\widehat{\mathbf{u}}^{-1}] + \mathbf{R}^T\mathbf{S}^{-1}\mathbf{R}. \tag{53}$$

Note that when $\widehat{\mathbf{u}}$ approaches $\mathbf{0}$, $-\mathbf{Q}$ becomes strongly diagonally dominant. When all elements are non-negative and the matrix is non-singular, $-\mathbf{Q}$ is positive-definite as required for the solution of Eq. (51) to be a maximum.

A conceptual, iterative solution for Eq. (51) can be written as,

$$\log(\mathbf{m}/\widehat{\mathbf{u}}^{n+1}) = -\mathbf{R}^T\mathbf{S}^{-1}(\mathbf{c} - \mathbf{R}\widehat{\mathbf{u}}^n)/\alpha, \tag{54}$$

[1] It should be noted that Skilling's development of ME in the context of Bayesian methods is much more advanced than the theory presented here. The following approach represents the "historic" application of simple ME regularization that compares most readily with other methods presented in this paper.

or

$$\hat{\mathbf{u}}^{n+1} = \mathbf{m} * \exp\left\{\mathbf{R}^T\mathbf{S}^{-1}(\mathbf{c} - \mathbf{R}\hat{\mathbf{u}}^n)/\alpha\right\}, \tag{55}$$

where log, exp, / and $*$ operate element by element (/ and $*$ denoting simple division and multiplication, respectively), and the superscripts denote successive approximations to $\hat{\mathbf{u}}$. No matter what values the components of $\hat{\mathbf{u}}^n$ and the corresponding data estimates $\mathbf{R}\hat{\mathbf{u}}^n$ may take, all components of the next iteration, $\hat{\mathbf{u}}^{n+1}$, will be non-negative. This demonstrates how ME regularization can naturally enforce positivity without the external constraints needed by other methods. It is also apparent in this form that a converged solution for $\hat{\mathbf{u}}$ depends on the value chosen for the Lagrange multiplier α. As $\alpha \to 0$, converged profiles approach the least squares solution, with their inherent degeneracy. As $\alpha \to \infty$, complete emphasis is given to entropic regularization which is maximized at the default profile $\mathbf{m}$.

The Bayesian philosophy of "updating" a current state of knowledge by incorporating new data with what is already known, forces one to think carefully about the prior information that may be available to help constrain profile reconstructions. One obvious constraint is the condition of positivity; physical concentration profiles cannot be negative. Also, real profiles cannot be unbounded. There may be a physical saturation limit or some information about the total mass of contaminant released in a spill that can be used to set a maximum. Given this emphasis on a well-defined prior model, the iterative solution method of Skilling begins with the assumption that $\hat{\mathbf{u}} = \mathbf{m}$, where $\mathcal{S} = 0$ and $\alpha = \infty$. To reach a given target value of α, the algorithm takes incremental steps away from the default in vector directions supported by the data. Successive solutions, corresponding to small adjustments in α towards the target, define a ME trajectory in the solution space of all possible $\hat{\mathbf{u}}$. Bayesian arguments also exist to support an *a priori* selection of the Lagrange multiplier, which is most often measured in terms of expectation values for the fit parameter χ^2, such as $\chi^2 = N$, the number of independent data points. Iteration stops when the solution has reached the corresponding value of α along the ME trajectory or when the data can no longer discriminate a meaningful direction for the next step.

Iteration along the ME trajectory proceeds by an algorithm based on a modified Newton's method. The scalar $\mathcal{Y}[\mathbf{u}]$ is first expanded about the current solution $\hat{\mathbf{u}}$ by Taylor's

series, which to the quadratic approximation is,

$$\mathcal{Y}[\hat{\mathbf{u}} + \delta\hat{\mathbf{u}}] = \mathcal{Y}[\hat{\mathbf{u}}] + \delta\hat{\mathbf{u}}^T \nabla\mathcal{Y}[\hat{\mathbf{u}}] + \frac{1}{2}\delta\hat{\mathbf{u}}^T \mathbf{Q}\delta\hat{\mathbf{u}} \tag{56}$$

