Sādhanā, Vol. 24, Parts 1 & 2, February & April 1999, pp. 73–96. © Printed in India.

Inverse modelling for parameter estimation and experiment design

V K GAUR

Indian Institute of Astrophysics, Sarjapur Road, Koramangala,
Bangalore 560 034 India
e-mail: gaur@cmmacs.ernet.in

Abstract. Inverse modelling forced itself on the attention of scientists in the 1960s with the advent of satellites and other revelatory technologies, despite their putative ill-posedness, when it became clear that estimation of parameters of a system not in themselves directly observed but extractable from their signatures in measured data, constituted an ineluctable problem of modern society. This paper begins by formulating the basic statement of inverse problems which have a generic form and leads through philosophical and analytical approaches to their possible solutions that are inherently non-unique. Finally, an example is provided for inverse modelling of the shear wave velocity structure of the crust beneath the ancient granites around Hyderabad from an analysis of reverberations caused by it that appear in the early part of broadband seismograms.

Keywords. Inverse modelling; parameter estimation; crustal structure; shear wave velocity.

1. Introduction

Inverse methods have been used for a long time to obtain valid answers to various problems of science and engineering. Geophysicists have used inverse methods to gain knowledge of the earth's internal structure by analysing anomalies in the natural and stimulated geophysical fields such as the travel times of seismic waves recorded at the surface, and engineers have studied transfer functions of a host of systems to characterise their inner structure. Gauss mentions using the method of least squares as early as 1795 to determine the orbital parameters of minor planets from observed data.

Inverse methods began receiving critical attention since the sixties when the launching of earth-orbiting satellites opened up the possibility of viewing the earth system as a whole from a vantage point in space and thereby determining the structure and working of its infinitely coupled solid and fluid spheres. A knowledge of some state parameters of the atmosphere and oceans, for example, is now routinely obtained by inverting radiance data measured from satellites. In order to examine how the unknown parameters or functions to

be inferred in various inverse problems are functionally related to measured data, we first examine the basic statement of two typical inverse problems, of determining the vertical temperature distribution in the atmosphere from satellite radiometric data and the internal structure of an object from the intensities of criss-crossing radiation traversing the body. Subsequently, we shall explore the source and nature of the limitations cast on the quality of inverse solutions as well as of their potential usefulness.

2. Determination of the vertical temperature distribution in the atmosphere from satellite radiometric data

The radiation $d(\lambda)$ at different wavelengths, measured by a sensor above the top of the atmosphere, can be expressed as an integrated effect of the thermal radiation emitted by various layers whose black body temperatures are $T(z)$. Here, we deduce the functional relationship between $d(\lambda)$ and $T(z)$.

A pencil of radiation of wavelength (λ) traversing a medium, will be attenuated by its interaction with matter i.e., absorption and scattering. Simultaneously, however, it will also intensify because of emission from the medium as well as multiple scattering that may eventually contribute to the incident ray. If scattering by air molecules can be neglected, which is usually justified in view of their small size, the change in the intensity $I(\lambda, \mu, \tau_z)$ of an incident beam in passing through an atmospheric layer of vertical thickness dz can be written using Kirchoff's law for absorption and Planck's law for emission in the layer.

$$dI(\lambda, \mu, \tau_z) = \mu^{-1}(-I + B)d\tau, \tag{1}$$

where $d\tau$ is the optical depth equal to kdz, k the absorption coefficient and $B(\lambda, T)$ the Planck radiance at temperature $T(z)$ prevailing at the altitude z; $\theta = \cos^{-1}\mu$ is the radiometer look angle.

Equation (1) is the fundamental equation of radiative transfer which we shall integrate to obtain the emergent radiation intensity at the top of the atmosphere. Rewriting (1) we get

$$\int_{\tau_0}^{\tau_t} d(Ie^{\tau/\mu}) = \int_{\tau_0}^{\tau_t} \mu^{-1} B\, e^{\tau/\mu} d\tau, \tag{2}$$

or,

$$I(\lambda, \mu, \tau_t) = I(\lambda, \mu, \tau_0)\exp[-(\tau_t - \tau_0)/\mu)] + \int_{\tau_0}^{\tau_t} \mu^{-1} B \exp[-(\tau_t - \tau_0)/\mu]d\tau.$$

The first term on the right can be neglected if τ_t is high enough, so that

$$I(\lambda, \mu, \tau_t) = \int_{\tau_0}^{\tau_t} \mu^{-1} B \exp[-(\tau_t - \tau_0)/\mu]d\tau = \int_{\tau_0}^{\tau_t} B(\lambda, \tau)dt(\tau), \tag{3}$$

where the new variable $t(\tau) = \exp[-(\tau_t - \tau_0)/\mu)]$ may be regarded as representing transmission. Further, transforming to a physically realizable variable such as pressure or altitude e.g., $t = \varphi(z)$ where $t = f(\tau)$, we obtain

$$I(\lambda, \mu, \tau_t) = \int_{z_1}^{z_2} B(\lambda, T(z))(\mathrm{d}t(z)/\mathrm{d}z)\mathrm{d}z, \tag{4}$$

or

$$I(\lambda) = \int_{z_1}^{z_2} G(\lambda)B(\lambda, T)\mathrm{d}z. \tag{5}$$

The kernel of this integral equation is the derivative of the transmission function. Besides, the equation is nonlinear as B also varies with wavelength. To linearize this equation, B is computed about some appropriate reference wavelength λ_0 using Taylor's expansions such as,

$$\begin{aligned} B(\lambda, T) &= B(\lambda_0, T) + \left.\frac{\partial B}{\partial \lambda}\right|_{\lambda_0} \Delta\lambda \\ &= \bar{B}(T) + \Delta B(\lambda), \end{aligned}$$

where $B(\lambda_0, T) = \bar{B}(T)$ is a function solely of $T(z)$, or

$$[I(\lambda) - \Delta B(\lambda)] = \mathrm{d}(\lambda) = \int_{z_1}^{z_2} G(\lambda, z)\ B_\lambda[T(z)]\,\mathrm{d}z. \tag{6}$$

The radiation $\mathrm{d}(\lambda)$ at different wavelengths, measured by a sensor above the top of the atmosphere, is thus expressed by (6) as an integrated effect of the thermal radiation emitted by various layers whose black body temperatures are $T(z)$.

Another example of inverse modelling is the problem of imaging the internal structure of objects from measurements made along criss-crossing rays traversing the object.

3. Computerized axial tomography (CAT)

This is an X-ray imaging device based on the mapping of X-ray opacity of body tissue, using measured attenuation of the beam intensity. The intensity of the beam diminishes with the distance travelled, at a rate proportional to the absorption coefficient of the medium traversed.

$$\frac{\mathrm{d}I}{\mathrm{d}l} = -m(x, y)I_0,$$

where m is the absorption coefficient of the medium at (x, y) and I_0, the X-ray source intensity. At the kth detector, the measured normalized intensity I_k/I_0 is given by

$$\frac{I_k}{I_0} = \exp\left\{-\int_{l_k} m(x, y)\mathrm{d}l\right\}. \tag{7}$$

Equation (7) is a nonlinear function of the unknown $m(x, y)$ which varies continuously along the beam. For a practical solution of the problem, however, it can be *linearized* by approximating the exponential with the first two terms of its Taylor's expansion. Thus, assuming the net absorption of X-rays to be small, $\exp(-x)$ can be replaced by $(1 - x)$,

$$\frac{I_k}{I_0} = 1 - \int_{l_k} m(x, y)\,\mathrm{d}l.$$

The measured quantity $(\Delta(I_k/I_0))$ can then be expressed by the following integral; or

$$\mathrm{d}_k = \frac{\Delta I_k}{I_0} = \frac{(I_0 - I_k)}{I_0} = \int_{l_k} m(x, y)\mathrm{d}l. \tag{8}$$

4. Seismic tomography of the earth

Similarly, the internal structure of the earth can be visualized from measurements of the travel times of seismic waves from known earthquake sources to a given array of seismographs. Since travel time between a source j and a receiver k can be determined for a given reference earth model, travel time residuals obtained by substracting the measured value from the theoretically computed value can be related to departures in the seismic velocity structure from that of the reference earth. Thus, for rays emanating from any given source,

$$\Delta t_k = (t_k)_{\text{Observed}} - (t_k)_{\text{Reference Earth}} = \int_{l_k} \left[\frac{1}{v_0 + dv_k} - \frac{1}{v_0}\right] \mathrm{d}l,$$

where $v_0(x, y, z)$ is the reference velocity.

$$\Delta t_k = \int_{l_k} \frac{1}{v_0} \left\{\left[1 + \frac{\mathrm{d}v_k}{v_0}\right]^{-1} - 1\right\} \mathrm{d}l$$

$$= \int_{l_k} \frac{1}{v_0} \left[-\frac{\mathrm{d}v_k}{v_0}\right] \mathrm{d}l, \tag{9}$$

or

$$\mathrm{d}_k = \Delta t_k = \int_{l_k} \frac{1}{v_0} \Delta m_k(x, y, z) \mathrm{d}l. \tag{10}$$

Equations (6) (8) and (10) have the form of a Fredholm equation of the first kind in which the unknown parameter or function $T(z)$ or $m(x)$, to be determined from observed data, multiplied by a kernel or Green's function appears within the integral. However, the quantity on the left representing measured data can only form a discrete set of values for given arguments. A straightforward solution of such equations to extract the unknown parameter or function indeed exists in the case of a few special kinds of Green's functions. Some of the well-known integral transform pairs given below enable one to express the unknown function in the Fredholm's equation as an integral of the measured data multiplied by an inverse Green's function. For example, the Laplace transform pair,

$$\mathrm{d}(y) = \int_0^\infty \exp[-yz] T(z) \mathrm{d}z; \quad T(z) = \frac{1}{2\pi} \int_{\alpha - j\infty}^{\alpha + j\infty} \exp[zy] \mathrm{d}(y) \mathrm{d}y, \tag{11}$$

and the Fourier transform pair,

$$\mathrm{d}(y) = \int_0^\infty \exp[-2\pi j yz] T(z) \mathrm{d}z; \quad T(z) = \frac{1}{2\pi} \int_0^\infty \exp[2\pi j zy] \mathrm{d}(y) \mathrm{d}y. \tag{12}$$

However, even these perfectly invertible integrals cannot be evaluated in practice, as measured data, being essentially discrete, cannot be expressed as a continuous function required for evaluation of the integral without making assumptions that may seriously vitiate the solution. In the case of integration over a complex plane, one would require data to be specified for complex arguments, which is manifestly impractical.

