Table 1.

	$\tilde{Z}$	$A = H_x^I$ $B = H_y^I$		Unit	
		Mean val.	σ	Mean val.	σ
$\lvert Z_{xx}\rvert$	$\sqrt{2}/2$	0.70	0.11	0.83	0.10
Φ_{xx}	$\pi/4$	0.78	0.15	0.77	0.17
$\lvert Z_{xy}\rvert$	0.5	0.88	0.10	0.55	0.08
Φ_{xy}	$\pi/2$	1.56	0.25	1.60	0.26
$\lvert Z_{yx}\rvert$	0.5	0.54	0.15	0.66	0.13
Φ_{yx}	$-\pi/2$	−1.51	0.28	−1.53	0.26
$\lvert Z_{yy}\rvert$	$\sqrt{2}/2$	0.71	0.14	0.78	0.09
Φ_{yy}	$3\pi/4$	2.23	0.19	2.31	0.14

Every component was perturbed choosing random numbers uniformly distributed over the range $(-D, +D)$.

$$F(k) = F^I(k) + F^N(k) \quad k = 1, \ldots, 50, \tag{5}$$

where the range was determined in such a way that the expectancy of the ratios:

$$(E_x^N, E_x^N)/(E_x^I, E_x^I) = (E_y^N, E_y^N)/(E_y^I, E_y^I) = \alpha, \tag{6a}$$

and

$$(H_x^N, H_x^N)/(H_x^I, H_x^I) = (H_y^N, H_y^N)/(H_y^I, H_y^I) = \beta. \tag{6b}$$

For α and β fixed, the calculations were repeated 50 times using independent sets of random numbers to generate the noise fields F^N. Mean values $\bar{Z}_{ij}$ and standard deviations σ_{ij} were computed. Results for $\alpha = 1$, $\beta = 0.5$ are presented in Table 1 where the proposed algorithm (called Unit) is compared to the "clean" reference $A = H_x^I$, $B = H_y^I$. The phases are given in radians.

Magnetotelluric Transfer Function Estimation Improvement by a Coherence-Based Rejection Technique EM1.5

Alan G. Jones, Energy, Mines and Resources Canada; and Hartmut Jödicke, Westfalen Wilhelms Univ., West Germany

Reliable estimation of the transfer function(s) of a linear system is of paramount importance in many fields of research, notably geophysics. The estimates are, however, biased by noise contributions on the various components of the system. Several techniques have been presented in the last 10 years for deriving more reliable and less biased estimates, particularly of the magnetotelluric impedance tensor elements. This paper addresses the question of whether it is useful, in certain circumstances, to reject some of the available realizations (or "raw estimates") from the smoothing procedure employed to determine the transfer function estimates. A rejection technique is proposed, based on the inspection of coherence functions as individual realizations are omitted, whereby more stable estimates of the transfer functions may be derived which exhibit less bias due to noise contributions.

The technique is illustrated by applying it to synthetic data and to real data (magnetotelluric impedance tensor element estimation). It is shown that superior estimates of the transfer functions result after applying the rejection technique—outliers are quickly located and rejected, random errors are reduced, bias errors are reduced, coherences are increased, and the estimates are more stable.

Introduction

It is relatively well known that the usual algorithms for estimating response functions give estimates that are biased by noise contributions occurring on the input and/or output realizations of the linear system under consideration. Recently, several techniques were proposed in the field of magnetotelluric (MT) data analysis to alleviate or circumvent the problems caused by noise-biased autospectral density estimates. The most powerful method to date appears to be the remote reference technique of Gamble et al. (1978), which utilizes a suggestion proposed by Akaike (1967) of employing time functions, so-called "instrumental variables" by Reiersøl (1950), that are correlated with the true input(s) and output(s) of the linear system but that are orthogonal to the noise components thereon. The "cross-frequency" analysis method of Dekker and Hastie (1981) appears promising, as do the cyclic signal-to-noise ratio enhancement techniques of Kao and Rankin (1977) and Lienert et al. (1980). Other methods of bias removal were discussed in Goubau et al. (1978).