Differentiation of Eq. (56) with respect to vector directions $\delta\hat{\mathbf{u}}$ yields the gradient at the new solution "point" (a complete concentration profile in M-dimensional vector space), which is forced to be stationary by setting $\partial\mathcal{Y}[\hat{\mathbf{u}} + \delta\hat{\mathbf{u}}]/\partial(\delta\hat{\mathbf{u}}) = \mathbf{0}$. Use of this condition in Eq. (56) shows that the next incremental step $\delta\hat{\mathbf{u}}$ is given by the equation $\nabla\mathcal{Y}[\hat{\mathbf{u}}] + \mathbf{Q}\delta\hat{\mathbf{u}} = \mathbf{0}$ whose solution is

$$\delta\hat{\mathbf{u}} = -\mathbf{Q}^{-1}\nabla\mathcal{Y}[\hat{\mathbf{u}}] = \left\{\alpha[\hat{\mathbf{u}}^{-1}] + \mathbf{R}^T\mathbf{S}^{-1}\mathbf{R}\right\}^{-1} \left\{\alpha(\log\mathbf{m}/\hat{\mathbf{u}}) + \mathbf{R}^T\mathbf{S}^{-1}(\mathbf{c} - \mathbf{R}\hat{\mathbf{u}})\right\}. \tag{57}$$

In order for the Newton direction $\delta\hat{\mathbf{u}}$ to move the solution towards the desired maximum, $\{\nabla\mathcal{Y}[\hat{\mathbf{u}}]\}^T\delta\hat{\mathbf{u}} < 0$, so by rearranging Eq. (57), $\delta\hat{\mathbf{u}}^T\mathbf{Q}\delta\hat{\mathbf{u}} > 0$, i.e., the Hessian matrix must be positive-definite. Press warns that far from the maximum there is no guarantee the $\mathbf{Q}$ will have this property, which may lead to steps away from the desired solution. Even when $\mathbf{Q}$ is positive-definite, large steps in the Newton direction can lead to points too far away for the quadratic approximation to be valid. So-called "quasi-Newton" methods treat these difficulties by building approximate Hessian matrices that are forced to be positive-definite, and by backtracking along the full Newton step to a point that lies with a trust region that supports the quadratic approximation. The Skilling algorithm defines the trust region with an effective radius of,

$$r_0^2 \geq (\delta r)^2 = \delta\hat{\mathbf{u}}^T[\hat{\mathbf{u}}^{-1}]\delta\hat{\mathbf{u}}. \tag{58}$$

The value of α appearing in $\mathbf{Q}$ is set just large enough that this constraint is obeyed for each iteration of α along the ME trajectory. Again, proximity to a well-defined default profile $m(z)$ encourages smooth convergence.

References

Backus G.E., and F. Gilbert, "Numerical Applications of a Formalism for Geophysical Inverse Problems," *Geophys. J. R. Atro. Soc.*, **13**, (1967).

Backus G.E., and F. Gilbert, *Geophys. J. R. Atro. Soc.*, **16**, 169-205 (1968).

Biemond, J., R.L. Lagendijk, and R. M. Mersereau, *Proceedings of the IEEE*, **78**, 856–883 (1990).

Craig, I.J., and J.C. Brown, *Inverse Problems in Astronomy*, Adam Hilger, Bristol, U.K. (1986).

Gull, S.F., and J. Skilling, "The Entropy of an Image," *Maximum-Entropy and Bayesian Methods in Inverse Problems*, University of Wyoming, pp. 287–301, 1985.

Gull, S.F., and J. Skilling, "Quantified Maximum Entropy: MEMSYS 3 Users' Manual," Maximum Entropy Data Consultants, Ltd., Royston, England, 1989.

Jaynes, E.T., "Where Do We Go from Here?" *Maximum-Entropy and Bayesian Methods in Inverse Problems*, C.R. Smith and W.T. Grandy Jr., eds., Reidel, Dordrecht, Netherlands, 1985.

Letellier, B., *Maximum Entropy a Posteriori Method for Estimating Soil Contamination Profiles*, Ph.D. Dissertation, Kansas State University, Manhattan ,KS; *in preparation.*

Loredo T.J., and R.I. Epstein, *Astrophysical Journal*, **336**. 896-919, (1989).

