Geophys. J. Int. (1989) **99**, 645–663

Transfer functions: smooth robust estimates by least-squares and remote reference methods

J. C. Larsen

Pacific Marine Environmental Laboratory, National Oceanic and Atmospheric Administration, 7600 Sand Point Way NE, Seattle, WA 98115-0070, USA

Accepted 1989 June 9. Received 1989 June 9; in original form 1987 November 9

SUMMARY

A method is described for finding the magnetotelluric transfer function that has the least amount of curvature consistent with most of the data and with a 1-D conductivity interpretation over the widest possible frequency range. This could be called an 'Occam' transfer function. It is represented by the transfer function for the best fitting 1-D conductivity model times a distortion function. The latter permits smooth departures of the transfer function from the 1-D case if the data are inconsistent with a 1-D interpretation. The transfer function, for single-station or remote reference magnetotelluric data, is found by a method of successive interations that is found to converge within six to eight iterations. The estimate of the transfer functions is made robust by using frequency and time weights that remove the effects of outliers in the time and frequency domain. If the weighted residuals for remote reference data satisfy certain necessary conditions for uncorrelated noise then the contribution to the noise by the electric and magnetic data can be estimated and used to evaluate the least-squares and remote reference estimates.

Examples illustrate the application of this method to artificial and real data. The latter consist of hourly cable voltage data from the Florida Straits, 1/256-s remote reference magnetotelluric survey data from the Phillippines and daily magnetic data from Tucson and Honolulu.

Key words: magnetotelluric transfer function

1 INTRODUCTION

Voltage measurements across the Florida Straits show fluid induced variations caused by changes in the Florida Current and geomagnetic induced variations caused by ionospheric sources. A technique is described whereby magnetic variations from a shore site can be used to separate the two types of variations using transfer functions. The usual least-squares method for estimating transfer functions was found to be unreliable due to the large fluid induced variations at the tidal and lowest frequencies. A simple analytic representation of the transfer function was therefore developed that could be robustly estimated using frequency and time weights. An example of the voltage variations across the Florida Straits (Fig. 1) shows that removing the geomagnetic and tidal variations and a few outliers (1.9 per cent of the values) eliminates most of the high frequency variations. This makes it possible to accurately observe the daily changes in the Florida Current. These non-tidal fluid induced variations are strongly correlated with velocity estimates of transport (Larsen & Sanford 1985) and are used to monitor the transport variations. The method of obtaining transfer functions was applied to other electromagnetic data in order to explore the general validity of the method and to determine the essential steps needed to obtain robust estimates.

Two basic methods have been used for estimating transfer functions. The standard method (Sims, Bostick & Smith 1971; Vozoff 1972) finds least-squares frequency band estimates using electric and magnetic data from a single site. These estimates minimize the variance of the residuals where the residuals are given by the observed electric data minus the values predicted from the magnetic data. These estimates are biased, however, by noise in the magnetic data, and the remote reference method using simultaneous local and remote magnetic data was developed in order to construct unbiased estimates (Gamble, Goubau & Clarke 1979; Clarke *et al.* 1983). These estimates minimize the modulus of the covariance between the locally and remotely derived residuals where the local residuals are given by the observed electric data minus the values predicted from the local magnetic data and the remote residuals are given by

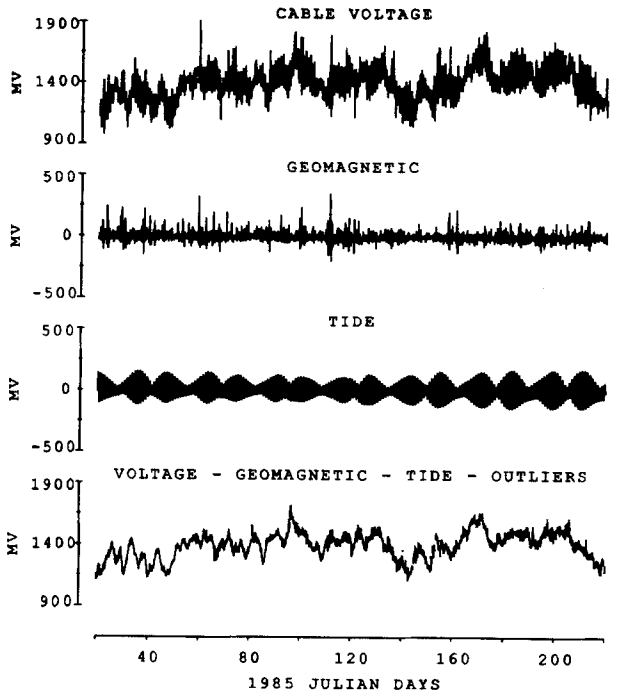

Figure 1. Voltages across the Florida Straits at 27°N (top curve), geomagnetic induced variations (second curve from top), tidal and solar variations (third curve from top), and non-tidal fluid induced variations (bottom curve) based on the voltages minus the geomagnetic and tidal variations and outliers.