However, all of the methods cited above are concerned with either *which* signals (or realizations) to employ to derive the estimates of the auto- and cross-spectral densities, or *how to improve* those estimates once they have been made. This paper discusses a procedure that, by its very nature, addresses itself to a more fundamental point than those proposed above, namely *how to construct* the auto- and cross-spectral densities. We propose a rejection technique, based on coherence, which results in transfer function estimates displaying less bias, smaller confidence intervals, and a higher overall coherence than the more usual analysis methods give. The technique complements those of Gamble et al., Dekker and Hastie, Kao and Rankin, Lienert et al., and Goubau et al. in that its employment does not prohibit utilizing any of the other methods—it can be used as one step in two-step processing of data.

Procedure

In this section, the procedure for the proposed transfer function estimation improvement technique is described with references to a single input/single output linear system. Extension to multiinput/multioutput systems is both simple and obvious.

Smoothed cross-spectral density estimates $\bar{C}_{xy}$ (dependence on frequency assumed) are constructed from averages of the raw estimates, for example,

$$\hat{C}_{xy} = \hat{x}^* \cdot \hat{Y},$$

and

$$\bar{C}_{xy} = \langle \hat{C}_{xy} \rangle. \tag{1}$$

(The averaging algorithm employed is not of importance for this work.)

From the averaged auto- and cross-spectra, an estimate $\hat{S}_d$ of the true response function S relating input $x(t)$ to output $y(t)$ is given by

$$\hat{S}_d = \frac{\bar{C}_{xy}}{\bar{C}_{yy}} \tag{2a}$$

and is well known to be *downward-biased* by noise component $n_x(t)$ on $x(t)$ as

$$E[\hat{S}_d] = \frac{S}{1 + E[R_x]} \tag{2b}$$

where E [] denotes expectation value, and R_x is the input noise power to signal power ratio. Alternatively, it is possible to estimate S by estimator $\hat{S}_u$ given by

$$\hat{S}_u = \frac{\bar{C}_{yy}}{\bar{C}_{yx}}, \tag{3a}$$

which can similarly be shown to be *upward-biased* by noise component $n_y(t)$ on $y(t)$ as

$$E[\hat{S}_u] = S(1 + E[R_y]) \tag{3b}$$

where R_y is the output noise power-to-signal power ratio. If $E[R_x] = E[R_y]$, then the estimator

$$\hat{S} = \sqrt{\hat{S}_u \hat{S}_d}, \tag{4}$$

i.e., the geometric mean of equations (3a) and (4a), is an unbiased estimate of S. Obviously, without any a priori information as to which noise-signal ratio is largest (i.e., R_x or R_y), the true value will lie in the range $[\hat{S}_d, \hat{S}_u]$, plus-minus the associated random error. The random error $\hat{r}^2$ may be estimated by any of the usual methods.

For n independent raw estimates of $\hat{C}_{xy}$, i.e., from n $(\hat{X}_i, \hat{Y}_i)$ pairs, the estimate of the ordinary coherence function between the input and the output is defined as

$$\hat{\gamma}^2_{xy} = \frac{|\bar{C}_{xy}|^2}{\bar{C}_{xx} \cdot \bar{C}_{yy}} = \frac{|\langle \hat{C}_{xy} \rangle|^2}{\langle \hat{C}_{xx} \rangle \langle \hat{C}_{yy} \rangle}. \tag{5}$$

Omitting one of the data pairs, say the jth $(\hat{X}_j, \hat{Y}_j)$, leads to a new estimate of the coherence derived from the $(n - 1)$ remaining $(\hat{X}_i, \hat{Y}_i)_{i \neq j}$ pairs, i.e.,

$$(\hat{\gamma}^2_{xy})_{i \neq j} = \left(\frac{|\bar{C}_{xy}|^2}{\bar{C}_{xx}\ \bar{C}_{yy}} \right)_{i \neq j}.$$

