ELSEVIER

Physics of the Earth and Planetary Interiors 98 (1996) 269–282

PHYSICS OF THE EARTH AND PLANETARY INTERIORS

Optimal one-dimensional inversion and bounding of magnetotelluric apparent resistivity and phase measurements

Robert L. Parker [a,*], John R. Booker [b]

[a] *Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, UCSD, La Jolla, CA 92093, USA*
[b] *Geophysics Program, Box 351650, University of Washington, Seattle, WA 98195, USA*

Received 21 November 1995; revised 30 January 1996; accepted 16 April 1996

Abstract

The properties of the log of the admittance in the complex frequency plane lead to an integral representation for one-dimensional magnetotelluric (MT) apparent resistivity and impedance phase similar to that found previously for complex admittance. The inverse problem of finding a one-dimensional model for MT data can then be solved using the same techniques as for complex admittance, with similar results. For instance, the one-dimensional conductivity model that minimizes the χ^2 misfit statistic for noisy apparent resistivity and phase is a series of delta functions.

One of the most important applications of the delta function solution to the inverse problem for complex admittance has been answering the question of whether or not a given set of measurements is consistent with the modeling assumption of one-dimensionality. The new solution allows this test to be performed directly on standard MT data. Recently, it has been shown that induction data must pass the same one-dimensional consistency test if they correspond to the polarization in which the electric field is perpendicular to the strike of two-dimensional structure. This greatly magnifies the utility of the consistency test.

The new solution also allows one to compute the upper and lower bounds permitted on phase or apparent resistivity at any frequency given a collection of MT data. Applications include testing the mutual consistency of apparent resistivity and phase data and placing bounds on missing phase or resistivity data. Examples presented demonstrate detection and correction of equipment and processing problems and verification of compatibility with two-dimensional B-polarization for MT data after impedance tensor decomposition and for continuous electromagnetic profiling data.

1. Introduction

The following paper is submitted to this special issue of PEPI honoring George Backus not because George has worked in the area of electromagnetic sounding (he never has), but because both of the authors have been profoundly influenced by his approach to science. We hope this paper reflects that influence.

Magnetotelluric (MT) data are commonly observed to depend on the coordinate system in which they are measured or to vary significantly between adjacent sites. Such observations make structural interpretation using a one-dimensional modeling assumption dubious and modern electromagnetic sounding experiments are usually interpreted in terms of two- or three-dimensional conductivity structures.

Inversions of data founded on a one-dimensional approximation for conductivity still have a place in structural interpretation. For instance, after the dis-

* Corresponding author.

PII S0031-9201(96)03191-3

torting effect of very shallow structure is accounted for, the shallow structure of sedimentary basins can be nearly one-dimensional (see Jones, 1988). If MT data are collected with dipoles that are comparable to the sounding depth, the measured response of two- or three-dimensional structures can be one-dimensional (Torres-Verdin and Bostick, 1992a, Torres-Verdin and Bostick, 1992b). Finally, once the distorting effect of crustal structure is removed, the response of the mantle below about 400 km appears to be one-dimensional within experimental error (e.g. Egbert and Booker, 1992).

However, the most important application of theoretical work on the one-dimensional inverse problem may be in data validation. Recently, Weidelt and Kaikkonen (1994) have shown that the criteria for consistency of a collection of MT impedance measurements at a single site is the same for B-polarization (data polarized with the electric field perpendicular to the strike of a two-dimensional structure) as it is for a one-dimensional MT sounding. In other words, one can test whether data are consistent with the assumption that they are B-polarized by applying exactly the same test that one would apply to test whether data are consistent with an assumption of one-dimensionality.

Testing for consistency of complex admittance data with the assumption of one-dimensionality is well understood in terms of solutions that have a finite number of infinitely thin, infinitely conducting layers, models said to be in the class D^+ (Parker, 1980). However, a great deal of electromagnetic induction data are in the form of apparent resistivity and impedance phase rather than the real and imaginary parts of the complex admittance used in the earlier work. In principle, these standard MT data can be converted to complex admittance because apparent resistivity is simply related to the magnitude of the complex admittance and impedance phase is simply related to the phase of the complex admittance. In practice, however, the relation between the uncertainties in standard MT data and in the real and imaginary parts of the complex admittance makes the conversion undesirable or impossible.

The uncertainties in the magnitude and phase of a complex quantity can usually be taken to be uncorrelated. Indeed, estimated uncertainties in phase measurements are sometimes much larger than in apparent resistivity measurements and the phase data can even be missing. This is because the estimation of phase is less robust to many types of equipment and time series processing problems than is apparent resistivity. On the other hand, the uncertainties in the real and imaginary parts of a complex quantity are extremely unlikely to be uncorrelated. Even when the magnitude of the complex admittance is well determined, the uncertainties in its real and and imaginary parts will be large or infinite when the uncertainties in phase are large or a phase datum is missing. The purpose of this paper is therefore to explore the consistency of one-dimensional systems when the data are apparent resistivity and impedance phase, rather than the real and imaginary parts of the complex admittance.

2. The integral representation

Our starting point is the fact that the MT response of any one-dimensional conductivity profile can be matched arbitrarily well at a finite number of frequencies by the response of a finite system of delta functions (Parker, 1980, Parker, 1994, Chapter 5). The admittance of such a system at radian frequency ω can be written

$$c(\omega) = a_0 + \sum_{n=1}^{N} \frac{a_n}{\lambda_n + \mathrm{i}\omega} \tag{1}$$

where the real constants $\lambda_n \geq 0$ and $a_n > 0$. In the work just mentioned, the measurements were assumed to be the real and imaginary parts of c. Let us suppose instead that we have measured the apparent resistivity ρ_a, and the impedance phase Φ, at a finite number of frequencies. These are related to the admittance by

$$\rho_a = \mu_0 \omega |c|^2 \tag{2}$$

and

$$c = |c| e^{\mathrm{i}(\Phi - \pi/2)} \tag{3}$$

As discussed in the Introduction we would like to treat ρ_a and Φ as separately measured quantities, rather than the real and imaginary parts of c. To

affect this separation we consider the natural log of the admittance:

$$G(\omega) = \ln c(\omega) = \ln|c(\omega)| + i(\Phi(\omega) - \pi/2) \tag{4}$$

$$= \frac{1}{2}\ln(\rho_a(\omega)/\mu_0\omega) + i(\Phi(\omega) - \pi/2) \tag{5}$$

Thus the real and imaginary parts of G are simply related to the two standard types of MT observations.