Miller, K., *SIAM Journal on Mathematical Analysis*, **1**, 52–74 (1970).

Parker, R.L., "Understanding Inverse Theory," *Annual Review of Earth and Planetary Science*, **5**, 35-64, (1977).

Phillips, D.L., *Journal of the Association for Computing Machinery*, **9**, 84–97 (1962).

Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes*, 2nd ed., Cambridge Univ. Press, Cambridge, 1992.

Shue, S.L., R.E. Faw and J.K. Shultis, "Fast Neutron Thermalization and Capture Gamma-Ray Generation in Soils," Proc. HSRC/WERC Joint Conference on the Environment, Albuquerque, NM, May 1996; Publ. by Great Plains-Rocky Mountain Hazardous Substance Research Center, Kansas State University, Manhattan, KS, 1996.

Shue S.L. and R.E. Faw, *Transport in Soil of Neutrons from 14-MeV, ^{252}Cf, and $^{241}Am/Be$ Sources*, HSRC-94-02-11, Department of Nuclear Engineering, Kansas State University, Manhattan, KS, 1996.

Shue S.L., R.E. Faw, and J.K. Shultis, "Thermal Neutron Intensities in Soils Irradiated by Fast Neutrons from Point Sources," submitted to *Isotope Geoscience*, 1997.

Shultis, J.K., R.E. Faw, and F. Khan, *Mathematical Models and Analysis for PGNAA of Soil Contamination*, HSRC-94-02-10, Department of Nuclear Engineering, Kansas State University, Manhattan, KS, 1996a.

Shultis,J.K., R.E. Faw and F. Khan, "Methods for Determining Soil Contaminant Profiles From Prompt-Neutron Gamma Activation Data," Proc. HSRC/WERC Joint Conference on the Environment, Albuquerque, NM, May 1996; Publ. by Great Plains-Rocky Mountain Hazardous Substance Research Center, Kansas State University, Manhattan, KS, 1996b.

Skilling, J., and Bryan, R.K., "Maximum Entropy Reconstruction: General Algorithm," *Monthly Notices of the Royal Astronomical Society*, **211**, 111-124, 1984.

Tikhonov, A. N., "The Approximate Solution of Fredholm Integral Equations of the First Kind," *Zh. Vych. Mat.*, **4**, 564–571 (1964).

Tikhonov, A. N., and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, Wiley, New York (1977).

Twomey, S., *Journal of the Association for Computing Machinery*, **10**, 97–101 (1963).

Twomey, S., *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements*, Elsevier, Amsterdam (1977).

Table 1. Calculated c_i values for chromium contaminated soil distributed with the bilinear profile shown in Fig. 2. Problem parameters: bilinear profile; dry dense soil of density 1.77 g/cm^3); $J_o = 1$; $\eta_i = 1$; $\theta_d = \pi/2$.

Energy (MeV)	Yield per capture	μ_i (cm^{-1})	c_i
0.7492	.1104	0.130030	0.19872
0.8351	.2686	0.123690	0.53652
7.9393	.1275	0.043362	1.46361
8.8841	.2697	0.042026	3.22336
9.7203	.1097	0.041009	1.35258