If $(\hat{\Upsilon}^2_{xy})_{i \neq j} > \hat{\Upsilon}^2_{xy}$, then the downward and upward biased estimates of the true transfer response function, derived from the reduced set of $(n - 1)$ pairs, can be considered to be superior to those estimates from the total data set of n pairs. Hence, deriving all the n coherence estimates $(\hat{\Upsilon}^2_{xy})_{i \neq j, 1, n}$, and ascertaining which value of j gives a *maximum* in $(\Upsilon^2_{xy})_{i \neq j}$, say the k_1th value, identifies the pair of raw estimates $(\hat{X}_{k_1}, \hat{Y}_{k_1}$ which causes the greatest decrease in total coherence. If $(\hat{\Upsilon}^2_{xy})_{i \neq k_1} > \hat{\Upsilon}^2_{xy}$, then $(\hat{S}_d)_{i \neq k_1}$ and $(\hat{S}_u)_{i \neq k_1}$ are superior estimates of S than $\hat{S}_d$ and $\hat{S}_u$, and vice versa.

For $(\hat{\Upsilon}^2_{xy})_{i \neq k_1} > \hat{\Upsilon}^2_{xy}$, the procedure is repeated for the remaining $(n - 1)$ estimates to detect the next pair $(\hat{X}_{k_2}, \hat{Y}_{k_2})$ whose omission causes the greatest increase in coherence. As before, if $(\hat{\Upsilon}^2_{xy})_{i \neq k_1, k_2} > (\hat{\Upsilon}^2_{xy})_{i \neq k_1}$, then $(\hat{S}_d)_{i \neq k_1, k_2}$ and $(\hat{S}_u)_{i \neq k_1, k_2}$ are superior to $(\hat{S}_d)_{i \neq k_1}$ and $(\hat{S}_u)_{i \neq k_1}$.

The cycle is repeated until a sufficient number of pairs, say m, $(\hat{X}_{k_1 \ldots k_m}, \hat{Y}_{k_1 \ldots k_m})$ have been rejected. The point at which to terminate the cycle, i.e., the definition of a "sufficient number," may also be by either subjective or objective criteria. A subjective criterion is, for example, to remove the worst 10 percent (or 20 or 30 percent, etc.) of estimates. An objective criterion would involve, for example, performing the rejection cycle until the confidence interval is minimized. Alternatively, the cycle could be terminated when the coherence function, estimated after rejecting m pairs, i.e., from the remaining $(n - m)$ pairs is observed to maximize, i.e.,

$$(\hat{\Upsilon}^2_{xy})_{i \neq k_1 \ldots k_{m-1}} < (\hat{\Upsilon}^2_{xy})_{i \neq k_1 \ldots k_m} > (\hat{\Upsilon}^2_{xy})_{i \neq k_1 \ldots k_{m+1}}.$$

However, this criterion should not be adopted because, in practice, the estimated coherence function does not maximize since data pairs are rejected due to the inherent bias associated with the estimate.

In this work, a form of the latter objective criterion is adopted because it leads to estimates of the true response function which display the highest SNRs, as inferred by the associated maximum in the estimate of the ordinary coherence. However, the estimate of the ordinary coherence function, as defined by equation (5) is known to be a biased function.

A more useful description of the coherence between two processes is offered by the estimate of the normalized transformed ordinary coherence function of Jones (1981). Between totally uncorrelated processes $E[\hat{N}_u] = 1$, hence the value of $\hat{N}_{xy}$ indicates directly the coherent-to-incoherent common signal ratio. The other properties of the NTOC function which make it preferable to ordinary coherence function, are detailed in Jones (1981). Utilizing the above described NTOC function, the *objective* criterion chosen for terminating the rejection cycle was after m pairs $(\hat{X}_{k_1 \ldots k_m}, \hat{Y}_{k_1 \ldots k_m})$ were identified and rejected such that the NTOC estimate maximized.