Next we look at the behavior of G in the complex ω plane. We assume that $a_0 > 0$ for the moment, which means there is an insulating layer at the surface (this interpretation follows from the way c behaves as ω tends to infinity). We shall return to the case with $a_0 = 0$ later. It is clear from Eq. (1) that the admittance c can be written as the ratio of two polynomials of degree N:

$$c(\omega) = a_0 \prod_{n=1}^{N} \frac{(\omega - i\nu_n)}{(\omega - i\lambda_n)} \tag{6}$$

$$= a_0 \prod_{n=1}^{N} \frac{(1 + \nu_n/i\omega)}{(1 + \lambda_n/i\omega)} \tag{7}$$

where $\omega = i\nu_n$ are the zeros of c. From Eq. (1) ν_n are real and positive, and the zeros interlace the poles at $i\lambda_n$ as seen from the positivity of the residues a_n in the expansion. We order them so that

$$0 \le \lambda_1 < \nu_1 < \lambda_2 < \nu_2 < \cdots \lambda_N < \nu_N \tag{8}$$

Combining Eq. (4) and Eq. (7) we see that

$$G(\omega) = \ln a_0 + \sum_{n=1}^{N} \ln\left(\frac{1 + \nu_n/i\omega}{1 + \lambda_n/i\omega}\right) \tag{9}$$

$$= \ln a_0 + \sum_{n=1}^{N} \ln(\omega - i\nu_n) - \sum_{n=1}^{N} \ln(\omega - i\lambda_n) \tag{10}$$

This expression shows that G is analytic in the complex ω plane except for the logarithmic branch points on the positive imaginary axis at $i\nu_n$ and $i\lambda_n$. We choose the branch cuts to join the successive poles and zeros of c, leading to the picture shown in Fig. 1, which depicts the situation when the number of poles is two. Because the number of poles and zeros is equal, our arrangement insures there is no cut extending to infinity; equivalently, G is analytic at infinity. We also see from Eq. (9) that, as $|\omega|$ becomes large, $|G - \ln a_0|$ tends to zero.

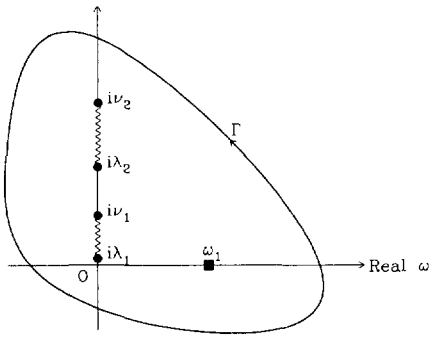

Fig. 1. Singularities of the integrand of Eq. (12) in the complex ω plane, for the case $N = 2$. The branch points at the zeros and poles of c are shown on the imaginary axis, together with the arrangement of the cuts. An isolated pole appears on the real axis. Also shown is a large contour surrounding all the singularities.

In the theory for the calculation of the best-fitting one-dimensional conductivity model (Parker, 1980), a central role is played by the representation of c as an integral over the spectral function, an expression first given by Weidelt (1972):

$$c(\omega) = \int_0^\infty \frac{da(\lambda)}{\lambda + i\omega} \tag{11}$$

We require a similar integral representation for G. To find it, we consider a contour integral on Γ, a path that encloses all the branch points and the point ω_1, which is real and positive:

$$I = \int_\Gamma \frac{G(\omega) - \ln a_0}{\omega - \omega_1} d\omega \tag{12}$$

If Γ is expanded to an arbitrary large circle centered at the origin, the integrand, which is analytic, vanishes faster than $|\omega|^{-1}$, and hence the contour integral must vanish as the radius is increased. This remains true when the contour is shrunk provided the path does not cross any of the cuts or the singularity at ω_1. Let us imagine now shrinking the contour onto the real and imaginary axes. The only contribution from the parts on the real axis comes from the pole at ω_1 because we can delete the two cancelling pieces leading to and from ω_1. On the imaginary axis we have a path down the negative real side and up the positive real side, as shown in Fig. 2. Of

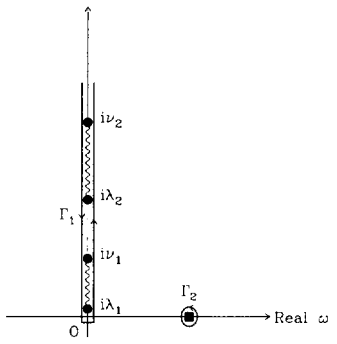

Fig. 2. As in Fig. 1, but with contour Γ separated into two parts: Γ_1 collapsed onto the imaginary axis, and Γ_2, concentrated around the pole at ω_1.

course, the integral retains the same value, namely zero, and so

$$I = 0 = \int_{\Gamma} = \int_{\Gamma_1} + \int_{\Gamma_2} \tag{13}$$

$$= 2\pi \mathrm{i}\left[G(\omega_1) - \ln a_0\right] + \left(\int_{\mathrm{i}\infty}^{-\varepsilon(1+\mathrm{i})} + \int_{-\varepsilon(1+\mathrm{i})}^{\varepsilon(1-\mathrm{i})} + \int_{\varepsilon(1-\mathrm{i})}^{\mathrm{i}\infty}\right) \tag{14}$$

where ε is real; in Eq. (14) we applied Cauchy's Residue Theorem (see, e.g. Korevaar, 1968) to the integral on Γ_2. As we allow the path Γ_1 to approach the imaginary axis, we must take the limit as ε tends to zero through real values; we find

$$0 = 2\pi \mathrm{i}\left[G(\omega_1) - \ln a_0\right] + \int_0^{\infty} \frac{\lim_{\varepsilon \to 0} G(\mathrm{i}\lambda + \varepsilon) - G(\mathrm{i}\lambda - \varepsilon)}{\mathrm{i}\lambda - \omega_1} \mathrm{id}\lambda \tag{15}$$

The function in the numerator of the integrand gives the discontinuity in G as one crosses the imaginary axis. The real part of G is continuous (it is $\ln|c|$), except at the poles and zeros, and therefore there is no contribution to the integral from the real part, except possibly from the branch points themselves. However, a short calculation shows the contribution from the branch points vanishes also, as $\lim_{|z|\to 0}|z|\ln|z| = 0$. To find the jump in imaginary part across the cut we reason as follows. As one circumnavigates a simple, isolated logarithmic branch point in a clockwise direction, the imaginary part changes by -2π, and this amount is independent of the radius of the circuit. Eq. (10) shows the system of cuts results from the sum of a set of logarithms each with a simple branch point. Hence the jump across a cut is just $-2\pi\mathrm{i}$, which we see if we move in a clockwise direction passing just above any of the branch points $\mathrm{i}\nu_n$. (We obtain the same answer passing in an anticlockwise direction below any $\mathrm{i}\lambda_n$, because of the negative sign in front of the corresponding log.) It follows that the numerator in the integrand in Eq. (15) is either zero, on the portions of the imaginary axis where G is analytic, or $-2\pi\mathrm{i}$ on the cuts. To summarize this behavior, let us introduce a real function $\mu(\lambda)$ that can take as its values only zero or one:

$$\mu(\lambda) = \begin{cases} 1, & \lambda_n \le \lambda \le \nu_n,\ n = 1,2\cdots N \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

Then we can write Eq. (15) as

$$0 = 2\pi \mathrm{i}\left[G(\omega_1) - \ln a_0\right] + \int_0^{\infty} \frac{-2\pi \mathrm{i}\mu(\lambda)}{\mathrm{i}\lambda - \omega_1} \mathrm{id}\lambda \tag{17}$$

.Dropping the superfluous subscript on ω and rearranging, we obtain

$$G(\omega) = \ln a_0 + \int_0^{\infty} \frac{\mu(\lambda)}{\lambda + \mathrm{i}\omega} \mathrm{d}\lambda \tag{18}$$

This is the integral representation we require.