the observed electric data minus the values projected from the remote magnetic data. These estimates are not biased if the noise is uncorrelated and are expected to be an improvement over the least-squares estimates. They can, however, be unreliable and give inferior results compared to the least squares estimates if the noise is correlated or the remote magnetic data are uncorrelated with the local magnetic data. Necessary conditions for the noise to be uncorrelated are derived using the locally and remotely derived residuals. These conditions, if satisfied, permit estimates of the noise contributed by the electric, local magnetic and remote magnetic data. A comparison of the noise can then be used to evaluate the quality of the least-squares and remote reference estimates.

Outliers are a common occurrence caused by instrumental and recording problems, and by source problems such as power generators, oceanic motional induced signals, treated here as noise, and geomagnetic events having high wavenumbers. Their presence can cause catastrophically poor estimates of the transfer function and its variance if standard least-squares or remote reference methods are used. A number of robust techniques have been developed (Egbert & Booker 1986; Chave, Thompson & Ander 1987; Chave & Thompson, 1989) for reducing the effects of outliers by using weighted section averaging. Weighted section averaging suffers, however, from the limitations that a section is downweighted even though there may only be a few very large outliers and that all sections may be contaminated by outliers. It has been found, for some data sets, that there are no clean sections of any useful duration. In the present method every datum in the frequency and time domain are examined and dealt with by frequency and time weights. The main goal is to find a set of weights that increase the coherence between the electric and magnetic data while, hopefully, downweighting only a small percentage (<20 per cent) of the data. The present method is basically the application of robust techniques to the iterative method used for determining smooth transfer estimates described in Larsen (1975, 1980).

The use of a smooth transfer function was inspired by the smooth admittance functions used in tidal analysis (Munk & Cartwright 1966) and is justified because the magnetotelluric (MT) transfer function is expected to be smooth and continuous for most geologic sites that are not close to surface discontinuities. The smooth continuous representation is found to be essential for estimating the geomagnetic variations at all frequencies, including frequencies where the fluid induced variations, such as the tides, overwhelm the geomagnetic variations.

The time domain electric and magnetic data are designated by, respectively $e(t)$ and $b(t)$ for $t = j\delta$, $j = 1, \ldots, 2J$ values with δ time increment and the frequency domain data are designated by $E(\omega)$ and $B(\omega)$ for radian frequency $\omega_j = j\pi/J\delta$, $j = 1, \ldots, J$ values where the mininum frequency is $\omega_1 = \pi/J\delta$ and the maximum is $\omega_J = \pi/\delta$. In general, the magnetic data must be complete but data gaps can be tolerated in the electric data provided these are replaced, in the iterative process, by the predicted values.

The data are preprocessed by removing the following: the mean and linear trend; the variations in the trend using a least squares cubic spline; the annual, tidal and solar diurnal variations and its harmonics, if they have a large presence in the residuals, using the known annual, tidal and solar diurnal frequencies, and the least-squares estimates of their amplitude and phase.

1.1 MT Relationship

The frequency domain electric data **E**, local magnetic data **B** and remote magnetic data **B**, are related through the MT relationships by

$$\mathbf{E} = \mathbf{B} \cdot \mathbf{Z} + \mathbf{R} \tag{1a}$$

and

$$\mathbf{E} = \mathbf{B} \cdot \mathbf{Z} + \mathbf{R}, \tag{1b}$$

where **Z** is the local MT transfer function tensor, **Z** is the remote MT transfer function tensor, **R** are the locally derived residuals and **R** are the remotely derived residuals. The **E**'s and **R**'s are J-vectors for J frequencies, the **B**'s are $J \times K$ matrices for K magnetic components (usually $K = 2$) and the **Z**'s are $K \times J$ matrices. Variables referring to the remote data are designated by underbars.