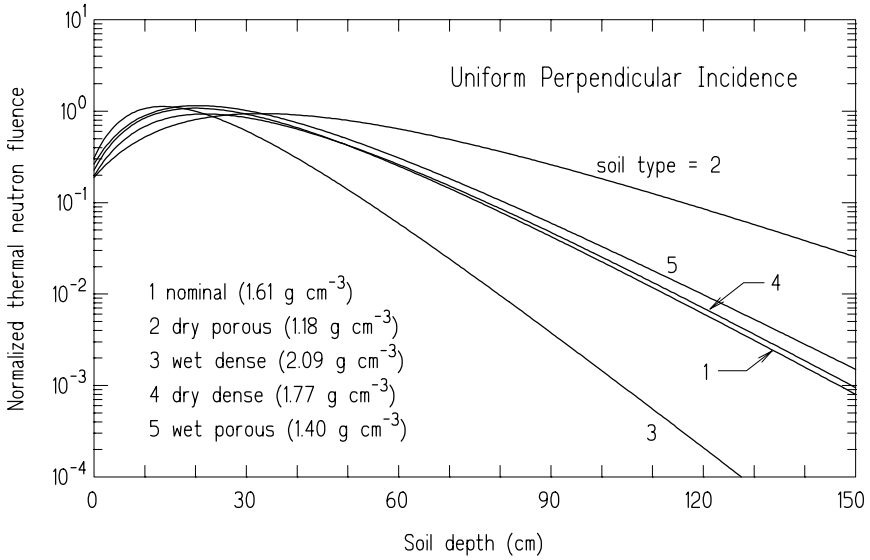

Figure 1. Thermal fluence profiles in the five representative soils used in this study. The soil surface is uniformly and perpendicularly irradiated by 14-MeV neutrons with a unit inward flow at the surface.

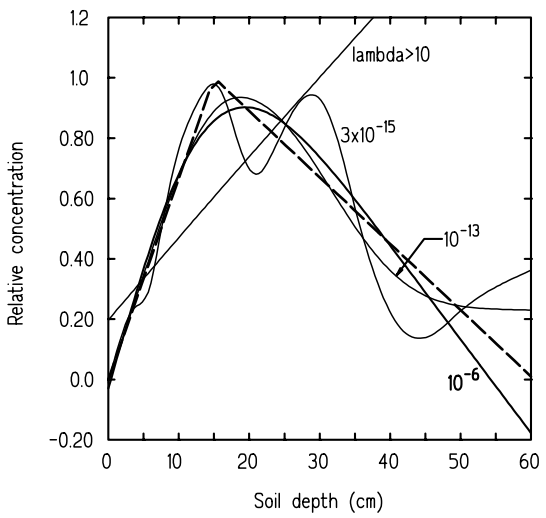

Figure 2. LR estimated profiles using linear smoothing for several values or λ. Exact bilinear profile is shown by the heavy dashed line.

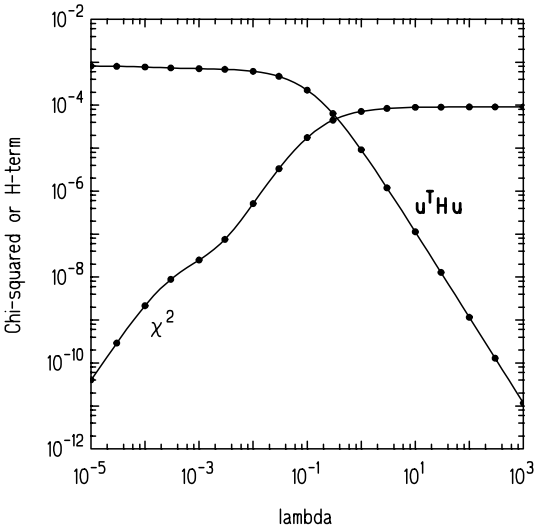

Figure 3. Variation of the accuracy (χ^2) and the smoothing regularization ($\widehat{\mathbf{u}}^T\mathbf{H}\widehat{\mathbf{u}}$) with λ for the results in Fig. 2.

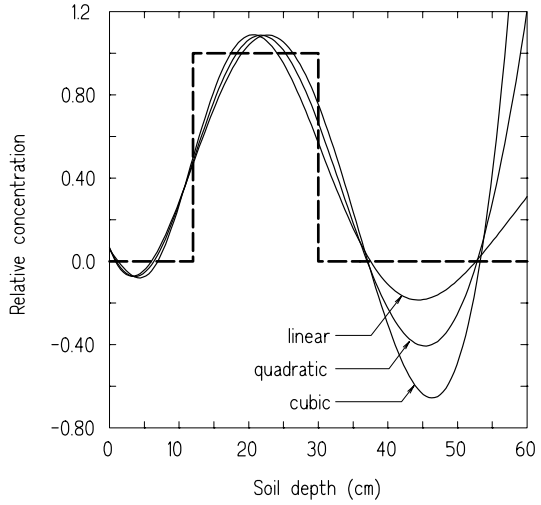

Figure 4. Optimal LR inversion ($\lambda \simeq 10^{-9}$) a step profile using three smoothing functionals.