Before discussing the use of this expression for solving the inverse problem, we turn briefly to the case in which $a_0 = 0$ in Eq. (1), which arises in the presence of a conductor right at the surface. Clearly Eq. (6) and subsequent steps are invalid for this case. To obtain an equivalent representation, we treat $\mathrm{i}\omega c(\omega)$ in place of $c(\omega)$. The purpose of this modification is to create a function tending to a constant value at infinity. We define

$$\tilde{G}(\omega) = \ln(\mathrm{i}\omega c(\omega)) \tag{19}$$

Then in place of Eq. (5) we have

$$\tilde{G}(\omega) = \frac{1}{2}\ln(\omega \rho_a(\omega)/\mu_0) + \mathrm{i}(\Phi(\omega) + \pi/2) \tag{20}$$

The reasoning that leads to Eq. (18) now applies

equally well to $\tilde{G}$; we find the corresponding representation:

$$\tilde{G}(\omega) = \ln A - \int_0^\infty \frac{\mu(\lambda)}{\lambda + i\omega} d\lambda \tag{21}$$

where the real constant A is given by

$$A = \sum_{n=1}^{N} a_n \tag{22}$$

3. Matching observations: exact data

To use either of these representations we must apply linear or quadratic programming. For clarity we shall restrict the discussion to Eq. (18), as the argument for Eq. (21) is virtually identical. The simplest case is that of exact measurements. Let us suppose we are given a total of N values of ρ_a and Φ at a number of frequencies; there need not be both kinds of observation at every frequency, or indeed, any observations at all of phase. There must be at least one apparent resistivity datum to set the magnitude scale of the solution, but there need be no more than one. The inverse problem to be solved is to decide whether the set of data values is compatible with any one-dimensional conductivity profile. The solution is similar to that given by Parker (1980) for c, applied to Eq. (18). A sequence of approximations is envisaged in which the integral is approximated by a sum over K samples in λ. Real and imaginary parts of Eq. (18) are set up as separate conditions to be satisfied, applying Eq. (5) to obtain G:

$$\alpha \mathbf{1} + \boldsymbol{L}_1 \boldsymbol{m} = \boldsymbol{g}_1; \quad \boldsymbol{L}_2 \boldsymbol{m} = \boldsymbol{g}_2 \tag{23}$$

where $\boldsymbol{m}$ is a K-vector of samples of the function μ, and $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are matrices approximating the real and imaginary parts of the integration operation; α is $\ln a_0$, and $\mathbf{1}$ is a vector consisting of all ones, with same dimension as $\mathbf{g}_1$, the vector of real parts of G; $\mathbf{g}_2$ is a vector of imaginary parts of G. The unknown $\boldsymbol{m}$ not only must obey Eq. (23) but also the condition

$$0 \le \boldsymbol{m} \le 1 \tag{24}$$

meaning that each component of $\boldsymbol{m}$ satisfies the constraints. Now we consider the feasibility of the linear program Eq. (23), Eq. (24) as the sampling in λ becomes ever more dense and the upper limit simultaneously approaches infinity. Every finite realization of this process is an elementary linear program. We assert that the feasibility of the program in the limit of dense sampling solves the existence problem. At first glance this is not obvious, because the condition Eq. (24) does not exactly correspond to $\mu(\lambda)$ taking values of only zero or one. However, an argument given by Parker (1994, Chapter 4) based on the Fundamental Theorem of linear programming, shows that no matter how large K becomes, no more than $N-1$ components of $\boldsymbol{m}$ can differ from the two extreme values in Eq. (24) in a so-called basic solution to the system (which must exist if any solution does). Therefore, as K grows, the basic solution vector $\boldsymbol{m}$ approximates ever more closely a function with only two values just as required. Furthermore, if a solution exists, there must be one with only finitely many transitions, because we are guaranteed (Parker, 1980) a model in D^+ also exists whenever there is an exact solution to a finite data set. We recall that models in D^+ have a finite number of transitions.

To apply these ideas to a particular data set, we must test both Eq. (18) and Eq. (21) because a surface insulating layer may be obligatory in the solution with one collection of observations, whereas with another, the model may require a surface conductor. The distinction between the two kinds of models is obviously artificial in the sense that a solution with a surface conductor whose conductivity is very small has a response almost identical to one with a surface insulator. Therefore, when a surface conductor is demanded, but the best-fitting solution is insulating, the admittance that develops in the limiting process has a pole a very long way out on the imaginary axis, with a very large residual, a_N. As the corresponding λ_N is much greater than any of the frequencies of observation, the contribution of this pole to the admittance, c, approximates that of the constant $a_0 = a_N/\lambda_N$; hence the constant has been effectively reintroduced into the response. Similarly, when $a_0 > 0$ is demanded in a solution, implying a surface insulator, the value of a_0 may tend to very small positive values in the limiting procedure, indicating that the other surface condition is more appropriate for the optimal solution.

4. Matching observations: noisy data

Naturally, exact matching of admittance values by a model is a purely theoretical exercise. With field measurements we must ask only that the agreement between model responses and observations be adequately close. We follow tradition by examining a weighted 2-norm as the measure of closeness. In place of Eq. (23) we consider an optimization problem:

$$\chi^2 = \min_{0 \le \boldsymbol{m} \le 1;\, \alpha \in \mathbb{R}} \left[\| \boldsymbol{S}_1^{-1} (\alpha \boldsymbol{1} + \boldsymbol{L}_1 \boldsymbol{m} - \boldsymbol{g}_1) \|^2 + \| \boldsymbol{S}_2^{-1} (\boldsymbol{L}_2 \boldsymbol{m} - \boldsymbol{g}_2) \|^2 \right] \quad (25)$$

where $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are diagonal matrices of the standard deviations of the measurements and the norm is the Euclidean length. As with the linear program for exact matching, the basic solution $\boldsymbol{m}$ of this quadratic program (an example of a Bounded-Variable, Least-Squares program, or BVLS; see Stark and Parker, 1995) can possess at most $N-1$ components that are not at one of the two limits, zero or one. The quantity minimized in Eq. (25) is a standardized squared discrepancy between model and measurement, and it is χ^2 distributed if the noise in the measurements is normally distributed. If the smallest possible misfit found from Eq. (25) would be attained or exceeded by chance in only 5% of random samples, we can reject the hypothesis that one-dimensional solutions exist at the 95% level.

The numerical techniques for solving the minimization posed in Eq. (25) are similar to those described by Parker and Whaler (1981). The imaginary axis is sampled initially in a geometric sequence, with largest and smallest values determined by the range of frequencies in the data set. Rather than using a numerical quadrature scheme to approximate the integral in Eq. (18), we can insist the function is constant on each of the open intervals between the sampling points and perform the integral exactly. Then we must specify that transitions can occur only just to the right of the sample points. After solving the optimization problem with the initial sampling, the axis is divided again, putting further potential transition points in the neighborhoods of those already found. The process may be repeated as often as necessary, but the minimum norm converges so rapidly that in practice one seldom needs more than a single stage of subdivision. A final refinement removes all the components in the solution vector that are not at one of the extreme values and substitutes a simple jump at a suitable intermediate value of λ chosen by linear interpolation. This last adjustment usually reduces the misfit by a small amount, less than 1%.