Equation (6) forms the basis of an inverse problem in which the parameters or functions to be determined are continuous while the data set is always discrete. As remarked earlier,

its solution can be directly obtained only for a few well-behaved Green's functions. In addition, the only way to handle discrete values of the integrand, even though they may make up a continuous function, would be to resort to numerical solutions. This necessarily calls for a discretization of the integral by approximating it as a sum, using one of the several procedures of numerical quadrature.

However, the discrete inverse formulation, which we shall return to shortly, retains all the potential of continuous inverse solutions as well as of their limitations, to which are of course added the peculiar consequences of discretization. An analysis of the former thus provides illuminating illustrations of the sources of uncertainties in the inverse solution, thereby opening up the possibility of handling them with insight.

5. Uniqueness and stability

Before we transit from the continuous to discrete inverse theory, we examine two analytic questions of great importance to solutions of inverse problems.

The problem of non-uniqueness is embedded in the very form of the basic equation (6) that is, in the possibility that there may exist non-trivial solutions $T^*(z)$ of $T(z)$ for which the integral vanishes (Green's functions being singular). Aspects of the solution represented by $T^*(z)$ would thus have no possibility of being determined from the data. What is more, its existence, even if real, may not be revealed in a numerical solution as even the true singular nature of the associated Green's function may be masked by discretization.

Numerical solutions of an inverse problem are thus invariably non-unique owing to this ban on retrieval of some parts of the solution.

The other serious concern in solving inverse problems is one of stability. This is best illustrated by examining how significantly the desired information about unknown parameters or functions may enter the measured data. Let us consider the integral equation (11) which has the negative real exponential function as the Green's function. This function decreases monotonically with a degree of smoothness determined by its derivative, i.e. the value of the exponent z. At large values of the argument y, this function is very small (unless z is small which would make the function even more smooth) and would have the effect of heavily reducing the contributions of the unknown function or parameters of the model to the measured data. A particular result of this is to make the data at widely varying values of the arguments insensitive to values of the function, thereby reducing the possibility of estimating them reliably. Conversely, infinitesimal variations in the data, which may as well be caused by the presence of errors, will result in wide fluctuations of the function or model being estimated. Thus, the inverse solution does not depend on the data continuously in the sense that small errors in it may lead to large variations in the solution.

The following example due to Twomey (1977) demonstrates this fact clearly. Let (13) be an error-free statement of two experimental results in which the quantities on the RHS denote the observed data and the coefficients of the two unknowns x and y are the numerical values of the two experimental conditions,

$$\begin{aligned} x + y &= 3, \\ 1.00001x + y &= 3.00001. \end{aligned} \tag{13}$$

The true values of x and y are obviously 1 and 2 respectively.

Let us now evaluate the values of the unknowns for 3 different erroneous situations:

(i) 0.001% error in measuring the coefficient of x in the second experiment.

$$x + y = 3$$
$$x + y = 3.00001$$

These equations are inconsistent and therefore have no solution.

(ii) 0.001% error in measuring the observed data in the second experiment.

$$x + y = 3$$
$$0.00001x + y = 3$$

Solution: $x = 0$; $y = 3$.

(iii) 0.001% errors in the measurement of both data and the coefficient in the second experiment.

$$x + y = 3$$
$$x + y = 3$$

Undetermined system, therefore many solutions ($x = 3 - y$).

6. Ill-posed problems

Endemic non-uniqueness and instability of inverse solutions violate two of the three basic conditions of well-posed problems enunciated by Hadamard in the 19th Century: Uniqueness, stability and existence of solutions. The third condition requires a proof of the existence of a solution. This is usually a difficult exercise and in its absence a solution is just assumed to exist. However, the problems of non-uniqueness and instability render most of the inverse problems ill-posed. For a long time, therefore, inverse problems were ignored as not being worthy of serious study, but the realization that most urgent problems of contemporary concern are essentially inverse, has spurred considerable interest in the search for regularization methods to obtain stable, albeit approximate, solutions which are sensible in some sense.

7. Numerical quadrature

In order to carry out numerical evaluation of a Fredholm integral such as (6), it is first approximated by a sum. This step, termed *numerical quadrature*, is accomplished by discretizing the interval (a, b) into shorter sub-intervals of desired fineness, marked by quadrature points $z_1, z_2, \ldots, z_m$, at which the respective values of the function $T(z)$ are $T_1, T_2, \ldots, T_m$. A number of quadrature formulae can be developed for this purpose depending on the choice of the interpolation functions. If this function is assumed to be linear, within each sub-interval, we may rewrite (6) as follows:

$$\int_a^b G(z, y)T(z)\mathrm{d}z = \sum_k G_{jk}\, T(z, k), \tag{14}$$

where j represents the argument of the observed data, and

$$G_{jk} = \left\{ \frac{-z_k - 1}{z_k - z_{k-1}} \int_{z_{k-1}}^{z_k} G(z, y)\mathrm{d}z + \frac{z_{k+1}}{z_{k+1} - z_k} \int_{z_k}^{z_{k+1}} G(z, y)\mathrm{d}z \right\}$$
$$+ \left\{ \frac{1}{z_k - z_{k-1}} \int_{z_{k-1}}^{z_k} zG(z, y)\mathrm{d}z - \frac{1}{z_{k+1} - z_k} \int_{z_k}^{z_{k+1}} zG(z, y)\mathrm{d}z \right\}.$$

The integrals $\int G(z, y)\mathrm{d}z$ and $\int zG(z, y)\mathrm{d}z$ in each sub-interval can, in turn, be evaluated numerically or analytically. With these qualifications, (6) can now be written as:

$$\mathbf{d} = \mathbf{GT}. \tag{15}$$

8. Nonlinearity

It must be noted here that often enough we find that the unknown function or parameter to be determined is nonlinearly related to data as (5). Considerable advances have now been made, following developments of global optimization methods, to address such nonlinear inverse problems directly but a more popular approach to their solution through quasi-linearization can be quite effective. For example, as in (5) where B happens to be a function of λ we quasi-linearized this equation locally about a reference wavelength λ_0.

In general, if $\mathbf{Gm} = \mathbf{d}$ is nonlinear, we expand the inner product $[\mathbf{G}, \mathbf{m}]$ about a reference (initial guess) model vector $\mathbf{m_0}$. Accordingly, $\mathbf{d} = [\mathbf{G}, \mathbf{m_0}] + [\partial\mathbf{G}/\partial\mathbf{m}|_{\mathbf{m_0}}, \Delta\mathbf{m}] + \mathbf{O}|\Delta\mathbf{m}^2|$. Neglecting the higher order terms in $\Delta\mathbf{m}$ and in view of the linearity of inner products,

$$\left[\left. \frac{\partial\mathbf{G}}{\partial\mathbf{m}} \right|_{\mathbf{m_0}}, \Delta\mathbf{m} \right] = [\mathbf{F}, \Delta\mathbf{m}] = [\mathbf{F}, \mathbf{m}] - [\mathbf{F}, \mathbf{m_0}],$$

or

$$\{\mathbf{d} - [\mathbf{G}, \mathbf{m_0}] + [\mathbf{F}, \mathbf{m_0}]\} = [\mathbf{F}, \mathbf{m}],$$

or

$$\bar{\mathbf{d}} = \bar{\mathbf{G}}\mathbf{m}, \tag{16}$$

where $\bar{\mathbf{d}}$ is the new reconstituted data vector and $\bar{\mathbf{G}}$ the new Green's function equal to the Frechet derivative of $\mathbf{G}$ at $\mathbf{m_0}$. A nonlinear problem can thus be reduced to a linear one.

9. Solution of discrete linear inverse problems

We have seen how the working equation of most inverse problems can, subject to approximations involved in discretization and linearization whenever called for, be reduced to the matrix equation

$$\mathbf{Gm} = \mathbf{d}, \tag{17}$$

where the N-dimensional vector $\mathbf{d}$ represents the data set, the $N \times M$ matrix $\mathbf{G}$ represents the Green's function or the physical theory relating the data to parameters of the system

(or model parameters m_k) and $\mathbf{m}$ represents the M-dimensional model vector consisting, for example, of k values of the atmospheric temperature $T(z)$ to be estimated at different attitudes z_k.