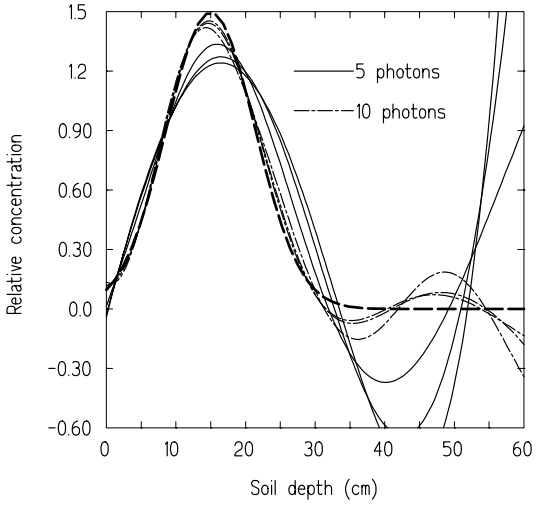

Figure 5. Optimal LR inversion ($\lambda \simeq 10^{-9}$) of a Gaussian profile using linear, quadratic and cubic smoothing functionals based on 5 and 10 photon energies.

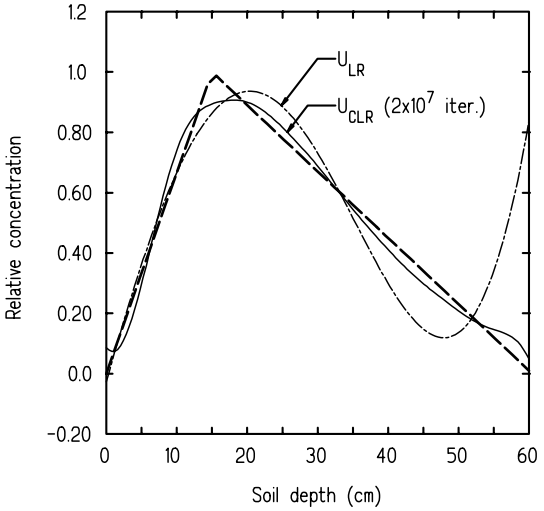

Figure 6. Optimal LR and CLR inversion for the bilinear profile with cubic smoothing. Exact profile is shown by the heavy dashed line.

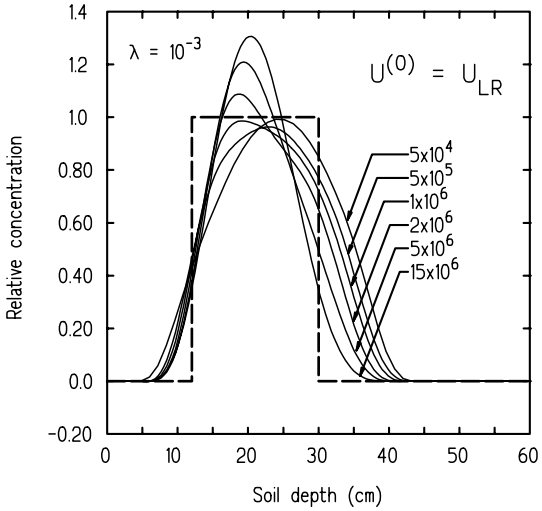

Figure 7. CLR inversion for the step profile after various number of iterations for $\lambda = 10^{-4}$. Starting profile was that obtained with LR inversion.

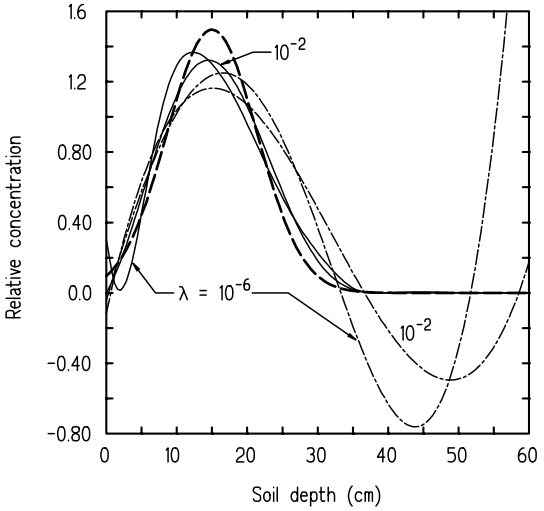

Figure 8. CLR profiles (solid lines) and LR profiles (broken lines) for a Gaussian profile for two different λ values. Cubic smoothing was used and CLR iterations continued until convergence.