This paper starts with frequency band estimates, then describes the method used to obtain robust estimates, then describes the method for constructing smooth estimates, then derives the conditions for estimating the noise contributed by the electric and magnetic data, and finally describes the application of the method to various types of data.

2 BAND AVERAGED ESTIMATES

Least-squares band averaged estimates of the local and remote transfer functions, minimize, respectively, $|\mathbf{R}^H \cdot \mathbf{R}|$ and $|\underline{\mathbf{R}}^H \cdot \underline{\mathbf{R}}|$, the variance of the residuals, where H is the Hermite transpose. The local and remote estimates are given, respectively, by

$$\mathbf{z} = (\mathbf{B}^H \cdot \mathbf{B})^{-1} \cdot (\mathbf{B}^H \cdot \mathbf{E}) \tag{2a}$$

and

$$\underline{\mathbf{z}} = (\underline{\mathbf{B}}^H \cdot \underline{\mathbf{B}})^{-1} \cdot (\underline{\mathbf{B}}^H \cdot \mathbf{E}) \tag{2b}$$

for N bands having $\hat{J}$ frequencies per band. Then $\mathbf{E}$ is a $\hat{J}$-vector, the $\mathbf{B}$'s are $\hat{J} \times K$ matrices for K magnetic components and the $\mathbf{z}$'s are $K \times N$ matrices. Note that the diagonal terms of $(\mathbf{B}^H \cdot \mathbf{B})$ and $(\underline{\mathbf{B}}^H \cdot \underline{\mathbf{B}})$ are biased by noise in the magnetic components.

Remote reference band estimates of the local and remote transfer functions minimize $|\underline{\mathbf{R}}^H \cdot \mathbf{R}|$: the modulus of the covariance between the locally and remotely derived residuals. This interpretation allows one to compute the transfer function without the replacement method of Gamble *et al.* (1979) and shows that the remote reference method demands local and remote transfer functions. Estimates are given, respectively, by

$$\mathbf{z} = (\underline{\mathbf{B}}^H \cdot \mathbf{B})^{-1} \cdot (\underline{\mathbf{B}}^H \cdot \mathbf{E}) \tag{3a}$$

and

$$\underline{\mathbf{z}} = (\mathbf{B}^H \cdot \underline{\mathbf{B}})^{-1} \cdot (\mathbf{B}^H \cdot \mathbf{E}) \tag{3b}$$

for N bands having $\hat{J}$ frequencies per band.

Note that the biasing of the remote reference estimate only vanishes if the noise is uncorrelated. This estimate cannot, therefore, be assumed to be superior to the least squares estimates unless the noise is known to be uncorrelated. For example, if $\mathbf{B}$ is noise free but $\underline{\mathbf{B}}$ consists only of noise that is uncorrelated with $\mathbf{B}$ then terms such as $(\underline{\mathbf{B}}^H \cdot \mathbf{B})$ and $(\mathbf{B}^H \cdot \underline{\mathbf{B}})$ are small and the remote reference estimate becomes unstable and clearly inferior to the least-squares estimate. Thus remote estimates may at times be much worse than the least-squares estimates.

Band averaged estimates depend on the assumption that the transfer function is constant over the band of frequencies from which it is estimated. The transfer function as a function of frequency can then be expressed in terms of the Heaviside function $\hat{H}$ as

$$Z(\omega) = \sum_{n=1}^{N} z_n[\hat{H}(\omega - \omega_{n\hat{J}-\hat{J}+1}) - \hat{H}(\omega - \omega_{n\hat{J}})], \tag{4}$$

where the band estimates are z_n for $n = 1, \ldots, N$ bands having $\hat{J}$ frequencies per band and $\hat{H}(\omega) = 0$ for $\omega < 0$ and $\hat{H}(\omega) = 1$ for $\omega \geq 0$. Hence, the transfer function $Z(\omega)$ estimated by (4) are not continuous but rather resemble a stair-case type function that is flat within the frequency bands but has offsets between bands. Therefore, a band average estimate may be incorrect if there is an appreciable variation of the transfer function within the frequency band.