In very few cases of practical interest is the solution of the matrix equation $\mathbf{Gm} = \mathbf{d}$ unique. Uniqueness is guaranteed only in the special case when the matrix $\mathbf{G}$ is non-singular and there can be no non-trivial solution of the homogeneous equation $\mathbf{Gm}=\mathbf{0}$. In such cases, the solution can be unambiguously written as $\mathbf{m} = \mathbf{G}_u^{-1}\mathbf{d}$, where $\mathbf{G}_u^{-1}$ is the unique inverse of matrix $\mathbf{G}$ and $\mathbf{G}_u^{-1}\mathbf{G}_u = I_M = \mathbf{G}_u\mathbf{G}_u^{-1}$. However, if the data contain errors and the vector $\mathbf{d}$ is contaminated by a random error vector, the covariance of the solution can be estimated from the relation $\mathrm{Cov}(\mathbf{m}) = \mathbf{G}^{-1}\mathrm{Cov}(\mathbf{d})(\mathbf{G}^{-1})' = \sigma^2\mathbf{G}^{-1}(\mathbf{G}^{-1})'$ in the case of identically distributed random errors whose variance is σ^2, $\mathbf{G}^{-1}$ being the derived inverse of $\mathbf{G}$, and the prime indicating a transpose.

In general, depending on the scheme of measurement defining the data argument, and parameterization of the model being inverted for, the matrix will not be square or even of full-rank. In such cases the above equation will have no unique solution and the matrix $\mathbf{G}$ will have no unique inverse. But, if a solution must be found, howsoever qualified, we must explore the possibilities of obtaining possible approximate solutions, and the specificities that such solutions will possess.

To elucidate this problem, we consider the three possible cases when the rank of the $N \times M$ matrix $\rho(\mathbf{G})$ is

(i) $\rho(\mathbf{G}) = M < N$,

(ii) $\rho(\mathbf{G}) = N < M$,

(iii) $\rho(\mathbf{G}) <$ (minimum of M and N).

(i) *Over-determined system*

The first set involves a system of N linear equations in M unknown model parameters. These may form a consistent or an inconsistent system, depending on whether the data vector $\mathbf{d}$ belongs or not to the column space of $\mathbf{G}$. If it does, then the system essentially reduces to M independent equations, thereby yielding a unique solution as in the case of a non-singular matrix.

If, on the other hand, the above system is inconsistent, as would most often happen because of perturbations in measured data introduced by random errors, one must seek the most acceptable approximate solution. A standard approach is to design an inverse of the matrix $\mathbf{G}$ which when substituted in the original equation will result in predicted values of data as close to the respective measured values as possible. Normative measures of 'closeness' or 'distance' are defined variously, each having its particular implications and leading to different inverses or solutions. One is thus faced with the challenge and opportunity of designing the best inverse for a given situation.

One way to approach an appropriate option is to ask how closely the inverse matrix $\mathbf{G}^{-1}$ maps the measured data vector into the predicted one (Menke 1989). For example, if $\mathbf{G}_l^{-1}$ is the best approximate inverse, then the estimated value of the model parameter will be

$$\hat{\mathbf{m}} = \mathbf{G}_l^{-1}\mathbf{d}, \tag{18}$$

and the corresponding values of the predicted data vector $\hat{\mathbf{d}}$ will be

$$\hat{\mathbf{d}} = \mathbf{G}\hat{\mathbf{m}} = (\mathbf{G}\mathbf{G}_l^{-1})\mathbf{d} = \mathbf{N}\mathbf{d}, \tag{19}$$

where, $\mathbf{N} = (\mathbf{G}\mathbf{G}_l^{-1})$.

Thus, we see that the degree of closeness of $(\mathbf{G}\mathbf{G}_l^{-1} = \mathbf{N})$ to the identity matrix $\mathbf{I}$, in fact, determines the fidelity with which a predicted data vector may approximate the observed one, and a sensible decision would be to design $\mathbf{G}_l^{-1}$ in such a way as to minimize the norm of $(\mathbf{I} - \mathbf{G}\mathbf{G}_l^{-1})$.

Let the desired inverse solution be $\mathbf{G}_l^{-1}$, such that $\| \mathbf{I} - \mathbf{G}\mathbf{G}_l^{-1} \|$ is a minimum. We first choose to minimize the Euclidean norm of the kth row vector of $(\mathbf{I} - \mathbf{G}\mathbf{G}_l^{-1})$, or of $(\mathbf{I} - \mathbf{N})$, where $\mathbf{N} = \mathbf{G}\mathbf{G}_l^{-1}$. Let,

$$\| I_{kp} - N_{kp} \| = S_k, \tag{20}$$

$$= \sum_p N_{kp}^2 - 2\sum_p N_{kp} I_{kp} + \sum_p I_{kp}^2. \tag{21}$$

To seek the desired $\mathbf{G}_l^{-1}$, therefore, and recalling that $\mathbf{N} = \mathbf{G}\mathbf{G}_l^{-1}$, we set

$$\partial S_k / \partial G_{qr}^{-1} = 0, \tag{22}$$

where G_{qr}^{-1} is the (q, r) element of $\mathbf{G}_l^{-1}$.

It may be noted that each of the terms in (2) is positive and the differential of the last term which is not a function of $\mathbf{G}_l^{-1}$, must be zero, (24) then reduces to

$$\frac{\partial}{\partial G_{qr}^{-1}} \left\{ \sum_p \left(\sum_m G_{km}\, G_{mp}^{-1} \right) \left(\sum_n G_{kn}\, G_{np}^{-1} \right) - 2 \left(\sum_t I_{kp}\, G_{kt}\, G_{tp}^{-1} \right) \right\} = 0,$$

or

$$\sum_p \left\{ \sum_m \sum_n G_{km}\, G_{mp}^{-1}\, G_{kn}\, \delta_{qn}\, \delta_{rp} + G_{km}\, G_{kn}\, G_{np}^{-1}\, \delta_{qm}\, \delta_{rp}) \right.$$

$$\left. -2 \sum_t I_{kp}\, G_{kt}\, \delta_{qt}\, \delta_{rp} \right\} = 0,$$

or

$$\sum_m G'_{qk}(G_{km} G_{mr}^{-1}) = G_{kq} I_{kp} \delta_{rp} = G'_{qk} I_{kr},$$

or

$$\mathbf{G}'\mathbf{G}\mathbf{G}_l^{-1} = \mathbf{G}'$$

or

$$\mathbf{G}_l^{-1} = (\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'. \tag{23}$$

This inverse is the same as would be obtained by minimizing the sum of squares of the misfits between the actual data elements and the corresponding values predicted from the

estimated model vector. For this reason, $\mathbf{G}_l^{-1}$ is called the least squares inverse. It can be shown that $\mathbf{G}_l^{-1}$ always gives a unique solution. A graphically illustrative example for a two dimensional model vector is given below.

It would be instructive to show that the corresponding model resolution matrix $\mathbf{R} = \mathbf{G}_l^{-1}\mathbf{G}$ which measures the fidelity with which estimated model parameters relate to true ones, is in fact the identity matrix, and that the normalized covariance of the estimated model parameters is $(\mathbf{G}'\mathbf{G})^{-1}$.

Indeed, it may often be desirable to minimize both $\| (\mathbf{I} - \mathbf{G}\mathbf{G}^{-1}) \|$ and some fraction say ε^2 of the normalized covariance of $\mathbf{m}$. If $\mathbf{E}$ represents the normalized data covariance matrix, the resulting inverse can be shown to be

$$\mathbf{G}_d^{-1} = (\mathbf{G}'\mathbf{G} + \varepsilon^2\mathbf{E})^{-1}\mathbf{G}'. \tag{24}$$

Equation (24) is identical to the damped least squares inverse. It will also be instructive to show that whilst $\mathbf{G}_\mathbf{l}^{-1}$ is not unique, the inverse solution always is. A proof of this for $\rho(\mathbf{G}) = 2$, is given below.

Least squares solution for $\rho(\mathbf{G})=2$, $\mathbf{N}>2$: Assuming that some measurable property (d) of a material is linearly dependent on, say, temperature (T), a number of measurements are made at different temperatures to determine the parameters of the linear model: $d = m_1 + m_2 T$.

The N measured values of data so generated can be expressed as:

$$\begin{bmatrix} d_1 \\ d_2 \\ \cdot \\ \cdot \\ d_N \end{bmatrix} = \begin{bmatrix} 1 & T_1 \\ 1 & T_2 \\ \cdot & \\ \cdot & \\ 1 & T_N \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_N \end{bmatrix},$$

or

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \varepsilon$$

where ε_i is the error associated with the measurement d_i.

It is clear that if $N = 1$, both m_1 and m_2 cannot be estimated. If $N = 2$, a solution can be obtained for m_1 and m_2 by assuming that the data is free from errors (neglecting ε_i). When $N > 2$, and the equations are inconsistent owing to random errors, the vector $\mathbf{d}$ does not belong to the column space of $\mathbf{G}$. Therefore, there is no unique solution. However, it would be sensible to estimate m_1 and m_2 by imposing the condition that the sum of the squares of error terms be a minimum. Accordingly,

$$\sum_{i=1}^{N} \varepsilon_i^2 = \mathbb{R} = \sum_{i=1}^{N} [d_i - m_1 - m_2 T_i]^2,$$

or

$$\mathbb{R} = \Big[\sum d_i^2 - 2m_1 \sum d_i - 2m_2 \sum d_i T_i + 2m_1 m_2 \sum T_i + N m_1^2 + m_2^2 \sum T_i^2\Big].$$

In a 3-dimensional coordinate system, $\mathbb{R}$ represents an elliptical paraboloid (figure 1) with just one minimum, thereby proving that the least squares solution is unique.