How much effect is the approximation by a normal error law in the noise likely to have? The random variable χ^2 is idealized as a sum of the squares of identically distributed random variables, Y_n, each obeying a normal distribution with standard error unity. Because N is usually large (greater than 20) the Central Limit Theorem (Rice, 1988) tells us that an asymptotic approximation χ^2 is itself Gaussian, whether or not the individual elements Y_n are. Because the variables in the sum have been standardized (normalized by the standard deviation), the mean of χ^2 is just N. A short calculation shows that the corresponding variance is $N(\mu_4 - 1)$, where μ_4 is the fourth central moment of the standardized distribution of Y_n; for Gaussian variables μ_4 is three. Thus the statement that χ^2 is more than 1.64 standard deviations above the mean (corresponding in the Gaussian approximation for χ^2 to 0.05 chance of occurring randomly) is equivalent to

$$\chi^2 \ge N + 1.64\sqrt{N(\mu_4 - 1)} \quad (26)$$

It is usually impossible to estimate μ_4 from the measurements with any precision. As we remarked, taking the variables Y_n to be Gaussian gives $\mu_4 = 3$.

We indulge in a simple model to give us some insight into the accuracies to be expected. Let us suppose that estimates of the real and imaginary parts of the admittance c at each frequency are perturbed by statistically independent, identically distributed Gaussian random components, with zero mean. Let the magnitude of the admittance at a certain frequency be $\bar{c}$, and the standard error in both real and imaginary part be σ_c; the ratio $\kappa = \sigma_c / \bar{c}$ is a measure of the accuracy of the admittance estimate. Modern equipment and processing methods can usually guarantee $\kappa < 0.1$; data with κ as large as 0.2 would be regarded as very noisy. Given this distribution in the admittance, we can compute the properties of the distribution of errors in $\ln \rho_a$, in

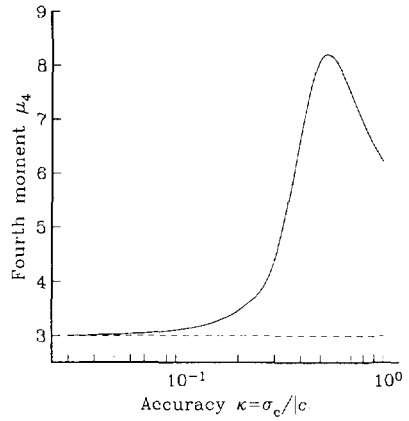

Fig. 3. Standardized fourth central moment μ_4 of the distribution of ρ_a when the admittance c is distributed as a bivariate uncorrelated Gaussian pdf in real and imaginary parts. The variance of the two Gaussian distributions is σ_c^2 and the mode of $|c|$ is $\bar{c}$. The ratio $\sigma_c/\bar{c}$ measures the relative accuracy of the admittance measurement.

particular the value of μ_4. The calculation of the fourth moment is a straightforward integration over the two-dimensional Gaussian probability density, which we performed numerically. We omit the details. The results, summarized in Fig. 3, show that unless the estimates are very noisy indeed, the Gaussian approximation for errors in $\ln \rho_a$ is satisfactory. For example, even if κ is a as large as 0.2, we find $\mu_4 = 3.466$ instead of three; then, when $N = 25$, Eq. (26) gives $\chi^2 > 37.88$ and the Gaussian approximation that $\mu_4 = 3$ yields $\chi^2 > 36.60$. Of course, in practice the estimators of c will not follow truly normal error distributions; the main concern arises from the presence of the singularity in the log function. Our simple illustration shows that, for reasonably accurate estimates of apparent resistivity, the singularity does little harm.

5. Bounding linear functionals

In addition to constructing the best-fitting model, the optimization technique can be employed to bound linear functionals of the representation function. In the case of D^+, Parker (1982) showed how this idea may be exploited to calculate the greatest depth from which information can be returned by an MT experiment. Here we examine the bounding of another set of linear functionals, the phases and apparent resistivities themselves.

Let us suppose only apparent resistivities are known; then we might ask: at a certain frequency ω, what is the largest and what is the smallest possible phase consistent with the apparent resistivity data? The value of the phase is just a linear functional of μ, namely Eq. (21), so these questions are answered by finding the maximum and minimum of the imaginary part of Eq. (21) at the frequency in question, subject to the constraint that the misfit is acceptable. Stated this way the problem is the optimization of a linear functional subject to a quadratic constraint. We can turn this into a sequence of more conventional quadratic programs as follows. We minimize the misfit taking a series of trial values for the unknown phase. The trial phase is swept through a range, at each value solving for the minimum misfit. The curve of minimum misfit vs. trial phase is convex and intersects the target misfit level twice, once at the lower bound on phase, once at the upper bound. To see why this is the case, we consider Fig. 4, which shows a typical curve: there are conductivity models corresponding to every point above the curve in the shaded zone, but none below because the curve gives the smallest possible misfit for that phase. Hence the segment of the horizontal line

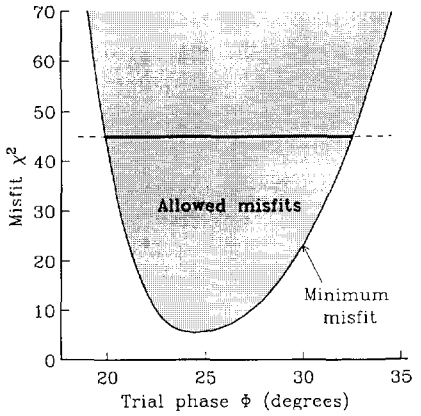

Fig. 4. The minimum misfit as a function of trial phase. The curve divides the plane into two regions: the shaded zone in which every point corresponds to a one-dimensional solution; the rest, where no solutions exist possessing the pair of values (Φ, χ^2). The horizontal line represents a target misfit level. The heavy segment covers the interval of allowed solutions and its endpoints are therefore the lower and upper bounds on permitted phase.

within the shaded zone delineates all permissible phases with that misfit, and its endpoints must be the bounding values. In exactly the same way we can find the upper and lower bounds at any frequency on either apparent resistivity or phase consistent with any finite collection of measurements. In practice, the requirement of an exact fit by the trial value can be replaced by the use of a very small standard error. These ideas were first used in a seismic problem by Stark and Parker (1987). The program BVLS mentioned above has been designed to be efficient for solving a series of closely related optimization problems (such as the set arising from a sweep through trial phases), because it can take advantage of a good starting approximation, which in this case can be based on the previous member in the sequence. We have been able to exploit this feature in our computer code for bounds on missing or highly uncertain data.