3 ROBUST ESTIMATES

The robustification of the least-squares or remote reference estimates of the transfer function and its variance is accomplished by frequency and time weights. The frequency weights consists of three parts: the prewhitening continuum weights $W(\omega)$, the line frequency weights $F(\omega)$ and the post whitening weights $V(\omega)$. The latter are sometimes found to be unnecessary. The time weights are $g(t)$. The weights are applied to the observed data in the following order: $W(\omega)$, $F(\omega)$, $g(t)$ and $V(\omega)$. The weights $W(\omega)$ prewhiten the continuum part of the residuals making them nearly independent. The frequency weights $F(\omega)$ eliminate any spectral peaks in the residuals, the time weights $g(t)$ eliminate the outliers in the time domain and the weights $V(\omega)$ postwhiten the weighted residuals. The frequency weights $F(\omega)$ are applied before the time weights $g(t)$ because time outliers are usually found to have less effect on $F(\omega)$ because the frequency domain noise in the prewhitened residuals is usually narrow-band and the time domain noise is broad-band.

Experiments with various real time series show that this sequence of weights usually reduces the residuals to one for which the values are nearly independent and identically distributed. This is desirable because the least squares estimates will then have minimum variance. Furthermore, the distribution of the weighted residuals are most commonly found to be approximately normal, which means that the statistical significance of the estimates can be readily made. Experience shows that the weights $F(\omega) = 0$ and $g(t) = 0$ play the most important role, followed by the continuum weight $W(\omega)$, then the weights $0 < F(\omega) < 1$ and $0 < g(t) < 1$ and then $V(\omega)$.

The weights are applied to the electric and magnetic data by the same steps. For example, the Fourier transform of the electric data are first frequency weighted yielding

$$E_{wF}(\omega) = F(\omega)W(\omega)E(\omega). \tag{5}$$

Hence frequency weighting is a simple multiplication in the frequency domain and a convolution integral in the time domain. Note that $E_{wF}(\omega) = E_{Fw}(\omega)$. The frequency weighted electric data are then time weighted yielding

$$e_{wFg}(t) = g(t)e_{wF}(t) \tag{6}$$

Hence time weighting is a simple multiplication in the time domain and a convolution integral in the frequency domain. Finally, the electric data are frequency weighted yielding

$$E_{wFgv}(\omega) = V(\omega)E_{wFg}(\omega) \tag{7}$$

3.1 Continuum weights

The coefficients w_m, $m = 1, \ldots, M$ of the continuum weights $W(\omega)$ are derived from the auto-regressive relationship

$$r_{\hat{F}w}(t) = r_{\hat{F}}(t) - \sum_{m=1}^{M} w_m r_{\hat{F}}(t - m\delta), \tag{8}$$

where δ is the time increment. The residuals $r_{\hat{F}}(t)$ are based on the Fourier inverse of $R_{\hat{F}}(\omega) = \hat{F}(\omega)R(\omega)$ with the frequency rejection weights $\hat{F} = 0$ for large residual lines

and $\hat{F} = 1$ otherwise (see next section). The use of rejection frequency weights allows one to use a low order ($M \leq 10$) representation of the continuum weights. The autoregressive relationship was found to be very effective and convenient for whitening the continuum part of the residuals. The continuum weights in the frequency domain are given by

$$W(\omega) = 1 - \sum_{m=1}^{M} w_m \exp(-im\delta\omega) \tag{9}$$

The coefficients w_m are found by solving the Yule–Walker equations in terms of the auto-correlation functions ρ_n of the residuals

$$\rho_n = \sum_{m=1}^{M} w_m \rho_{|m-n|}$$

for $n = 1, \ldots, M$, where $\rho_n = \sum_t r_{\hat{F}}(t) r_{\hat{F}}(t - n\delta)$, using the fast and accurate Levinson algorithm (see Thomson 1977 for details).

3.2. Rejection weights

The main purposes of the frequency rejection weights $\hat{F}(\omega)$ and time rejection weights $\hat{g}(t)$ are to help determine the continuum weight $W(\omega)$ and to identify the large outliers.

The set of frequency and time rejection weights used to remove the residual lines and outliers are defined in general by

$$\hat{y}(x) = \begin{cases} 1 & \text{for } x < \hat{x}_M \\ 0 & \text{for } x > \hat{x}_M \end{cases}$$

where $\hat{y}$ are, respectively, the rejection frequency weights $\hat{F}$ or the rejection time weights $\hat{g}$, and x are the scaled frequency residuals $|R_{w\hat{w}}(\omega)|/\varepsilon$ or time residuals $|r_{\hat{F}w}(t)|/\varepsilon$. The threshold level is given by

$$\hat{x}_M(p, J) = \{J[1 - (p/J)^{1/(J-1)}]\}^{1/2}$$

(Fisher 1929) at the $100(1-p)$ per cent level for J normally distributed terms.