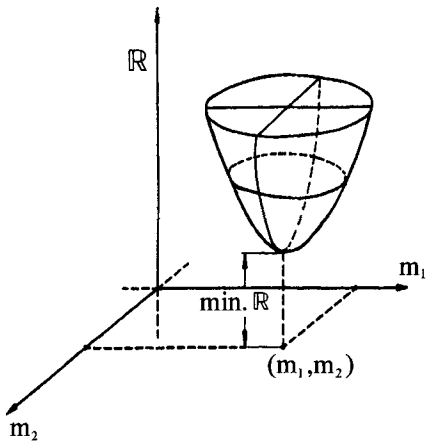

Figure 1. The mean square prediction error ℝ plotted as a function of model parameters of a 2-dimensional model space, describes an elliptical paraboloid which manifestly has a minimum for a unique pair of model parameters that constitute the least squares solution.

The estimates of the model parameters m_0, m_1 can be obtained by equating to zero, the differentials of ℝ w.r.t. m_1 and m_2. This would yield,

$$\hat{m}_1 = \frac{\sum d_i \sum T_i^2 - \sum T_i \sum d_i T_i}{N \sum T_i^2 - (\sum T_i)^2},$$

$$\hat{m}_2 = \frac{N \sum d_i T_i - \sum T_i \sum d_i}{N \sum T_i^2 - (\sum T_i)^2}.$$

(ii) *Under-determined system*

When the rank of the matrix is $N < M$, the data vector **d** will always belong to the column space of **G**. Therefore, **m** will always have a solution, albeit a non-unique one. For, **G** in this case is so conditioned that in operating on an M-dimensional model vector, it can only illuminate the resultant transformation in a lower N-dimensional data space, leaving the remaining $(M - N)$-dimensional subspace totally obscure. The operation **Gm** in this case annihilates all information belonging to this null subspace which is orthogonal to the N-dimensional illuminated subspace and is formed by $(M - N)$ linearly independent vectors, each comprising M elements.

The N-dimensional measured data thus reflect only a partial view of the model from a lower dimensional space and its inversion to estimate the model can only provide an incomplete image of the model. However, we have two options to make the best of this fuzzy situation. One can, in the first place, ask what definition of the model can best be extracted from the available data. Alternatively, if some additional guiding insights are available, one may attempt building up its missing dimension by designing an appropriate, even though wholly arbitrary, combination of the row vectors of the null space. In effect this would amount to adding $(M - N)$ independent but wholly arbitrary equations to raise the rank of the matrix to M, thereby contriving a unique inverse. The resulting solution would, however, be legitimate only to the extent that the added equations can be justified.

A search for the best possible solution, in turn, requires that the estimated model vector $\hat{\mathbf{m}}$ be as close to the true one as possible, since $\hat{\mathbf{m}} = \mathbf{G}_m^{-1}\mathbf{d} = \mathbf{G}_m^{-1}\mathbf{G}\mathbf{m} = \mathbf{R}\mathbf{m}$, **R** must

be as nearly equal to the identity matrix as possible. This can be ensured by minimizing the Euclidean norm of $\| \mathbf{I} - \mathbf{R} \|$. Let S_k represent the Dirichlet spread of the kth row of $\mathbf{R}$, then

$$S_k = \sum_p (I_{kp} - R_{kp})^2 = \sum_p [I_{kp}^2 + R_{kp}^2 - 2I_{kp}R_{kp}]. \tag{25}$$

Since each S_k is positive, we can minimize the total spread of $\mathbf{R} = \sum_k S_k$ by minimizing each S_k respectively with respect to each of the elements G_{qr}^{-1} of the desired inverse $\mathbf{G}_m^{-1}$. Accordingly,

$$\frac{\partial}{\partial G_{qr}^{-1}} \left[\sum_p \left\{ I_{kp}^2 + \sum_m \sum_n G_{km}^{-1} G_{mp} G_{kn}^{-1} G_{np} \right\} - 2I_{kp} \sum_t G_{kt}^{-1} G_{tp} \right] = 0, \tag{26}$$

or

$$\sum_x G_{rp} G_{xp} G_{qx}^{-1} = G_{rp} I_{qp}, \tag{27}$$

or

$$\sum_p \sum_x G_{qx}^{-1} (G_{xp} G'_{pr}) = G'_{qr}. \tag{28}$$

Thus elements G_{qr}^{-1} of the desired inverse $\mathbf{G}_m^{-1}$ must satisfy (28) or, $\mathbf{G}_m^{-1} = \mathbf{G}'(\mathbf{G}\mathbf{G}')^{-1}$.

The above G_m^{-1} will be recognized as the Minimum Norm solution obtained by minimizing the length of the model vector, or what is the same, by excluding any contributions from the null space.

It would be instructive to show that in this case the counterpart data resolution matrix, $\mathbf{N}$, is the identity matrix, and that the normalized model covariance is $\mathbf{G}'(\mathbf{G}\mathbf{G}')^{-2}\mathbf{G}$.

However, when errors in data or uncertainties in the adopted mathematical model are known to be significant, it may be desirable to seek an inverse that would simultaneously minimize both the Euclidean norm of the solution, as well as some fraction of the model covariance. The resulting solution is given by the damped minimum norm inverse $\mathbf{G}'(\mathbf{G}\mathbf{G}' + \varepsilon^2\mathbf{E})^{-1}$, so named by analogy to (24), where $\mathbf{E}$ as before represents the data error covariance matrix.

Another choice of an acceptable solution (Menke 1989) that can be obtained from available data alone without the aid of any *a-priori* assumptions, is to seek a unique solution of some weighted average of the model parameters even though a unique solution of the model parameters themselves is not possible. For example, if $\mathbf{m}_p$ be a particular partially illuminated solution of the model and $\mathbf{m}_{0i}$ the $(M - N)$ independent vectors of the null space, a general solution using arbitrary coefficient α_i can be written as follows:

$$\mathbf{m}_{\text{gen}} = \mathbf{m}_p + \sum_i^{M-N} \alpha_i \mathbf{m}_{0i}, \tag{29}$$

and

$$\langle m \rangle = \omega' \mathbf{m}_{\text{gen}} = \omega' \mathbf{m}_p + \sum_i^{M-N} \alpha_i \, \omega' \mathbf{m}_{0i}, \tag{30}$$

where $\langle m \rangle$ is the weighted average which can be made unique by determining the particular vector ω_p' which reduces all its products with $\mathbf{m}_{0i}$ to zero, so that,

$$\langle m \rangle = \omega_p' \mathbf{m}_p. \tag{31}$$

That at least one such averaging vector ω_p' does exist, can be clearly seen by recalling the i desired relations $\omega' \mathbf{m}_{0i} = 0$ which provide the necessary $(M - N)$ constraints.

Estimates of weighted averages of model parameters indeed have a clear physical meaning, and therefore significance, particularly if they possess some natural ordering in space and time, such as the depth of an irregular subsurface horizon below points on the earth's surface.

Finally, if some insightful knowledge about the nature of the model parameters is independently available or can be assumed as reasonable, one may approach the problem by building up a legitimate image of it, as an opportunity for creative design by exploiting the presumed attributes of the model in an imaginative way.

(iii) *Partially determined system*

Most real world inverse problems are essentially under-determined but often include a subspace of model parameters which may be either over-determined or inconsistently incorporated. Data from a tomography experiment, for example, may result in such a mixed system if some compartments of the discretized region lie in the path of more than adequate number of rays and others are sparsely sampled. In such a case, a subspace $S_0(\mathbf{m})$ of model parameters will remain unilluminated by the measured data, while only a subspace $S_p(\mathbf{d})$ of the data space may be spanned by the column space of $\mathbf{G}$, whose rank p is less than the minimum of N and M.

Partially determined problems of this kind can of course be handled by using the aforementioned strategies if over- and under-determined parts of the mixed system could be separated. This is indeed possible using the spectral or Singular Value Decomposition (SVD) of an $(N \times M)$ matrix $\mathbf{G}$ in terms of the two sets of orthonormal eigenvectors $\mathbf{U}$ and $\mathbf{V}$ of $\mathbf{GG}'$ and $\mathbf{G}'\mathbf{G}$ respectively, and their eigenvalue matrix Λ^2. To understand its significance, we construct an augmented $(N + M) \times (N + M)$ matrix $\mathbf{S}$ from $\mathbf{G}$ and its transpose $\mathbf{G}'$, and examine its characteristic equation

$$\mathbf{SW} = \begin{bmatrix} \mathbf{0} & \mathbf{G} \\ \mathbf{G}' & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix} = \Lambda \mathbf{W},$$

or

$$\begin{bmatrix} \mathbf{0} & \mathbf{G} \\ \mathbf{G}' & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_i \\ \mathbf{V}_i \end{bmatrix} = \lambda_i \begin{bmatrix} \mathbf{U}_i \\ \mathbf{V}_i \end{bmatrix}, \tag{32}$$

which, in turn, yield

$$\begin{aligned} \mathbf{GV} &= \mathbf{U}\Lambda, \\ \mathbf{G}'\mathbf{U} &= \mathbf{V}\Lambda, \end{aligned} \tag{33}$$

where Λ is the diagonal matrix of the eigenvalues of $\mathbf{S}$.