Application

Synthetic data. In order to illustrate the rejection procedure, synthetic data were generated consisting of true signals plus randomly added noise contributions of various degree and intensity. The true signals were $X = (3, 2)$ and $Y = (4, 7)$, i.e., a sinusoidal input of $\sqrt{13}$ units amplitude with an initial phase of 33.7 degrees, and an output of $\sqrt{65}$ units amplitude with an initial phase of 60.3 degrees. Hence the linear system was a ×2.24 amplifier with a 26.6 degree phase shift, i.e., $S = (2, 1)$. The synthetic data were generated by adding random numbers from a Gaussian distribution to a certain percentage of 100 (X, Y) true values given above. The noise was added randomly to the four traces, i.e., Re(X), Im(X), Re(Y), and Im(X), independently of each other. Hence, the data represent estimates, at a single frequency, of X and Y from 100 independent realizations of the process. That the noise is added to the four data series in a totally independent manner may be considered a worst possible case since a highly noise-contaminated realization will result in a noise degraded pair, either [Re(X_i), Im(X_i)] or [Re(Y_i), Im(Y_i)].

For the example to be considered in depth, 50 percent of the data were perturbed by noise contributions of standard deviation 2, e.g., 68 percent of the "noisy" values of trace Re(X) lay in the range 3 ± 2. The input signal power was 13 units, the output signal power was 65 units, and the noise power on both components was 4 units, i.e., $(2^2 + 2^2)/2$, the factor of 2 arising from only 50 percent of the data containing noise contributions. Hence, the input SNR was 13/4 = 3.25, and the output SNR was 65/4 = 16.25, and from equations (2b) and (3b),

$$E[\hat{S}_d] = \frac{(2, 1)}{\left(1 + \frac{1}{3.25}\right)} = (1.53, 0.76)$$

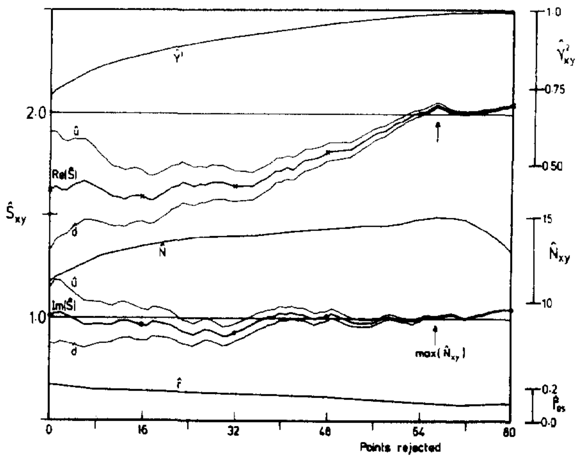

FIG. 1. Example of the application of the rejection technique applied to synthetic data. The true transfer function is given by $S = (2, 1)$. The variations of Re($\hat{S}$) and Im($\hat{S}$) are illustrated as noisy pairs (X_i, Y_i) and rejected (full lines with crosses every 16 points). They are bounded by their respective upward $\hat{S}_u$ and downward $\hat{S}_d$ biased estimates [equations (2a) and (3a), respectively]. Variations of the estimates of the ordinary coherence $\hat{\gamma}^2_{xy}$, the normalized transformed ordinary coherence $\hat{N}_{xy}$, and the radius of the circle of the confidence at the 95 percent confidence level $\hat{r}$ as complex pairs are rejected are also shown.

and

$$E[\hat{S}_u] = (2, 1)\left(1 + \frac{1}{16.25}\right) = (2.12, 1.06).$$

Also, from equation (5),

$$E[\hat{\gamma}^2_{xy}] = \frac{1}{(1 + E[R_x])(1 + E[R_y])} = 0.72,$$

and, the expectation value of the random error at the 95 percent confidence level is

$$E[\hat{r}^2_{0.95}] = \frac{2}{198}\, 3.04\,(1 - 0.72)\left(\frac{65 + 4}{13 + 4}\right) = 0.035.$$