6. Examples

Our first example uses the COPROD data, an MT site from Scotland (Jones and Hutton, 1979), that has been widely distributed for the purpose of comparing one-dimensional inversions (e.g. Constable et al., 1987). The data are plotted in Fig. 5. These response estimates are typical of MT data before the advent of robust digital time series methods (see Jones et al., 1989) in that most of the phase errors are larger than expected for the stated uncertainties in apparent resistivity. For instance, the smallest uncertainty in ρ_a is 3.8% at 320 s. Assuming that the error ellipse for complex impedance is a circle, the uncertainty $\delta\Phi$ (radians) corresponding to an uncertainty in ρ_a is

$$\delta\Phi = \tan^{-1}\left(0.5\frac{\delta\rho_a}{\rho_a}\right) \approx 0.5\frac{\delta\rho_a}{\rho_a} \qquad (27)$$

for small errors. Thus the uncertainty in Φ should be 1.1°. The actual claimed uncertainty is 6°. At 1450 s, $\delta\Phi$ is actually smaller than predicted from $\delta\rho_a$, whereas at seven periods, the stated phase errors are so large (23°) that they appear to place essentially no constraint on Earth structure.

Fig. 5(a) shows the 95% confidence bounds on Φ, when the data are the phase, excluding the seven phases with very large uncertainties. The response of the D^+ model is also shown. The D^+ minimum χ^2 is 4.44, substantially less than the just-acceptable value of 16.9. There are two immediate conclusions. First, the included phase data are, by themselves, consistent with a one-dimensional model. This is not particularly useful, because one would be unlikely to base an interpretation of this MT site on such limited phase data. Second, the phase datum at 585 s is inconsistent with the other included phase data (adding the excluded Φ data does not alter this conclusion). Its most probable value lies well above the upper bound and its error bar barely overlaps the upper bound. A surprise is that all of the excluded phase data lie very close to the response of the D^+ model and well within their bounds. Furthermore, the supposedly one standard error bars of the excluded data extend well outside their 95% bounds. We can conclude, with a high degree of certainty, that the excluded phases are completely consistent with the included phases and that the uncertainties for the excluded phases have been overestimated.

Fig. 5(b) shows the 95% confidence bounds on Φ when only the ρ_a data are used. To reduce complexity, the bound symbols in this figure (and all subsequent figures) have been replaced by dotted lines connecting the bounds, although the bounds have still been calculated only at the measurement periods. The D^+ χ^2 is 5.2 compared with the just-acceptable χ^2 of 25. Thus the ρ_a data, by themselves, are also consistent with one-dimensionality. Phase bounds from ρ_a are larger than from Φ and consequently the Φ datum at 585 s is consistent with the ρ_a. It should be noted, however, that the Φ values suggest a phase maximum at about 60 s, whereas the ρ_a values predict monotonically rising phase at short period. This inconsistency between the resistivity and phase causes Φ at 96 s to lie outside its bound. It is again obvious that the seven Φ, excluded in the first exercise, have error bars that have been overestimated.

To test the hypothesis that the Φ errors can be replaced with those predicted from the ρ_a, we first calculated the bounds on Φ with the modified errors. These data pass the one-dimensionality test although both the datum at 96 and 585 s are clearly above their 95% confidence upper bounds. We next performed a joint inversion of ρ_a and Φ and calculated the 95% confidence bounds shown in Fig. 5(c) and

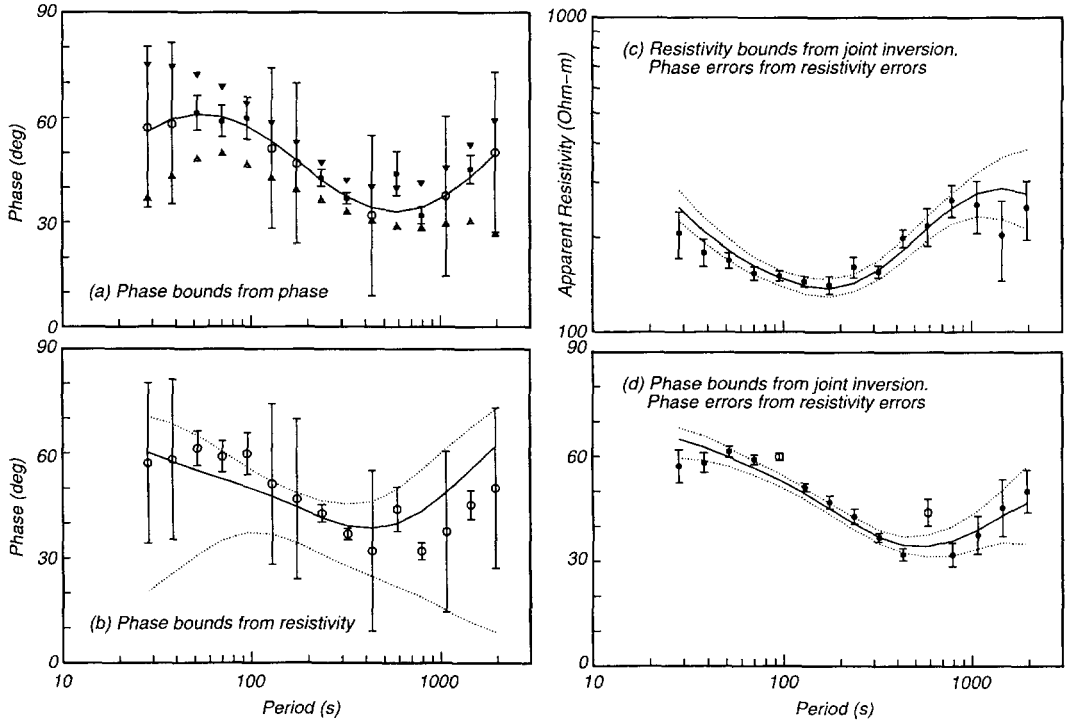

Fig. 5. Data, bounds and D^+ responses for the COPROD data. (a) Bounds on Φ computed from inversion of Φ. (b) Bounds on Φ computed from ρ_a. (c) and (d) Bounds on ρ_a and Φ computed from joint inversion of ρ_a and Φ. In (a), ▲, lower bounds; △, upper bounds. In (b)–(d) dotted lines connect the bounds that are still computed only at the measurement periods. In each case, continuous lines connect the predicted D^+ responses at the measurement periods; ●, data included in the calculation; ○, data excluded. The error bars in (a), (b) and (c) are the claimed errors, whereas the phase errors in (d) are those predicted from the ρ_a.

Fig. 5(d). The two suspected phase outliers just identified were excluded from this computation because this significantly decreases the minimum χ^2. Each bound is the limiting value of the datum at which χ^2 equals its just-acceptable value, assuming that all the rest of the data and uncertainties are unchanged. Reducing the minimum χ^2 increases the range between the upper and lower bounds at an excluded outlier and thus provides a stronger test of its incompatibility. The minimum χ^2 is 18.7, whereas the just-acceptable value is 26.3. Thus reducing the error estimates for phase does not cause the data to fail the test for one-dimensionality. The two phase outliers remain clearly outside their bounds. However, six ρ_a and six additional Φ also lie very close to, or outside their bounds, indicating that the reduced errors are somewhat too optimistic. Furthermore, the misfit to the D^+ responses, and the relation of the data to their bounds, both appear non-random. These are manifestations of the incompatibility of phase and apparent resistivity already noted, which has been exacerbated by the reduced errors.