The above relations, in turn, yield

$$\mathbf{G}'\mathbf{G}\mathbf{V} = \mathbf{V}\Lambda^2, \tag{34}$$

$$\mathbf{G}\mathbf{G}'\mathbf{U} = \mathbf{U}\Lambda^2, \tag{35}$$

which enable us to interpret $\mathbf{V}$ and $\mathbf{U}$ as the set of orthonormal eigenvectors of $\mathbf{G}'\mathbf{G}$ and $\mathbf{G}\mathbf{G}'$ respectively, so that $\mathbf{U}\mathbf{U}' = \mathbf{I}_N$ and $\mathbf{V}\mathbf{V}' = \mathbf{I}_M$. Using these properties of $\mathbf{U}$ and $\mathbf{V}$, we can then write (43) as follows:

$$\mathbf{G} = \mathbf{U}\Lambda\mathbf{V}'; \mathbf{G}' = \mathbf{V}\Lambda\mathbf{U}'. \tag{36}$$

The vectors $\mathbf{U}$ and $\mathbf{V}$ respectively span the full spaces $S(\mathbf{d})$ and $S(\mathbf{m})$ of data and model parameters. But the rank of $\mathbf{G}$ being p, only the subspaces $S_p(\mathbf{d})$ and $S_p(\mathbf{m})$ spanned by eigenvectors $\mathbf{U}_p$ and $\mathbf{V}_p$ corresponding to non-zero eigenvalues, contain any information about predicted data or model parameters. Thus, we see that $\mathbf{G}$ can be written as

$$\mathbf{G} = \mathbf{U}\Lambda\mathbf{V}' = \mathbf{U}_p\Lambda_p\mathbf{V}'_p, \tag{37}$$

where Λ is partitioned as

$$\begin{bmatrix} \Lambda_p & 0 \\ 0 & 0 \end{bmatrix}. \tag{38}$$

The null spaces S_0 $(\mathbf{d})$ and S_0 $(\mathbf{m})$ are similarly spanned by the eigenvectors $\mathbf{U}_0$ and $\mathbf{V}_0$ which must be respectively orthogonal to $\mathbf{U}_p$ and $\mathbf{V}_p$.

In seeking an acceptable solution of $\mathbf{G}\mathbf{m} = \mathbf{d}$, or $\hat{\mathbf{m}} = \mathbf{G}_p^{-1}\mathbf{d}$, we therefore look for an inverse that would ensure that $\hat{\mathbf{m}}$ has no component in $S_0(\mathbf{m})$, and the predicted data has no component in $S_p(\mathbf{d})$. Using the condition that the null spaces S_0 both of data and the model are respectively orthogonal to the illuminated S_p vector space, it can be shown that one such inverse, also called the natural inverse of a partially determined linear system, is given by

$$\mathbf{G}_p^{-1} = \mathbf{V}_p\Lambda_p^{-1}\mathbf{U}'_p. \tag{39}$$

The respective values of $\mathbf{R}$, $\mathbf{N}$ and the normalized $\mathrm{Cov}_n(\mathbf{m})$ are, in turn, given by

$$\mathbf{R} = \mathbf{V}_p\mathbf{V}'_p, \tag{40}$$

$$\mathbf{N} = \mathbf{U}_p\mathbf{U}'_p, \tag{41}$$

$$\mathrm{Cov}_n(\hat{\mathbf{m}}) = \mathbf{V}_p\Lambda_p^{-2}\mathbf{V}'_p. \tag{42}$$

Whilst SVD provides a simple way of identifying the null vectors of $\mathbf{G}\mathbf{m} = \mathbf{d}$ and thereby the number p needed for constructing the inverse, a problem often arises from the endemic character of most real world data kernels in that the small eigenvalues decrease very smoothly making it difficult to distinguish between those that are actually zero from the near-zero ones.

These near-zero eigenvalues are the prime source of instability in the inverse solutions as can be discerned from (39), since they introduce high frequency oscillations. This problem can be circumvented either by specifying a cut-off eigenvalue which will amount to reducing the dimension of the operational eigenspace or by enhancing the near-zero eigenvalues, which will have the effect of enforcing smoothness on the inverted solution.

If only the very small eigenvalues are excluded, the solution will be generally close to the natural solution and have good variance, as the covariance of the estimated model (42) is extremely sensitive to the smallest non-zero eigenvalue. However, the model and data resolution would deteriorate, exposing the inevitable trade-off between resolution and variance, which every inverse problem has to deal with by an appropriate parameterization of the model.

If on the other hand, one chooses to enhance the eigenvalues by an amount proportional to the error in the data vector, it would be desirable to seek an inverse that would simultaneously minimize the norm of the data misfit vector and that of the solution roughness represented by the vector of some-order difference of the parameter components. As before, trade-off exists between the error misfit and the degree of smoothness. The resulting inverse in this case is given by

$$(\mathbf{G}'\mathbf{G} + \varepsilon^2\mathbf{H})^{-1}\mathbf{G}', \tag{43}$$

where $0 < \varepsilon^2 < 1$ is the trade-off parameter, and $\mathbf{H} = \mathbf{K}'\mathbf{K}$, $\mathbf{K}$ being the finite difference coefficient matrix. For example, the following $\mathbf{K}$ matrix of order $(M-2)\times M$ would represent a second-order difference.

$$\mathbf{K} = \begin{bmatrix} -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & . \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & . \\ 0 & 0 & -1 & 2 & -1 & 0 & . & . & . \\ 0 & 0 & 0 & -1 & 2 & -1 & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \end{bmatrix}_{(M-2)\times M}.$$

It will be recognized that $\mathbf{K} = \mathbf{I}$ reproduces the damped least squares solution, also known as ridge regression or Marquardt, while $\mathbf{K}$ constituted from first-order differentials leads to a solution known as Occam's solution.

10. Generalized inverse

It may be mentioned here that the various inverses discussed above belong to a class of generalized inverses $\mathbf{G}^{-g}$ which determine exact or approximate solutions of basic equation $\mathbf{Gm} = \mathbf{d}$ depending on the structure of $\mathbf{G}$. They satisfy the relation $\mathbf{G}\mathbf{G}^{-g}\mathbf{G} = \mathbf{G}$, in addition to satisfying additional conditions or constraints imposed on them. For example, in deriving $\mathbf{G}_m^{-1}$ for the under-determined problem, we made the demand that the model resolution matrix $\mathbf{R}$ which affects a weighted averaging of the true model in yielding the estimated parameters be as close to an identity matrix as possible, and proceeded to minimize $\| \mathbf{I} - \mathbf{R} \|$ or $\| \mathbf{I} - \mathbf{G}_m^{-1}\mathbf{G} \|$. However when data or model parameters have a natural ordering, one would prefer that the off-diagonal values in $\mathbf{R}$ are inversely proportional to their distance from the diagonal. This would make the model estimates more nearly representative of the local averages. A generalized inverse designed to accomplish this by damping the side lobes of the row spread in $\mathbf{R}$ was constructed by Backus and Gilbert and bears their name.

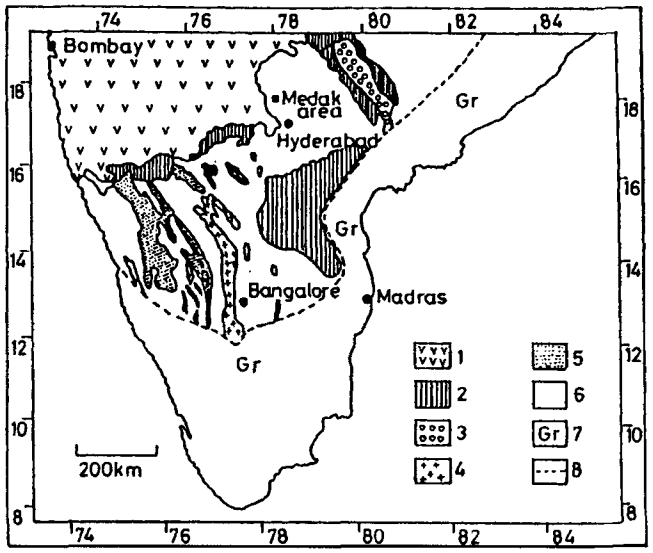

Figure 2. Simplified geological map of the southern Indian peninsular shield from Sarvothaman & Leelanandam (1987). (1) Deccan basalts; (2) Proterozoic formations; (3) Gondwanas; (4) Younger granites; (5) Dharwar Supergroup metasediments/volcanics; (6) Unclassified crystallines – granite-gneiss; (7) Granulite facies; (8) Granite-granulite boundary.