It is instructive to examine the variation of the parameters as noisy data pairs are rejected for one particular run. Deriving the smoothed cross-spectral densities and then deriving $\hat{S}_d$, $\hat{S}_u$, $\hat{S}$, $\hat{r}^2_{0.95}$, $\hat{\gamma}^2_{xy}$, and $\hat{N}_{xy}$ for the total 100 pair data set gave

$$\left.\begin{aligned}\hat{S}_d &= (1.33,\ .88),\\ \hat{S}_u &= (1.88,\ 1.55),\end{aligned}\right\}\ \hat{S} = (1.58,\ 1.01)$$
$$\hat{r} = 0.210,$$
$$\hat{\Upsilon}^2 = 0.73,$$

and

$$N = 12.08.$$

The 100 data pairs (X_i, Y_i) under consideration here were processed by the cyclic rejection procedure. The variation of $\hat{S}_d$, $\hat{S}_u$, $\hat{S}$, $\hat{r}$, $\hat{\Upsilon}^2_{xy}$, and $\hat{N}_{xy}$ as data pairs were rejected, up to a total of 80 pairs, is illustrated in Figure 1. As shown in the figure, $\hat{N}_{xy}$ maximized after 67 of the pairs were rejected, at which point the estimates of the various parameters were

$$\hat{S}_d = (2.04,\ 1.02),$$
$$\hat{S}_u = (2.07,\ 1.04),$$
$$\hat{S} = (2.05,\ 1.03),$$
$$\hat{r} = 0.105,$$
$$\hat{\Upsilon}^2 = 0.99,$$

and

$$\hat{N} = 15.0.$$

It may appear rather strange that 67 percent of the data pairs were rejected when only 50 percent were noise-contaminated. However, it must be remembered that the noise was added to each of the four traces independently. The estimates are obviously far superior to the original ones from the total data set. The other points to note are: $\hat{\Upsilon}^2_{xy}$ does not maximize, $\hat{S}$ becomes worse as more points are rejected, $\hat{r}^2$ tends to increase as too many points are rejected, and $\hat{N}_{xy}$ displays only one maximum. These features were also unanimously displayed in *all* test runs made with those noise characteristics, and also in all runs with various combinations of percentage noise contributions and noise power.

Magnetotelluric data. The elements of the impedance tensor $\underline{Z}$ are commonly derived by using multiple linear regression methods. For example, for element Z_{yx} there exist two distinct solutions depending upon on which component (either E_y or H_x) the noise contribution is minimized in a least-squares sense. These solutions are, for element Z_{yx},

$$\hat{Z}^1_{yx} = \frac{\bar{C}_{xe} \cdot \bar{C}_{yy} - \bar{C}_{ye} \cdot \bar{C}_{xy}}{\bar{C}_{xx} \cdot \bar{C}_{yy} - \bar{C}_{xy} \cdot \bar{C}_{yx}}, \qquad (6a)$$

$$\hat{Z}^2_{yx} = \frac{\bar{C}_{ye} \cdot \bar{C}_{ey} - \bar{C}_{yy} \cdot \bar{C}_{ee}}{\bar{C}_{yx} \cdot \bar{C}_{ey} - \bar{C}_{yy} \cdot \bar{C}_{ex}}, \qquad (6b)$$

(dependency on frequency assumed), where e denotes E_y, x denotes H_x, and y denotes H_y. The two solutions are equivalent to considering, respectively, E_y as output and (H_x, H_y) as inputs, equation (6a); and H_x as output and (E_x, H_y) as inputs, equation (6b); and they minimize noise on their respective output. Both solutions are biased by noise occurring on their respective inputs. Assuming that the coherence between H_x and H_y is small, which is important for obtaining stable estimates of the impedance tensor elements, $\hat{Z}^1_{yx}$ is *downward-biased* mainly (but not exclusively) by noise on H_x, whereas $\hat{Z}^2_{yx}$ is *upward-biased* mainly by noise on E_y. The phases of $\hat{Z}^1_{yx}$ and of $\hat{Z}^2_{yx}$ are, however, not affected by noise contributions, and are equal.

For a multiple input/single output linear system, there is a choice of which coherence function to employ for picking out the noise degraded realizations. The ordinary coherence function between the output and one of the inputs is considered very unsuitable, however, because it does not give a correct measure of the correlation existing between the two components when any other inputs have a measurable effect on the output. The two coherence functions that appear to be reasonable to employ are (1) the *multiple* coherence between the output and all the inputs, or (2) the *partial* coherence between the output and the input of interest. For the example to be illustrated here, the former of these was chosen, i.e., the multiple, or "predicted," coherence.