The second example uses data from a continuous electromagnetic profile in the Arbuckle Mts. region of south–central Oklahoma. The profile involves 93 nominally 300 m dipoles laid end to end with the magnetic field measured in the perpendicular direction. We wanted to test whether these data are compatible with their interpretation as B-polarization induction.

Our first step was to use two nearby tensor MT sites to determine the regional electric strike. We used the impedance tensor decomposition of Smith (1995) and found that the strike deviated about 20° from the normal to the profile. Unfortunately, there are no cross-dipoles for this profile, so the data cannot be rotated into the coordinate perpendicular

to strike and thus are probably not quite B-polarization.

The next step was to apply the consistency test of Weidelt and Kaikkonen for B-polarization (i.e. the test for compatibility with D^+). We chose a single span (number 60 counting from the southwest end) that had small claimed error bars and exhibited a significant response to the main lateral structure within the profile. The data are shown in Fig. 6. Although the original data consist of 40 estimates from 2.60×10^{-3} to 1820 s, the estimates at periods greater than 114 s or less than 0.02 s are poorly constrained at most spans and we have not used them. The minimum χ^2 for a joint inversion of ρ_a and Φ is 96.5, which exceeds the 95% confidence limit of 69.8. Consequently, these data fail the B-polarization test.

To see whether this is a problem with the data estimates or inadequacy of the B-polarization assumption, we performed the bounding exercises shown in Fig. 6. Fig. 6(a) and Fig. 6(b) are the 95% confidence bounds on ρ_a and Φ when the data are the ρ_a, excluding a suspected outlier at 0.222 s. The minimum χ^2 is 2.15, well below the 95% confidence value of 37.7. Thus the resistivity data, by themselves, easily pass the B-polarization consistency test. Fig. 6(c) and Fig. 6(d) reverse the bounding exercise, with 95% confidence bounds on ρ_a and Φ when the data are the Φ, excluding a suspected outlier at 0.222 s. The minimum χ^2 is 5.5, which is again well below its 95% confidence value. Thus the phase data also easily pass the B-polarization test. The failure of ρ_a and Φ to jointly pass the test must therefore be due to some form of inconsistency between them. In Fig. 6(b), we see that several of the Φ values between 1 and 0.1 s exceed their upper bounds. Thus one possible source of the incompatibility is a systematic problem with estimation of short period phase. However, Fig. 6(c) suggests a much simpler explanation. There is a tear in the ρ_a estimates which coincides with the boundary between measurement bands. The high-frequency ρ_a,

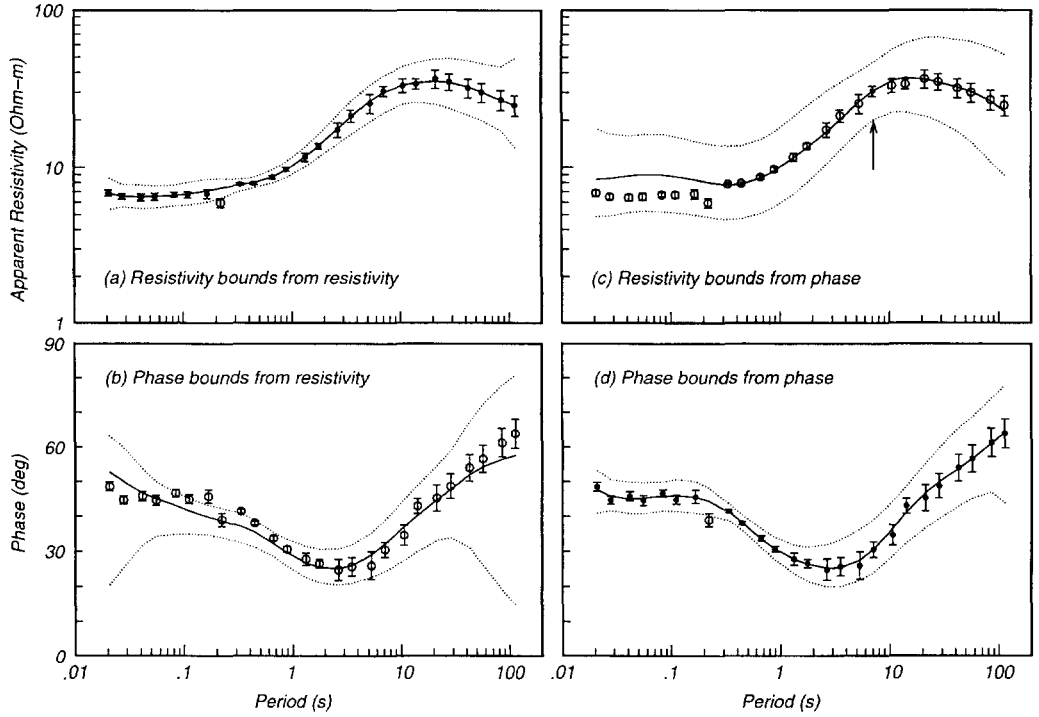

Fig. 6. Bounds on ρ_a and Φ for an electromagnetic profile span. The dotted lines connect the bounds at the measurement periods; the continuous line connects the D^+ responses at the measurement periods. ●, Data included in each calculation; ○, data excluded. (a) and (b) invert ρ_a; (c) and (d) invert Φ with the single ρ_a indicated by the arrow.

although individually compatible with the Φ at the 95% confidence level, have an almost constant downward bias from the D^+ responses predicted by the phase data. This suggests either a consistent statistical bias owing to noise, or an instrument gain calibration problem. As these data have been processed using a remote magnetic reference (Gamble et al., 1979) located 7 km from the primary magnetic sensors, a bias explanation seems unlikely. The median ratio of the eight high-frequency ρ_a to the predicted D^+ responses is 0.75. Using this factor to adjust the high-frequency ρ_a leads to a minimum χ^2 of only 19.9 for joint inversion of all the ρ_a and Φ. As the just-acceptable χ^2 at the 95% level is 69.8, the modified data easily pass the B-polarization test and the only datum not consistent at the 95% confidence level is the Φ at 0.222 s. The circumstantial evidence for a gain calibration problem or a consistent statistical bias is therefore fairly strong. One could try to distinguish between calibration and bias by examining data from other spans in the same way. One cannot, however, unambiguously assign a gain problem to the high-frequency data. A shift of the long period data in the opposite direction would give the same improvement in data misfit.

The final example is a tensor MT site from the Northern California Coast Ranges. After impedance tensor decomposition to determine regional electrical strike, the data were rotated to strike and assigned to the B-polarization and E-polarization (electric field parallel to strike) based on the direction of the real part of the horizontal to vertical magnetic field transfer function (variously called the induction vector or the tipper). The chosen direction for the B-polarization is also consistent with what one would expect from the regional geology.