11. Design of experiments

We have seen that most real world problems are inverse and do not possess a unique solution. The challenge in all these cases is to design the best approximate solution. However, inverse formalisms are also exciting because they provide a creative opportunity to explore the character of the data spaces in which a particular model space may be mapped, as well as the view of the model space from that of measured data. This visualization is indeed provided by the form of the data kernel matrix **G** which is partly dependent on the way data is measured but not on the data itself. The illuminating potential of **G** can therefore be explored in order to bring out all the intrinsic limitations of possible solutions and their significance and quality, in advance of the data collection exercise. In situations where the solution can or is to be obtained on the basis of experimental data, such an analysis provides a creative opportunity to design the experiment appropriately which would, in turn, effect desirable structuring of the data kernel. Indeed, analysis of data inversion and adaptive experiment design have very high potential, not always fully exploited, to push the sharpness of solutions to the limit of resolution theoretically permitted in a given situation.

Finally, actual results of seismogram inversion to obtain the shear velocity structure $V_s(z)$ beneath the Archaean Granites around Hyderabad (figure 2) (Gaur & Priestley 1997) is presented below as an example of inverse modelling.

12. Velocity structure beneath Hyderabad

The groundmotion record **G(t)** at a site generated by an earthquake (seismogram) is known to begin with the first arriving compressional waves (P) followed by its reverberations generated by the layered structure beneath the recording site (a few hundred kilometres deep) and later by the slower shear and surface waves. The reverberations (figure 3a)

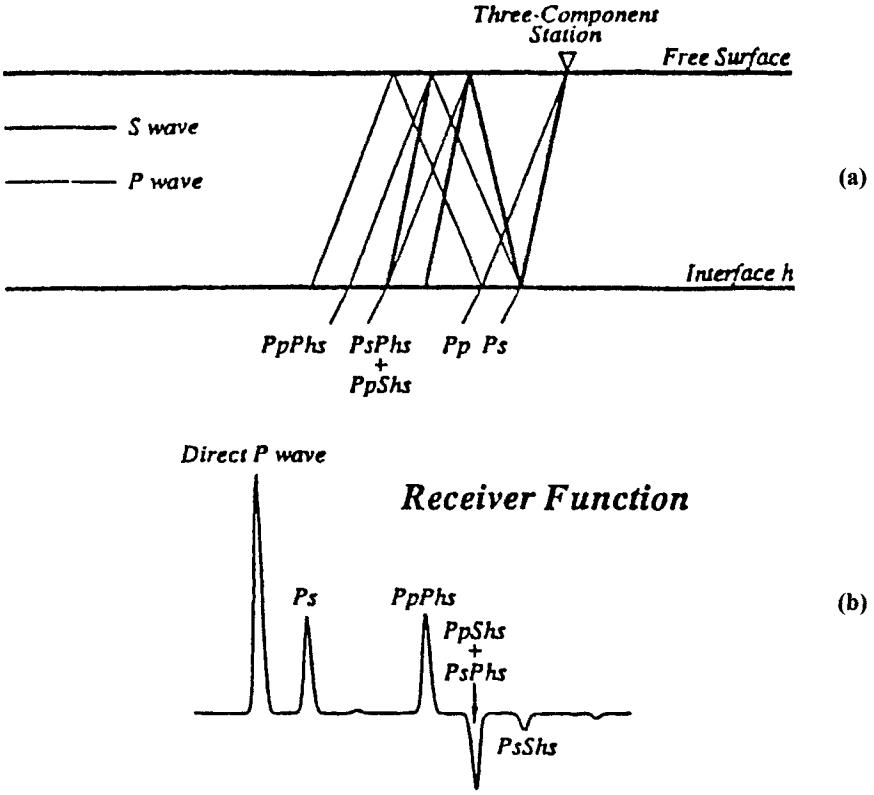

Figure 3. (**a**) Ray diagram showing the main P-S converted phases which comprise the Receiver Function for a single layer overlying a half space; (**b**) showing the corresponding waveform.

that represent a convolution of the incident (P) wave with the velocity structure **m(z)** beneath the recording site are themselves quite weak, but yield discernible signals of P to S converted phases (P_s) after deconvolution of the vertical component of the P wave from the horizontal component of the reverberations (figure 3b) recorded before the onset of the next dominating shear (S) waves, which being slower than the P waves lag behind it depending on the distance of the earthquake from the site. These selectively extracted converted phases are called Receiver Functions, denoting an attribute of the seismogram that is wholly reflective of the properties of the earth directly beneath the receiver (recording seismograph).

The Receiver Function thus obtained over a horizontally stratified earth appears (Ammon 1991) as a scaled version of the radial component of ground motion with the P multiples entirely eliminated. The deconvolution is accomplished by division in the Fourier transformed frequency domain and retransformation of the quotient back to the time domain. Let $V(t)$, $R(t)$ represent a ray description of the vertical and radial components of ground motion respectively and $V(W)$, $R(W)$ their Fourier domain counterparts. Then,

$$V(t) = \sum_k v_k S(t - t_k), \tag{44}$$

$$R(t) = \sum_k r_k S(t - t_k), \tag{45}$$

and the Fourier domain Receiver Function

$$H(w) = R(w)/V(w), \tag{46}$$

where $S(t)$ is the source time function and t_k the instant of arrival of the kth ray, $k = 0$ representing the direct P phase.

In practice, of course, one must band limit $V(w)$, $R(w)$ by using a Gaussian filter $F(w)$ of appropriate width, and also forestall any instability in deconvolution that may arise from division by spectral values of $V(w)$ that are either zero or very small. This can be done by constraining the lowest value of the denominator to remain above a practicable value set by a small parameter called the water level parameter (Clayton & Wiggins 1976).

Accordingly, one may rewrite (46) as

$$H(w) = [R(w)V^*(w)F(w)]/\phi(w), \tag{47}$$

where, V^* is the complex conjugate of V, and

$$F(w) = f \exp(-w^2/4a^2), \tag{48}$$

is the Gaussian filter normalized to unit amplitude in the time domain by the factor f, and having a width a, and,

$$\phi(w) = \max[V(w)V^*(w), c.\max\{V(w)V^*(w)\}], \tag{49}$$

$\phi(w)$ is set to the greater of the two quantities in parenthesis on the RHS of (49).

The corresponding time-domain Receiver Function $h(t)$ can then be shown (Ammon 1991) to be given by,

$$h(t) = (r_0/v_0)[\delta(t) + r_{sk}(t - t_k)],$$

where r_0 and v_0 are the vertical and horizontal amplitudes of the direct P phase, and r_{sk} those of the various converted shear wave phases.

It may be noted that a judicious choice of a and c is critical in determining the shape of the Receiver Function owing to their effects on the waveform spectra. In addition, the choice of c has to be a trade-off between the extent of such modification and stability of the deconvolution process. Options for such a choice can be delineated by abstracting Receiver Functions for a suite of a and c values and examining their signal quality, particularly of the self-deconvolved vertical component.

The factor (r_0/v_0) scales the radial Receiver Function and therefore clearly depends on the epicentral distance as well as on the extent of contamination by scattered waves. In the original treatment of Receiver Function analysis, this quantity used to be normalized to unity, thereby obliterating the effect of varying epicentral distances (through the incidence angle of P waves), which is advantageous when stacking Receiver Functions from events covering a large geographical spread. This approach, however, forfeits valuable information implicit in this quantity, particularly that concerning the near surface velocity structure and estimates of contamination by scattering. Modern analysis of Receiver Functions, following Ammon (1991), therefore retains their true amplitudes and preserves the ratio (r_0/v_0) of the radial and vertical amplitudes of the P wave. In practice, this is accomplished by deconvolving the vertical P wave from itself, which is seldom the expected

delta function owing to spectrum modification by the water level and Gaussian parameters, and using its maximum amplitude to normalize the radial Receiver Function.

Before inverting Receiver Functions for velocity structure, it is desirable to reduce the effect of random errors introduced by the computational processes and earth noise. This can be achieved by stacking a number of Receiver Functions arising from closely spaced events (in our case within 2°). If records of an adequate number of such close events are not available and the signal-to-noise-ratio of the Receiver Function is unsatisfactory, recourse could be had to form composite Receiver Functions from a wider geographical spread of earthquakes, but in that case it would be necessary to suppress the range effect on individual Receiver Functions by normalizing the zero lag amplitude ratio (r_0/v_0) to unity.

It may be remarked here that whilst the tangential Receiver Function over a horizontal earth should ideally be zero, small amplitudes would generally appear in it owing to the presence of small levels of scattered energy in the seismograms. Significant finite amplitudes of tangential Receiver Functions therefore suggest departures from horizontal stratification and systematic variations of coherent phases, if present, can be used to infer the nature of inhomogeneity. The relative amplitudes of tangential and radial Receiver Functions are therefore an important guide in parameterizing the velocity structure to be inverted for.

Inversion of Receiver Functions to obtain the shear wave velocity structure is accomplished by parameterizing the latter in a manner such that the inverse problem is overdetermined i.e., the number n of data points d_j, is greater than the number m of unknown parameters of the velocity structure. For a layered earth model, this is normally defined as consisting of m layers each of constant thickness and uniform velocity m_k. The thickness of the layer is chosen to be such that it can be clearly resolved by the data phases being inverted. For example, in the case of predominantly 1.0 s period shear waves having a wave length of over 3.5 km in the upper lithosphere, a layer thickness of 1–2 km should be quite satisfactory. When the earth model is so defined, determination of the velocity structure is reduced to the problem of estimating the unknown velocities in each of the m layers.