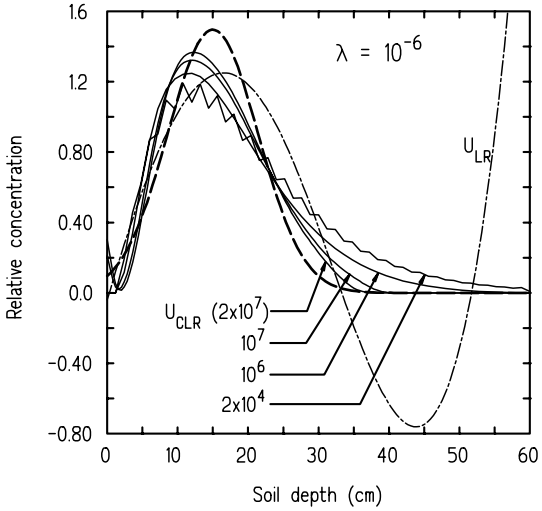

Figure 9. CLR inversion of a Gaussian profile after various number of iterations. Starting profile is $u_{LR}^{(0)}(z) = 0$. Also show by the broken line is the LR calculated profile.

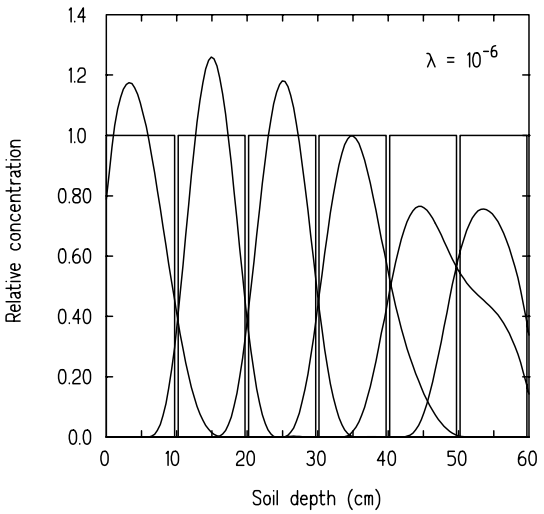

Figure 10. Converged CLR profiles (peaked curves) for a uniformly contaminated stratum at various depths. The rectangular profiles are the exact profiles (solid lines rather than dashed lines are used for clarity).

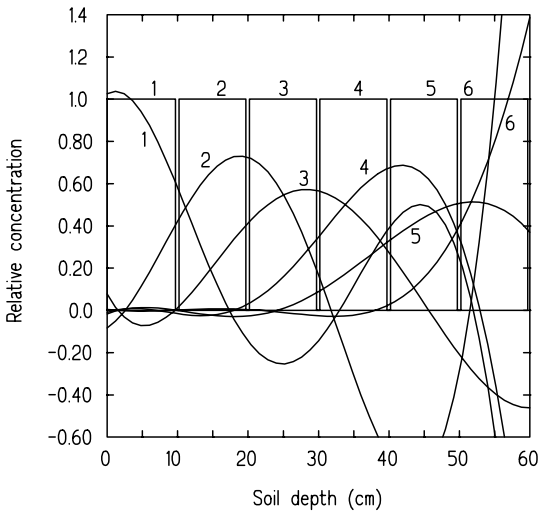

Figure 11. The LR profiles for the uniformly contaminated strata of Fig. 10. The number pairs indicate the corresponding inverted profile with the exact profile.