The time series were processed using a robust technique closely related to 'method 7' outlined by Jones et al., 1989, which is based, in turn, on the work of Egbert and Booker (1986). The modifica-

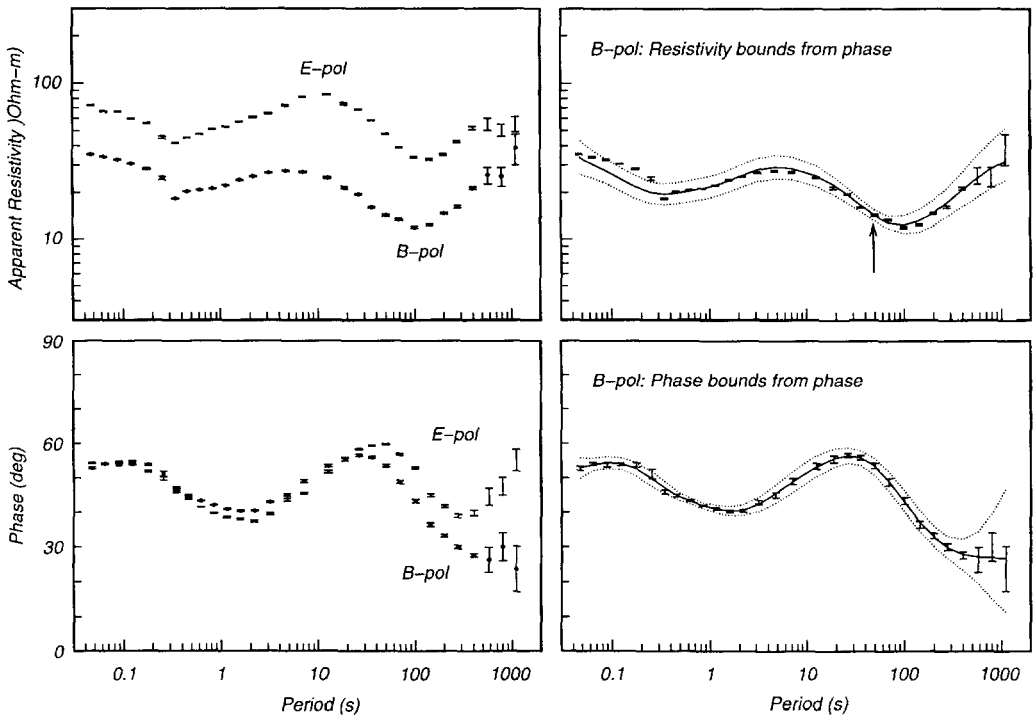

Fig. 7. On the left: data and claimed errors from an MT site in the Northern California Coast Ranges. The data have been rotated 10° clockwise from their geomagnetic measurement coordinates. B-Polarization data have east–west electric field and their most probable value is indicated (•). On the right: bounds on ρ_a and Φ calculated from the B-polarization Φ and the single ρ_a indicated by the arrow. The dotted lines connect the bounds at the measurement periods; the continuous line connects the D^- responses.

tions are in the so-called dead band around 1 s, where the weakness of the signal can lead a simple robust method to reject the desired data in favor of noise. The response estimates are plotted in Fig. 7. The median claimed error in ρ_a is 1% (0.3° in Φ); only the three long period data exceed 2%. The ρ_a data for the two polarizations are offset, but have nearly identical period dependence, and the Φ data are almost identical at periods below 10 s. This suggests that the shallow structure is nearly one-dimensional below a very shallow, non-inductive structure that distorts the electric field (see Smith (1995) and references contained therein). At the longer periods, the period dependence of the ρ_a is no longer parallel and the Φ data split, indicating multi-dimensional regional structure.

Non-inductive surface distortion mixes the strike-aligned electric fields of the regional structure except when the impedance is in the regional strike coordinates (see Smith, 1995). Even if the responses of both the B- and E-polarizations pass a one-dimensionality test, there is no guarantee that the response of a distorted field in non-strike coordinates will also pass. In fact, distorted phases have been observed to lie outside the range of 0–90° required by any one-dimensional response. Thus application of the B-polarization consistency test provides a constraint on the legitimacy of the choice of strike as well as the assumption of regional two-dimensionality. This is in addition to the consistency test for the assumption of shallow, non-inductive distortion of two-dimensional regional responses provided by the impedance tensor decomposition used to determine the strike.

Joint inversion of the B-polarization ρ_a and Φ with the claimed errors gives a minimum χ^2 of 1097, far in excess of the just-acceptable value. Switching the polarizations gives a minimum χ^2 of 527. Thus neither polarization passes the consistency test for B-polarization. In fact, neither ρ_a nor Φ

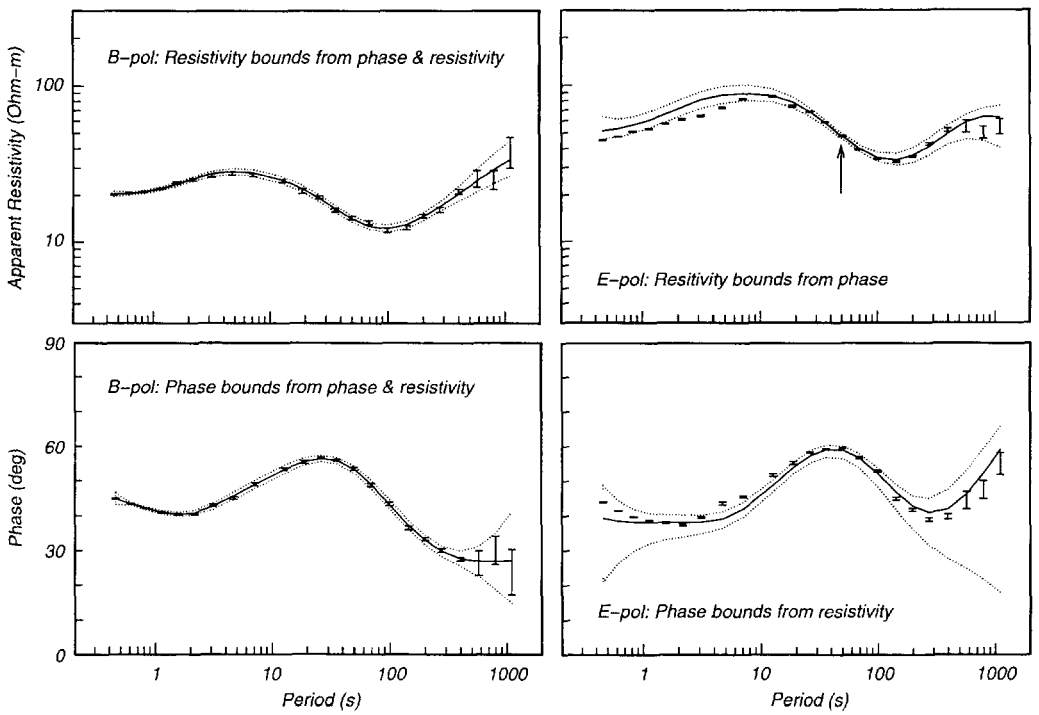

Fig. 8. Bounds on ρ_a and Φ for the California MT site. The dotted lines connect the bounds at the measurement periods; the continuous line connects the D^+ responses. Only the 23 longest periods are used (see text). On the left: inversion of B-polarization ρ_a together with Φ. On the right: the upper panel shows bounds for ρ_a computed from inversion of the E-polarization Φ plus the single ρ_a indicated by the arrow. The lower panel shows bounds for Φ computed from inversion of the E-polarization ρ_a.

alone for either polarization passes the test at the 95% level.