13. Data and analysis

Three-component broadband seismograms recorded at Hyderabad between May 1989 and March 1996 were obtained from the Geoscope Data Centre at Paris. Both radial

Table 1. Data for the period 1990–95.

Year	Julian day	Latitude	Longitude	Depth (km)	Magnitude (m_b)
1990	150	45.841	26.668	89	6.7
1990	364	−5.097	150.967	179	6.6
1992	246	−6.046	112.138	625	5.9
1992	255	−6.087	26.651	11	6.7
1992	292	−6.279	130.214	119	5.8
1993	015	43.30	143.691	102	6.9
1994	104	−6.587	129.771	166	5.8
1994	194	−7.532	127.770	159	6.5
1994	271	−5.786	110.352	638	5.9
1994	319	−5.589	110.186	561	6.2
1995	235	18.856	145.218	595	6.3

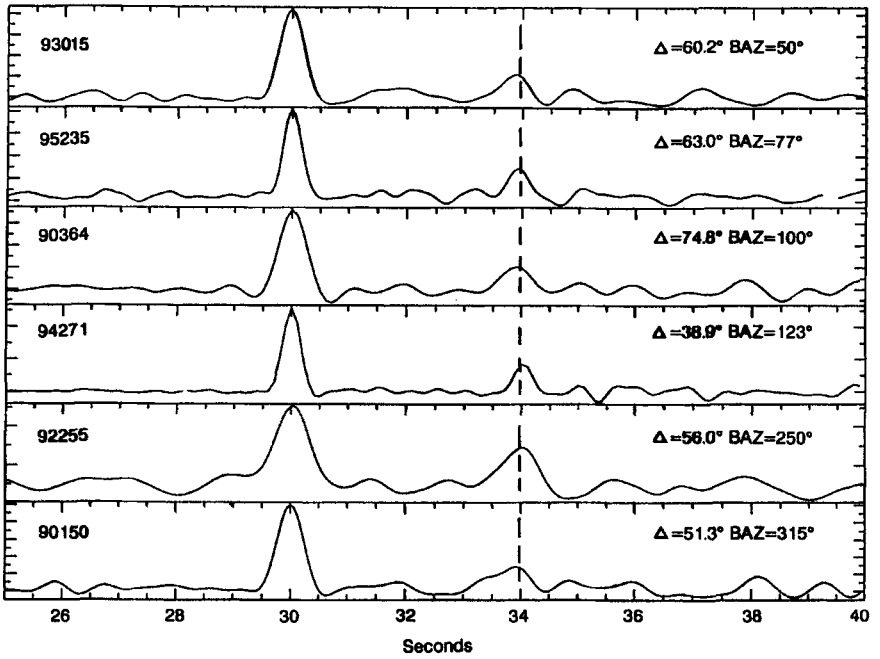

Figure 4. HYB single-event radial Receiver Functions ($c = 0.001$, $a = 5.0$) plotted as a function of event azimuth. The higher frequency Gaussian was chosen in this case to better resolve time differences in P_s-P delay. The distance and azimuth of the events from HYB are given to the right of each trace.

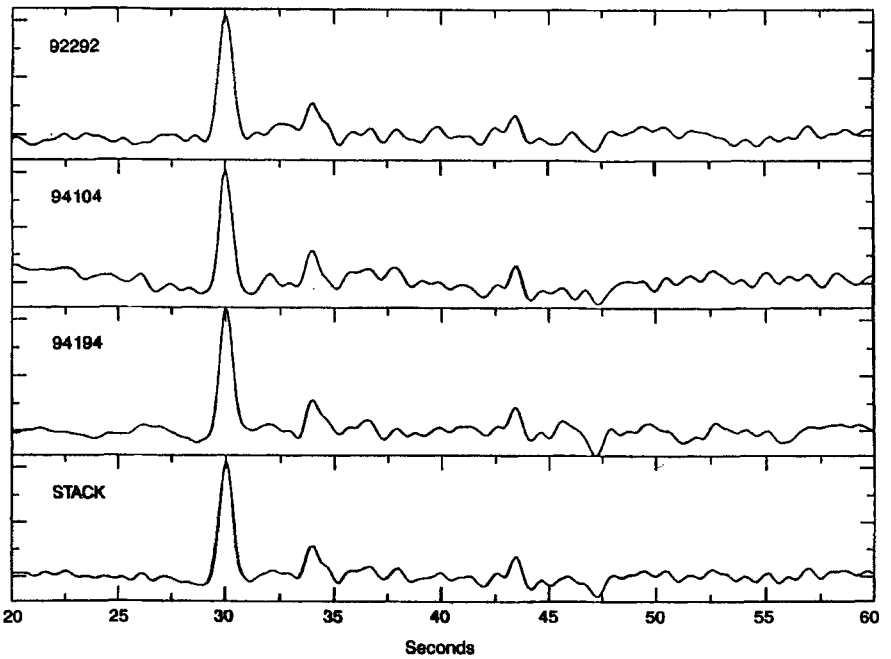

Figure 5. Three single-event radial Receiver Functions ($c = 0.001$, $a = 2.5$) and the resulting radial stack for events lying between latitude 6.35 and 7.5° south and longitude 128 and 130° east.

and transverse Receiver Functions were generated from a large number of those for which broadband records were available, using different values of the Gaussian width a (5, 2.5, 1.0) and the water-level parameter c (0.0001, 0.001, 0.01, 0.1). After a close scrutiny of the signal quality of these Receiver Functions, 11 of them (table 1), were selected for further analysis.

Figure 4, shows the Receiver Functions ($c = 0.001$, $a = 5.0$) obtained from six events, including both shallow and intermediate focus, from different azimuths. Their radial components show a remarkable coherence of phases. The time delays (3.0 ± 0.1 s) of the P_s phase varies systematically with azimuth. This variation can be explained by two different velocity models: variations in the velocity structure of the mid to lower crust or variations in the Moho depth. However, possible variations, as will be shown later, are estimated to be small, and coupled with the small relative amplitudes ($< 10\%$) of the tangential Receiver Functions (figure 4), warrant the crust to be modelled as a predominantly one-dimensional structure on which small perturbations may be superposed.

Figure 5 shows the radial Receiver Functions ($c = 0.001$, $a = 2.5$) of 3 events clustered between 128° and 130° E and 6.35° and 7.5° S, as well as the stacked radial Receiver Function from these three events. Figure 6 compares this with another radial Receiver

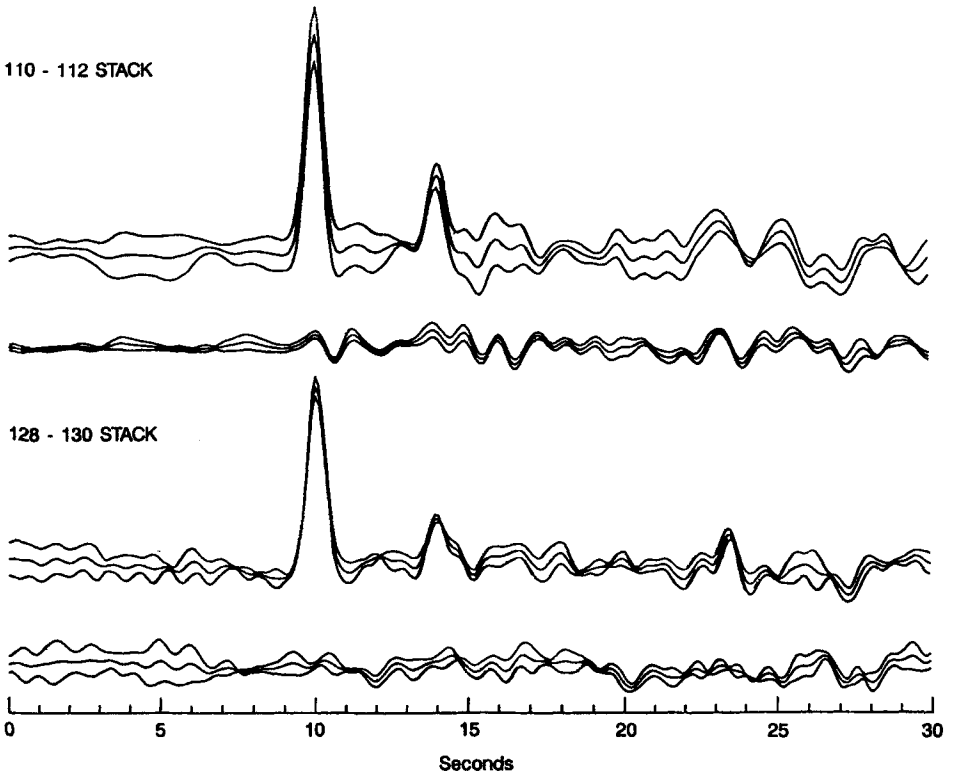

Figure 6. Comparison of two radial (the upper waveform in each) and tangential Receiver Function stacks used in the simultaneous inversion for the crustal structure beneath HYB. The upper pair of Receiver Functions are from three events lying between 5.6 and 6.0° south and longitude 110 and 112° east. The lower pair of Receiver Functions are from the events shown in figure 5. The mean and ±1 standard deviation are shown.