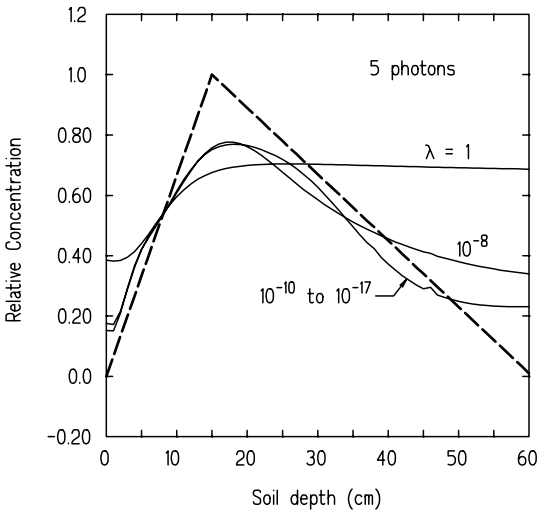

Figure 12. BG bilinear profile for several values of λ. Simulated data were based on the 5 most abundant Cr capture-gamma photons.

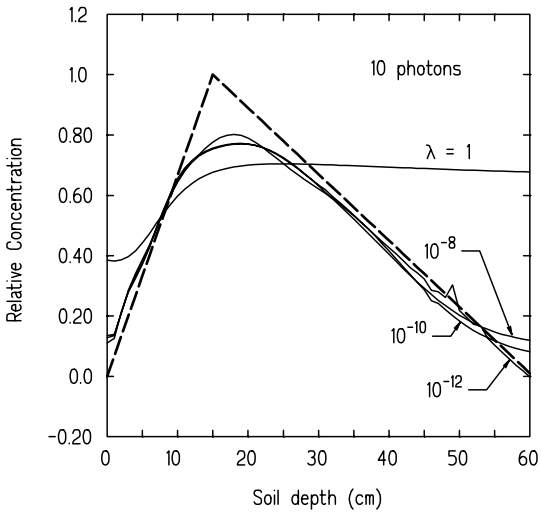

Figure 13. BG bilinear profile for several values of λ. Simulated data were based on the 10 most abundant Cr capture-gamma photons.

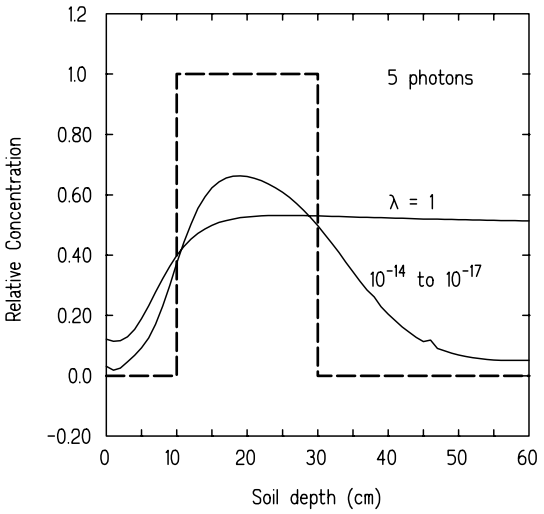

Figure 14. BG bilinear profile for several values of λ. Simulated data were based on the 5 most abundant Cr capture-gamma photons.

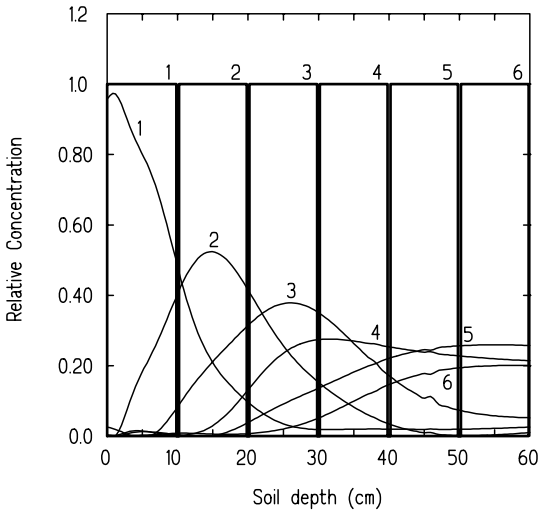

Figure 15. BG profiles (peaked curves) for a uniformly contaminated stratum at various depths. The rectangular profiles are the exact profiles (solid lines rather than dashed lines are used for clarity).