In order to detect the spectral lines in the residuals one needs to first whiten the continuum using preliminary continuum weights $W(\omega)$. These are derived from the residuals $r_{\hat{F}}(t)$ using the rejection weights from the previous iteration. It was found that $W(\omega)$ can be contaminated by undetected lines for the first few iterations and a modification $\hat{W}(\omega) = 1/\varepsilon(\omega)$ to the continuum weight was therefore necessary. The $\varepsilon(\omega)$ is a linear interpolation to the robust estimates of the scale ε_n for $n = 1, \ldots, N$ frequency bands where ε_n is equal to 1.483 times the median of $|R_{\hat{F}w}(\omega)|$ for the nth frequency band. See Hogg (1979) and Chave *et al.* (1987) for a discussion about robust estimates of scale. It is assumed here that the mean is zero. The new frequency rejection weights $\hat{F}(\omega)$ are then derived from the scaled frequency residuals $|R_{w\hat{w}}(\omega)|/\varepsilon$ where ε is the robust scale of $|R_{w\hat{w}}(\omega)|$.

In order to detect large outliers in the residuals one examines the scaled time residuals $|r_{\hat{F}w}(t)|/\varepsilon$ using the rejection weights $\hat{F}$, continuum weights W and a robust estimate of the scale ε equal to 1.483 times the median of $|r_{\hat{F}w}(t)|$. The rejected or missing values are then replaced by the predicted values (Appendix C) in order to lessen, for the next iteration, the rejection of good values adjacent to the outliers or missing values.

3.3 Frequency and time weights

These weights are used to reduce the effects of individual outliers in the time and frequency domain. In general, the weighted scaled frequency or time residuals are given by $x' = y(x)x$ where the y are, respectively, the frequency weights $F(\omega)$ or time weights $g(t)$. The x are the scaled frequency residuals $|R_w(\omega)|/\varepsilon$ using continuum weights $W(\omega)$ based on $r_{\hat{F}}(r)$ and a robust ε equal to 1.483 times the median of $|\mathcal{R}e\{R_w(\omega)\}|$ and $|\mathcal{I}m\{R_w(\omega)\}|$ or the scaled time residuals: $r_{Fw}(t)/\varepsilon$ using frequency weights $F(\omega)$, continuum weights $W(\omega)$ based on $r_F(t)$, and a robust ε equal to 1.483 times the median of $|r_{Fw}(t)|$.

In general the weights satisfied $0 \leq y \leq 1$ and are assumed to decrease for increasing $|x|$, i.e. $dy/dx \leq 0$ for $x > 0$ and $dy/dx \geq 0$ for $x < 0$. In addition the binwidths dx for the unweighted residuals should map into positive binwidths dx' for the weighted residuals, i.e. $dx'/dx > 0$. In the following discussion let x be positive for simplicity's sake as the case for negative x is similar in development.

A common procedure is to assume an analytical form for $y(x)$ (see Holland & Welsch 1977; Hogg 1979 for a list of commonly used weights). A simple example is the rejection weights $y(x) = 1$ for $0 \leq x \leq A$ and $y(x) = 0$ for $A < x$ for a given A. Then $dx'/dx = 1$ for $0 \leq x \leq A$ and $dx'/dx = 0$ for $A < x$. These weights do not satisfy the condition of positive bin widths for $A < x$. They also have two additional problems. They do not change the distribution for $x \leq A$ and they truncate the tails of the distribution for $A < x$ which can play havoc with the estimates if there is a high density at $x = A$ (Huber 1964).

Another class of weights have therefore been used for changing the tails of the distribution so that it is more normally distributed. An example is the Huber weights (Huber 1964) where $y(x) = 1$ for $0 \leq x \leq A$ and $y(x) = (A/x)^{1/2}$ for $A < x$ for a given A. Then $dx'/dx = 1$ for $0 \leq x \leq A$ and $dx'/dx = 1/2$ for $A < x$. These weights satisfy the basic condition of positive bin widths for the weighted residuals and are applicable for a distribution that is normal in the center and Laplacian in the tails. Huber weights, however, are not very effective in dealing with large outliers because they decrease rather slowly with increasing x.