To see what was going on, we doubled the claimed errors for the B-polarization Φ (except at the three longest periods). This reduces the minimum χ^2 to 14.3 and we calculated the Φ and ρ_a bounds shown in Fig. 7. It is immediately obvious that there may be a problem with the ρ_a. Statistical downward bias of the ρ_a in the so-called dead band around 1 s could be a problem, because the magnetic remote reference failed at this site. However, bias primarily affects magnitude and not phase of a transfer function. Thus the high consistency of the ρ_a and Φ in this period range argues strongly against bias. Because the large deviation of the ρ_a from the D^+ response predicted by Φ varies smoothly with frequency, we suggest that the most likely cause is an incorrectly applied filter correction. We have not pursued this further and will not use the seven highest-frequency data in the rest of the discussion.

Once the seven high-frequency data are deleted, the minimum χ^2 for Φ drops to 10.2, well below the 95% value of 36.4. Thus the phase data clearly pass the B-polarization test. Unfortunately, both the ρ_a data alone and joint inversion of ρ_a and Φ fail (but not by much) at the 95% confidence level. We have therefore arbitrarily doubled the claimed errors (except at the three longest periods) and computed the bounds from joint inversion of ρ_a and Φ. The results are shown on the left in Fig. 8. At this error level (which is still less than 2% (0.6° in Φ) for most of the data), the full data set passes the B-polarization test (the minimum χ^2 is 36.8; the just-acceptable value is 62.8). Furthermore, all the data lie within their 95% bounds.

It is not known under what conditions an E-polarization response should pass the one-dimensional response test. However, it is certainly worth trying to see if this polarization can be B-polarization, because tensor composition is ambiguous in this regard.

The E-polarization data (minus the seven shortest periods) do not pass the one-dimensional response test either jointly or individually. One must triple the claimed errors (except at the longest three periods) to obtain the χ^2 sufficiently low that the 95% bounds encompass most of the data. Even then, the three Φ centered on 10 s lie at or above their bounds. Excluding them from the calculation permits all the remaining ρ_a and Φ to lie within their 95% bounds with errors reset to twice their claimed errors. Both ρ_a and Φ individually pass the one-dimensionality test with doubled errors (except at the three longest periods). The 95% bounds for ρ_a predicted from Φ and for Φ predicted from ρ_a are shown on the right of Fig. 8. The incompatibility between the two responses is clear. Either the three Φ centered on 10 s already noted are too high by about 3° or there is an approximately 10% downward bias in the ρ_a with periods shorter than 10 s. Downward bias off the ρ_a has to be a serious contender because of the absence of a remote reference at this site.

Although these MT data do not pass the B-polarization test for either polarization, their claimed errors are so small that they still would be regarded as very good data, even after their error bars have been sufficiently increased to pass the test. Thus, other than possessing error estimates that are somewhat too optimistic, there appears to be no reason to think that the 'B-polarization' data are not compatible with two-dimensional B-polarization interpretation. With more qualifications, the 'E-polarization' data also do not appear to show deviations from a one-dimensional response that cannot reasonably be attributed to statistical bias or error estimation that is too optimistic.

7. Conclusions

We have shown how to recast the problem of finding the best fitting (D^+) model from real and imaginary admittance to apparent resistivity and phase. This inverse problem leads naturally to bounds on resistivity and phase for a given set of MT data. Because two-dimensional B-polarization responses must also have a $D+$ response, the inverse problem and related bounds can be used to test the validity of both one- and two-dimensional modeling assumptions and to detect data collection problems.

Acknowledgements

The COPROD data were provided by Alan Jones; the electromagnetic profile data were collected by an industrial contractor for EXXON and provided by

Leonard Srnka; the California MT data were collected by Harve Waff and Dean Livelybrooks and processed by Gary Egbert. This research was supported by the Office of Basic Energy Sciences, Department of Energy Grant DE-FG06-92ER14231, with additional support by the National Science Foundation Grant EAR87-08365.

References

Constable, S.C., Parker, R.L. and Constable, C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from EM sounding data. Geophysics, 52: 289–300.

Egbert, G.D. and Booker, J.R., 1986. Robust estimation of geomagnetic transfer functions. Geophys. J. R. Astron. Soc., 87: 173–194.

Egbert, G.D. and Booker, J.R., 1992. Very long period magnetotellurics at Tucson Observatory: implications for mantle conductivity. J. Geophys. Res., 97: 15099–15112.

Gamble, T.D., Goubau, W.M. and Clarke, J., 1979. Magnetotellurics with a remote reference. Geophysics, 44: 53–68.

Jones, A.G., 1988. Static shift of magnetotelluric data and its removal in a sedimentary basin environment. Geophysics, 53: 967–978.

Jones, A.G. and Hutton, R., 1979. A multi-station magnetotelluric study in southern Scotland—I. Fieldwork, data analysis and results. Geophys. J. R. Astron. Soc., 56: 329–349.

Jones, A.G., Chave, A.D., Egbert, G., Auld, D. and Bahr, K., 1989. A comparison of techniques for magnetotelluric response function estimates. J. Geophys. Res., 94: 14201–14213.

Korevaar, J., 1968. Mathematical Methods. Academic Press, New York.

Parker, R.L., 1980. The inverse problem of electromagnetic induction: existence and construction of solutions based upon incomplete data. J. Geophys. Res., 85: 4421–4428.

Parker, R.L., 1982. The existence of a region inaccessible to magnetotelluric sounding. Geophys. J. R. Astron. Soc., 68: 165–170.

Parker, R.L., 1994. Geophysical Inverse Theory. Princeton University Press, Princeton, NJ.

Parker, R.L. and Whaler, K., 1981. Numerical methods for establishing solutions to the inverse problem of electromagnetic induction. J. Geophys. Res., 86: 9574–9584.

Rice, J.A., 1988. Mathematical Statistics and Data Analysis. Brooks–Cole, Monterey, CA.

Smith, J.T., 1995. Understanding telluric distortion matrices. Geophys. J. Int., 118: 219–226.

Stark, P.B. and Parker, R.L., 1987. Velocity bounds from statistical estimates of $\tau(p)$ and $X(p)$. J. Geophys. Res., 92: 2713–2719.

Stark, P.B. and Parker, R.L., 1995. Bounded-variable least-squares: an algorithm and applications. Comp. Stat., 10: 129–141.

Torres-Verdin, C. and Bostick, Jr., F.X., 1992a. Implications of the Born approximation for the magnetotelluric problem in three-dimensional environments. Geophysics, 97: 587–602.

Torres-Verdin, C. and Bostick, Jr., F.X., 1992b. Principles of spatial surface electric field filtering in magnetotellurics: Electromagnetic Array Profiling (EMAP). Geophysics, 97: 603–622.

Weidelt, P., 1972. The inverse problem of geomagnetic induction. Z. Geophys., 38: 257–289.

Weidelt, P. and Kaikkonen, P., 1994. Local 1-D interpretation of magnetotelluric B-polarization impedances. Geophys. J. Int., 117: 733–748.