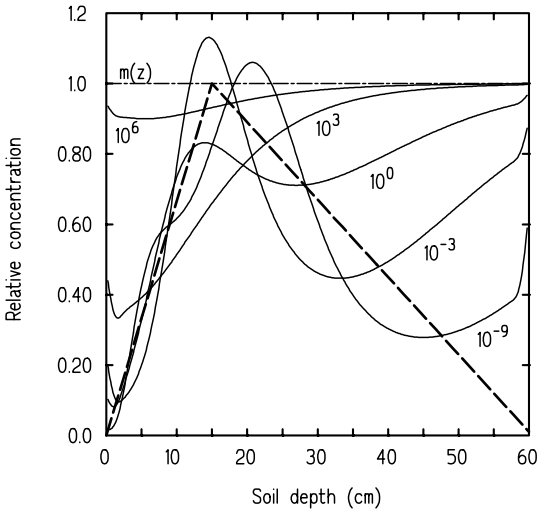

Figure 16. ME computed bilinear profiles for several values of the data fit metric χ^2/N. The assumed default profile is $m(z) = 1$.

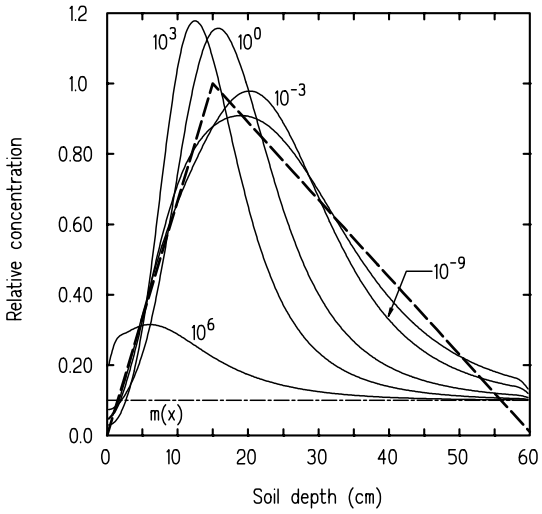

Figure 17. ME computed bilinear profiles for several values of the data fit metric χ^2/N. The assumed default profile is $m(z) = 0.1$.

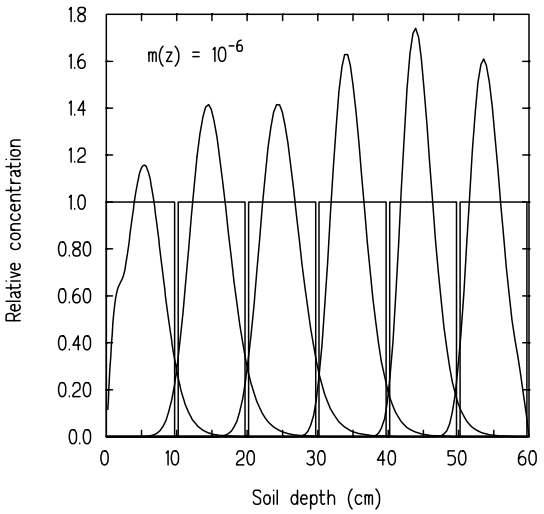

Figure 18. Optimal ME computed profiles for a uniformly contaminated layer (rectangles) at various depths.

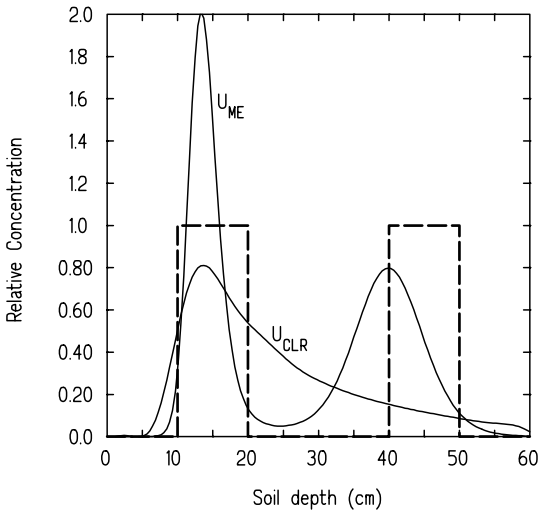

Figure 19. Optimal ME and CLR computed profiles for two separated uniformly contaminated layers (dashed line).