Estimates of Z_{yx} from equations (6a) and (6b), for a typical data set are illustrated, in terms of MT apparent resistivity,

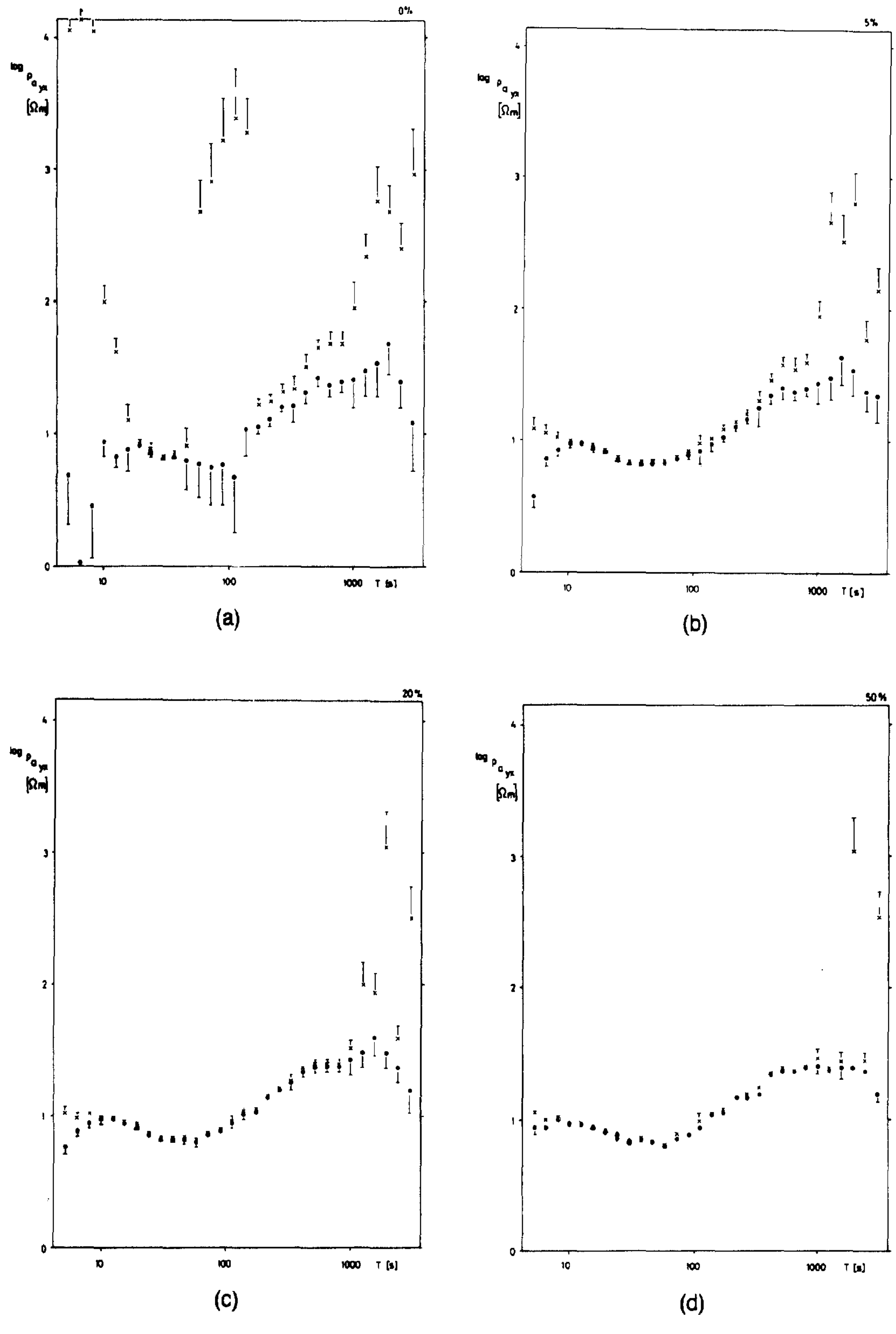

FIG. 2. (a) The two solutions for $\hat{\rho}_{a_{yx}}$, indicated by dots ($\hat{\rho}_a{}^1_{yx}$), and crosses ($\hat{\rho}_a{}^2_{yx}$) when all the available data are employed. The standard deviations of $\hat{\rho}_a{}^1_{yx}$ and $\hat{\rho}_a{}^2_{yx}$ are indicated by the one-sided error bars drawn downward and upward, respectively. (b) Same as (a) but for the three solutions that result after the worst 5 percent of the available data have been rejected. (c) Same as (b) but when 20 percent of the available data have been rejected. (d) Same as (b) but when 50 percent of the available data have been rejected.

in Figure 2a. The one-sided error bars are the standard deviations of the estimates, and are drawn upward for the solution "noise power in H_x minimized" (equation 6b), and downward for the solution "noise power in E_y minimized" (equation 6a). These estimates can be described as those that one would derive from a "normal" analysis of the available data.

The available estimates were treated by the rejection method described above to isolate and reject the worst realizations from 5 to 50 percent of the total. Figures 2b–2d display the apparent resistivities for 5, 20, and 50 percent rejection. The following points are worthy of note: the standard deviations of the estimates became quite small with increasing number of rejected realizations (this was mainly due to the increased multiple coherence). The shape of the two ρ_a curves became smooth. Obvious outliers in the period range 60–100 s were detected and rejected at the 5 percent stage. For the short periods ($T < 10$ s) and the long periods ($T > 1500$ s), an important improvement is accomplished. At these periods, the bias error was extremely large due to the low SNRs because of (1) weak natural signals for periods $T < 10$ s, and (2) decreasing resolution of the magnetic field variations for periods $T > 1500$ s due to the sensors employed (induction coils).