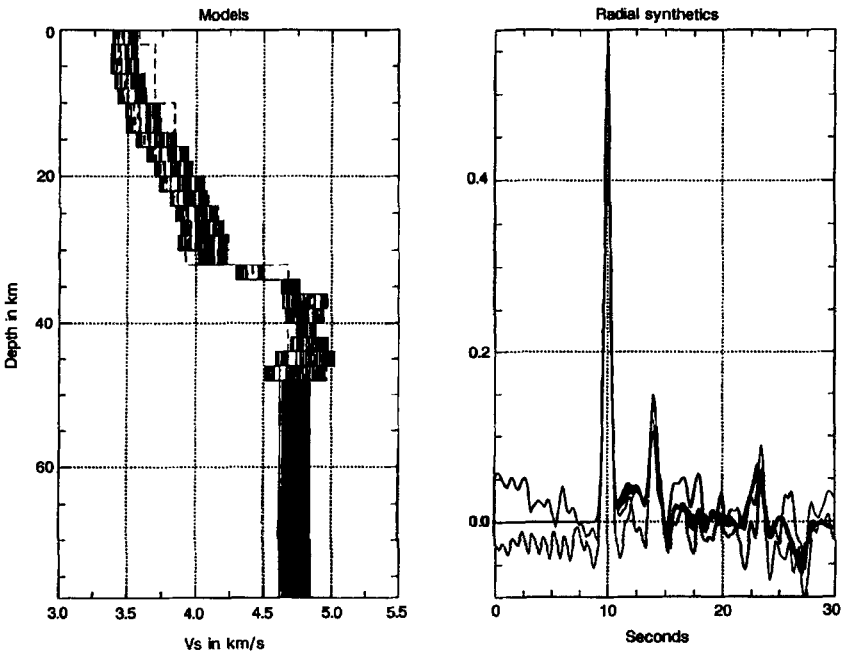

Figure 7. (**a**) Initial crustal models from the simultaneous inversion of the two radial Receiver Functions shown in figure 6. The dotted line denotes the initial starting model; (**b**) The fit of the resulting model synthetic radial Receiver Functions to the ± one standard deviation bounds.

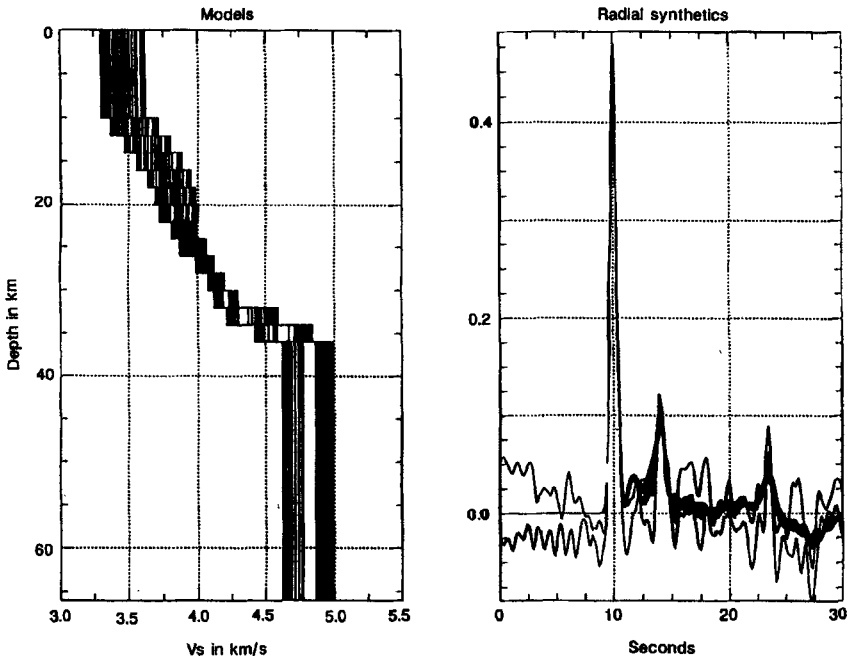

Figure 8. (**a**) Crustal models from the simultaneous inversion of the two radial Receiver Functions shown in figure 6, after constraining the surface layer to be 10 km thick; (**b**) The fit of the resulting model synthetic radial Receiver Functions to the ± one standard deviation bounds.

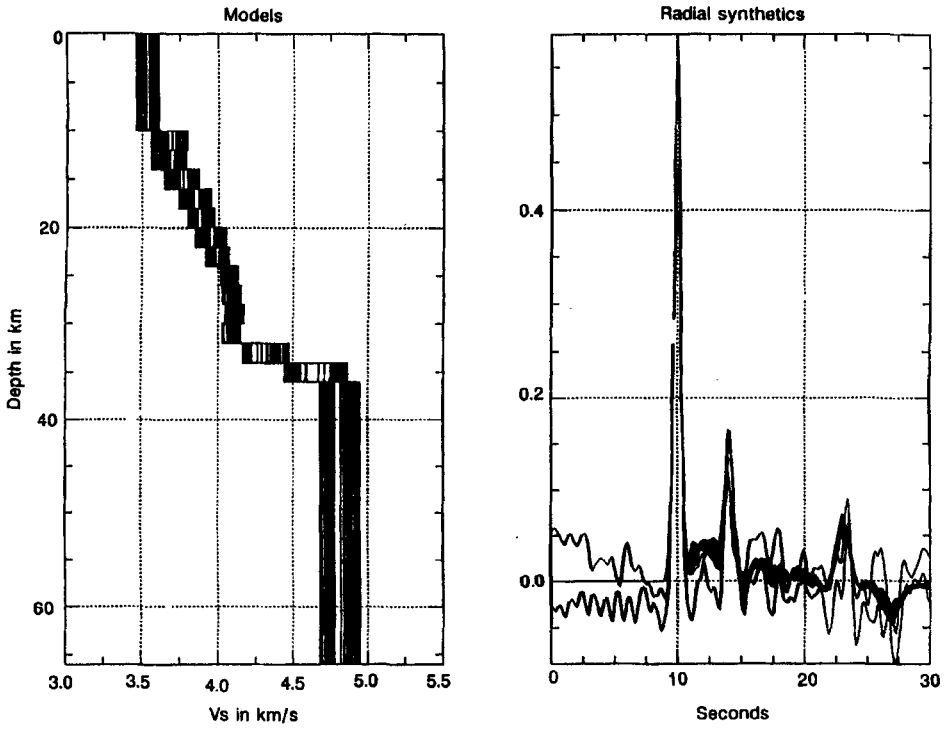

Figure 9. (**a**) Final crustal models from the simultaneous inversion of the two radial Receiver Functions shown in figure 6, after constraining the surface layer to be 10 km thick, and the shear wave velocity in this layer to lie in the range 3.46 to 3.61 km/s; (**b**) The fit of the resulting model synthetic radial Receiver Functions to the ± one standard deviation bounds.

Function ($c = 0.001$, $a = 2.5$) stack composed of 3 events lying between 110° and 112° E and 5.6° and 6° S.

For inverting the stacked Receiver Functions, we adopted a starting reference model (dotted line in figure 7) based on some knowledge of the surface and average crustal velocities around Hyderabad (personal communication with Sri Nagesh and S S Rai). From this, we generated a family of 20 new initial models (Ammon *et al* 1991) by adding to the reference model, a 2-component perturbation vector, with the purpose of obtaining a group of models which whilst being significantly different from each other would be expected to share some essential attributes of the parent model. The 2-component vector used to perturb the initial model consisted of a cubic perturbation vector scaled to a maximum velocity perturbation of 1.0 km/s, and a random velocity change up to 20% of cubic perturbation.

Synthetic Receiver Functions were then computed using the same values a and c (2.5, 0.001) selected earlier, and a P wave ray parameter (horizontal slowness) of 0.063 s/km corresponding to a source approximately 60° away, to iterate the inversion process. Figure 7 shows the crustal models obtained after simultaneous inversion of the two stacked Receiver Functions shown in figure 6, for each of the initial models. Examination of several hundred such inversions, however, showed that the S wave velocity in the first 10 km was substantially uniform (3.48 ± 0.12 km/s).

This warranted investigation of a more constrained solution space in which shear wave velocity in the upper 10 km whilst free to excurse between the limits prescribed earlier, was held uniform. The result is shown in figure 8. The velocity range in the top layer that characterizes about 75% of all these models lay within a narrower limit $3.46 \leq V_s \leq 3.61$ km.

Assuming that this range of top layer velocity shared by a majority of velocity solutions may approximate the real one more closely, we then imposed an additional constraint that the shear wave velocity in the top 10 km thick crustal layer should vary only within these limits.

The inverted velocity model of the Hyderabad crust, subject to constraints suggested by these data, is shown in figure 9.

References

Ammon C J 1991 The isolation of receiver effects from teleseismic *P* waveforms. *Bull. Seism. Soc. Am.* 6: 2504–2510

Clayton R W, Wiggins R A 1976 Source shape estimation and deconvolution of teleseismic body waves. *Geophys. J.R. Astron. Soc.* 47: 151–177

Gaur V K, Priestley K 1997 Shear wave velocity structure beneath the Archaean granites around Hyderabad, inferred from Receiver Function analysis. *Proc. Indian Acad. Sci. (Earth Planet. Sci.)* 106: 1–8

Menke W 1989 *Geophysical data analysis: Discrete inverse theory* (London: Academic Press)

Twomey S 1977 *Introduction to the mathematics of inversion in remote sensing and indirect measurements: developments in geomathematics* (Amsterdam: Elsevier Scientific) vol. 3