Another class of weights, called redescending weights, are therefore used such as the Tukey biweight $y(x) = 1 - (x/A)^2$ for $0 \leq x < A$ and $y(x) = 0$ for $x \geq A$. Then $dx'/dx = 1 - 3(x/A)^2$ for $0 \leq x < A$ and $dx'/dx = 0$ for $A \leq x$ for a given A. These weights do not satisfy the condition of positive binwidths for $A/\sqrt{3} \leq x$ and will contaminate the distribution if there is a high density at $x = A/\sqrt{3}$. This class of weights was found, for some of the examples, to yield distributions that had small secondary peaks at the location where $dx' = 0$.

Data adaptive weights are therefore used here. These are defined by $y(x) = 1$ for $0 \leq x < x_l$, $0 \leq y(x) \leq 1$ for $x_l \leq x \leq x_M$ and $y(x) = 0$ for $x_M < x$. This representation assumes there is a region, $0 \leq x < x_l$ where the distribution is

normal, a region $x_I \leq x \leq x_M$ where the weighted distribution can be made normal and a region, $x_M < x$, where the residuals are rejected. This is a reasonable assumption and is found to be valid for all the examples because most of the whitened residuals are found to be normally distributed. The frequency and time weights are therefore only necessary for modifying the tails of the distribution so that the tails are consistent with the normal distribution. It is therefore assumed that at least 50 per cent of the residuals are normally distributed and that only 1 per cent of the residuals should be rejected if the distribution is normal. The threshold levels are then $|x_I| = 0.7$ and $|x_M| \geq 2.6$ for the time weights. The frequency weights are assumed to be real and the same for the real and imaginary parts of the frequency residuals. This avoids introducing any spurious phase shifts, but means examining the modulus of residuals. This has a Rayleigh distribution if the real and imaginary parts are normally distributed. The threshold levels are then $x_I = 1.1$ and $x_M \geq 3.0$ for the frequency weights.

The weights $y(x)$ for $x > x_I$ are estimated from the tails of the distribution. Let N_i be the observed number of values of x in the ith bin and $N_i' = CS_i$ be the expected number of values in the ith bin for the weighted residuals x' where the ith bin is $x_{i-1} \leq x \leq x_i$, $x_i = i\Delta$ and Δ is the binwidth. The $x_i' = y_i x_i$ and $y_i = y(x_i)$ are, respectively, the values of the weighted residuals and weights at the boundaries of the bins. The C is a scaling factor based on the normal part of the distribution and is given, for the minimum χ^2 misfit, by

$$C = \left[\sum_{i=-I}^{I} (N_i^2/S_i) \Big/ \sum_{i=-I}^{I} S_i \right]^{1/2} \tag{10}$$

where the χ^2 misfit (Press *et al.* 1986) is

$$\chi_I^2 = \sum_{i=-I}^{I} (N_i - N_i')^2 / N_i'$$

The

$$S_i = \int_{x_{i-1}'}^{x_i'} f(x)\, dx \tag{11}$$

is the contribution to the cumulative distribution from the ith bin using the distribution function $f(x)$ that is Rayleigh, $f(x) = x \exp(-x^2/2)$ for the frequency residuals and is normal, $f(x) = \exp(-x^2/2)$ for the time residuals.

For the possibly non-normal region $i > I$, the weights are $y_i \leq 1$ and the conditions $dy/dx \leq 0$ and $dx'/dx > 0$ require $y_{i-1}[1 - \Delta/x_i] < y_i \leq y_{i-1}$. The value of y_i for the bin, $i = I + 1$, is found by varying y_i by small increments within these limits and choosing the y_i that yields the smallest χ^2 misfit between the observed N_i and the expected number $N_i' = CS_i$ by computing S_i (11) and using the estimated scaling factor C (10). The value of y_i for the next bin, $i = I + 2$, is then found and the process repeated up to $i = M$ that is at least greater than the 1 per cent threshold level and stopped when the χ^2 misfit between N_i and N_i' exceeds the 95 per cent confidence level for $I + M - 1$ degrees of freedom. Continuous values of $y(x)$ are then generated by a linear interpolation to the estimated values y_i for $i = I, \ldots, M$ with $y_I = 1$ and $y_M = 0$. Finally the rejection weights are included by setting $g(t) = 0$ for $\hat{g}(t) = 0$ and $F(\omega) = 0$ for $\hat{F}(\omega) = 0$.