References

Akaike, H., 1967, Some problems in the application of the cross-spectral method, *in* Advanced seminar on spectral analysis of time series: B. Harris, Ed., John Wiley and Sons Ltd., 81–107.

Dekker, D. L., and Hastie, L. M., 1981, Sources of error and bias in a magnetotelluric depth sounding of the Bowen Basin: Phys. Earth Plan. Int., **25**, 219–225.

Gamble, T. D., Goubau, W. M., and Clarke, J., 1978, Magnetotellurics with a remote reference: Geophysics, **44**, 53–68.

Goubau, W. M., Gamble, T. D., and Clark, J., 1978, Magnetotelluric data analysis: Removal of bias: Geophysics, **43**, 1157–1166.

Jones, A. G., 1981, Transformed coherence functions for multivariate studies: IEEE Trans. Acoust., Speech, Signal Proc. **ASSP-29**, 317–319.

Kao, D., and Rankin, D., 1977, Enhancement of signal-to-noise ratio in magnetotelluric data: Geophysics, **42**, 103–110.

Lienert, B. R., Whitcomb, J. H., Phillips, R. J., Reddy, I. K., and Taylor, R. A., 1980, Long term variations in magnetotelluric apparent resistivities observed near the San Andreas fault in southern California: J. Geomag. Geoelectr., **32**, 757–775.

Reiersøl, O., 1950, Identifiability of a linear relation between variables which are subject to error: Econometrica, **18**, 375–389.

Estimation of the Magnetotelluric Impedance Tensor by the ℓ_1 and ℓ_2 Norms EM1.6

Francis J. Turk, Motorola, Inc.; James C. Rogers and Charles T. Young, Michigan Technological Univ.

Magnetotelluric impedance tensor estimates have conventionally been calculated using the least-squares (ℓ_2) formulation. However, noise on one or more channels can lead to biased and scattered estimates. In this paper, the use of the least-absolute value (ℓ_1) norm in estimating the impedance tensor is investigated. Use of the remote reference technique in recent years has reduced the effects of correlation between noise on different channels in most MT situations. However, significant amounts of conventional 4-channel MT data still remain to be processed to reduce the effects of noise contamination.

The ℓ_1 problem is formulated as a linear programming (LP) problem, using a modified version of the Simplex method of linear programming. Unlike the ℓ_2 norm, which minimizes the summed squared errors, the ℓ_1 norm seeks to minimize the summed absolute value of the errors. Therefore, estimates so obtained should be less susceptible to outlying points or "flyers." The processing technique is illustrated by applying noise to a set of MT data related by a known impedance. Comparisons of the ℓ_1 and ℓ_2 estimates are made for various noise levels and types.

In the magnetotelluric (MT) method of geophysical prospecting, a linear relationship is assumed to exist between horizontal components of the electric field **E** and the magnetic field **H**. The system is described at any frequency by the impedance tensor [Z] where

$$\begin{bmatrix} \mathbf{E}_x \\ \mathbf{E}_y \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_{xx} & \mathbf{Z}_{xy} \\ \mathbf{Z}_{yx} & \mathbf{Z}_{yy} \end{bmatrix} \begin{bmatrix} \mathbf{H}_x \\ \mathbf{H}_y \end{bmatrix}, \tag{1}$$

$$[\mathbf{E}] = [\mathbf{Z}][\mathbf{H}], \tag{2}$$

and Cartesian coordinates with z directed downward are indicated. In principle, two independent observations of the four fields yield a solution. However, since all physical measurements are contaminated by noise, redundancy is used to improve the estimate. This work describes and compares two types of impedance tensor estimation: the least-squares (ℓ_2) method and the least absolute-error (ℓ_1) method. Comparisons of the two are presented for single station data, although the results are applicable to remote stations as well.

Impedance estimation by the least-squares method

With N observations of the fields, the least squares (ℓ_2) approach minimizes the term

$$[Z_2] = [Z]\colon \sum_{i=1}^{N} (\mathbf{E}_i - Z\,\mathbf{H}_i)(\mathbf{E}_i \;P\; Z\,\mathbf{H}_i)^*. \tag{3}$$

As pointed out by Sims (1969), this will minimize the error caused by noise on the **E** channels. Similar minimizations for the other channels yield four equations to solve for the elements Z_{xx} and Z_{xy}. Of the six pairs, the two which use the $\mathbf{E}_x$-$\mathbf{H}_y$ and $\mathbf{E}_y$-$\mathbf{H}_x$ reference field pair become unstable as the geology becomes one-dimensional. Using the $\mathbf{H}_x$-$\mathbf{H}_y$ reference field pair, the ℓ_2 estimate for $\mathbf{Z}_{xy}$ becomes

$$\hat{Z}_{xy} = \frac{\langle \mathbf{H}_x\mathbf{H}_x^*\rangle\langle \mathbf{E}_x\mathbf{H}_y^*\rangle - \langle \mathbf{H}_x\mathbf{H}_y^*\rangle\langle \mathbf{E}_x\mathbf{H}_x^*\rangle}{\langle \mathbf{H}_x\mathbf{H}_x^*\rangle\langle \mathbf{H}_y\mathbf{H}_y^*\rangle - \langle \mathbf{H}_x\mathbf{H}_y^*\rangle\langle \mathbf{H}_y\mathbf{H}_x^*\rangle}, \tag{4}$$

where

$$\langle \mathbf{AB}^*\rangle = \frac{1}{N}\sum_{i=1}^{N} \mathbf{A}_i\mathbf{B}_i^*. \tag{5}$$

This will give an unbiased estimate of Z_{xy} provided the fields are not polarized over the period of measurement, and noise on the channels is random and independent of the MT signals. In practice, to reduce the significance of periods with poor signal, a weighted average of equation (5) is normally computed. The weighting can be based on the coherency between orthogonal **E-H** pairs, the multiple coherency between measured and predicted **E** fields, etc. However, the circumstances under which each weighting type should be used are not well understood. After computing the weighted average of equation (5), the crosspower files often cannot be reprocessed with a different weight type since they are stored cumulatively. To use a different weighting, a new run must be started, or provisions must be made for storing all weighted crosspowers for later computation.

Use of absolute error minimization

An alternative to least-squares minimization proposed for geophysical data equation (2) when the data contain outlying