The normalcy of the weighted residuals in the frequency and time domain, excluding frequency values for which $F = 0$ and time values for which $g = 0$, are examined by computing the χ^2 misfit between the observed values N_i and expected values N_i' and comparing it to the 95 per cent confidence level and by plotting the distributions of N_i and N_i'. These plots are used to visually examine the normalcy of the weighted residuals. The iterative steps by which the weights are generated are outlined in Appendix E.

3.4 Weighted MT relationship

The application of the weighting to the MT relationship for a single magnetic component yields

$$E_{wFgv}(\omega) = V(\omega)\{G(\omega - \hat{\omega}) * [B_{wF}(\hat{\omega}) Z(\hat{\omega})]\} + R_{wFgv}(\omega) \tag{12}$$

where $G(\omega)$ is the Fourier transform of $g(t)$ and $*$ represents a convolution integral over the available frequencies.

In order to solve this equation by the method of successive iterations let the next estimate of the transfer function Z_i be equal to a perturbation function U_i times the previously estimated Z_{i-1}, i.e. $Z_i = U_i Z_{i-1}$, and solve (12) in terms of U_i. Equation (12), ignoring the subscript i, then becomes

$$E_{wFgv}(\omega) = V(\omega)\{G(\omega - \hat{\omega}) * [B_{ZwF}(\hat{\omega}) U(\hat{\omega})]\} + R_{wFgv}(\omega), \tag{13}$$

where $E_{wFgv}(\omega)$ and $B_{ZwF}(\omega) = F(\omega)W(\omega)Z(\omega)B(\omega)$ are derived from the previous $i - 1$ estimates of the weights and transfer function. If Z_{i-1} is a reasonably good approximation of the transfer function, i.e. $Z_i(\omega) \approx Z_{i-1}(\omega)$, then the perturbation function approximates a constant that can be brought outside the convolution integral. The approximation $G * (B_{ZwF} U) \approx B_{ZwFg} U$ will then be valid and (13) becomes

$$E'(\omega) \approx P'(\omega) U(\omega) + R'(\omega), \tag{14}$$

where the weighted electric data are designated by

$$E'(\omega) = E_{wFgv}(\omega) \tag{15}$$

and the weighted predicted data are designated by

$$P'(\omega) = B_{ZwFgv}(\omega) \tag{16}$$

The approximate weighted MT relationship in matrix form for the local and remote weighted data can then be written, respectively, as

$$\mathbf{E}' \approx \mathbf{P}' \cdot \mathbf{U} + \mathbf{R}' \tag{17a}$$

and

$$\mathbf{E}'' \approx \mathbf{P}'' \cdot \underline{\mathbf{U}} + \underline{\mathbf{R}}'', \tag{17b}$$

where $'$ and $''$ indicates that the weights for the local and remote magnetic data are, in general, different. The $\mathbf{E}$'s and $\mathbf{R}$'s are J-vectors for J frequencies, the $\mathbf{P}$'s are $J \times K$ matrices for K magnetic components and the $\mathbf{U}$'s are $K \times J$ matrices. In applying these weights to the electric data the missing and large outliers were replaced by the predicted values (Appendix C). This was found to be necessary in order to prevent leakage from the gaps to the adjacent

values. The magnetic data is assumed to be free of outliers and missing values.

Note that data windows are equivalent to time weights $g(t)$. Using the windowed series $\mathbf{E}_g$ and $\mathbf{B}_g$ and the MT relationship $\mathbf{E}_g = \mathbf{B}_g \cdot \mathbf{Z} + \mathbf{R}_g$ implicitly implies the approximation $G*(BZ) \approx B_g Z$. This may not be valid unless the window $g(t)$ is so smoothly varying that its Fourier transform, $G(\omega)$, consists of a narrow peak around $\omega = 0$.

The application of the continuum weights on the electric and magnetic data can be carried out in the time domain using (8) or in the frequency domain using (9). The latter is subject to round off errors and the time domain process is therefore used for the examples. In applying the weights, the order of weighing W and F can be reversed. Then the weights W, g and V are applied sequentially in the time domain. The weighted electric and predicted data are then designated, respectively, by $E'(\omega) = E_{Fwgv}(\omega)$ and $P'(\omega) = B_{ZFwgv}(\omega)$.

4 SMOOTH ESTIMATES

Smooth transfer functions are represented by

$$Z(\omega) = D(\omega)\bar{Z}(\omega), \tag{18}$$

where $\bar{Z}(\omega)$ is a 1-D model transfer function described in the next section and $D(\omega)$ is a smooth distortion function represented by

$$D(\omega) = \sum_{n=1}^{N} d_n S_n(\omega) \tag{19}$$

that allows smooth departure of the transfer estimates from $\bar{Z}(\omega)$. The coefficients d_n are, in general, complex and the $S_n(\omega)$ are various frequency dependent functions described later. In general, $\bar{Z}(\omega)$ is needed for the convergence of the transfer estimate to a 1-D model transfer function if the weighted data are found to be consistent with a 1-D interpretation because $D(\omega)$ by itself will usually be a crude approximation to a 1-D model transfer function.

4.1 1-D model transfer function

Necessary and sufficient conditions for the existence of a 1-D model transfer function are given by Weidelt (1986) and the conditions for single and adjacent pairs of transfer estimates have been modified here (Appendix A) to allow for noise in the estimates. These modified 1-D Weidelt conditions are used to establish the widest 1-D frequency range. The 1-D model transfer function for this 1-D frequency range has the analytic representation (Weidelt 1972; Parker 1980)

$$\bar{Z}(\omega) = \sum_{n=1}^{N-1} \bar{z}_n (1 + i a_n/\omega)^{-1} - i\omega \bar{z}_N \tag{20}$$

for real and positive coefficients $\bar{z}_n$ and a_n. The low frequencies limit is $\bar{Z}(\omega) \approx -i\omega(\sum_{n=1}^{N-1} \bar{z}_n/a_n + \bar{z}_N)$ and the high frequency limit is $\bar{Z}(\omega) \approx -i\omega\bar{z}_N$. These are transfer functions for an insulating layer above a perfect conductor at, respectively, depths $\sum_{n=1}^{N-1} \bar{z}_n/a_n + \bar{z}_N$ and $\bar{z}_N$.

The conductivity model represented by (20) is called a D^+ model (Parker 1980) and consists of the stack of horizontally conducting thin layers separated by insulators. The a's, $\bar{z}$'s and the optimum number of layers N are found by a least-squares procedure using the D^+ inversion algorithms described in Parker & Whaler (1981) and applied here to 30 or fewer estimates of Z and its variance distributed over the 1-D frequency range. The 1-D model transfer functions are then generated by (20) for the full frequency range using the estimated a's and $\bar{z}$'s.

4.2 Distortion function

The distortion function should have a representation that makes it possible for the transfer estimates to converge to a 1-D model transfer function if the weighted data are found to be consistent with a 1-D interpretation. The form of the frequency dependent function, $S_n(\omega)$, is therefore derived from the ratio of transfer functions for different types of 1-D models.

For example, the ratio of the transfer functions, $\bar{Z}'$ and $\bar{Z}''$, for two D^+ models can be expressed in terms of partial fractions as

$$\bar{Z}'/\bar{Z}'' = \sum_{n=1}^{N-1} d_n (1 + i b_n/\omega)^{-1} - i\omega d_N$$

because (20) is itself a ratio of two polynomials. Thus $\bar{Z}'/\bar{Z}''$ is also a ratio of two polynomials and the frequency dependent function for the distortion (19) can be given by

$$S_n(\omega) = (1 + i b_n/\omega)^{-1} \tag{21a}$$

for $n = 1, \ldots, N-1$, and $S_N(\omega) = -i\omega$. I have not been successful in applying (21a) because the representation is found to be sensitive to the number and values of the b's. The dependency of the partial fractions on frequency suggests, however, that (21a) can be approximated by a power series in frequency. The frequency dependent term can then be written as

$$S_n(\omega) = (2\omega/\omega_J - 1)^{n-1} \tag{21b}$$

and is found to give good results.

Layered models of thickness h_n and conductivity σ_n where $\sigma_n h_n^2$ is the same constant for each layer are called H^+ models (Parker 1980). If the conductivities are close to a uniformly conducting half space, then the transfer function for N layers can be approximated (Larsen 1981) by

$$\bar{Z}(\omega) \approx (-i\omega/\mu_0\sigma_0)^{1/2} \exp \sum_{n=1}^{N} r_n S_n(\omega),$$

where $(-i\omega/\mu_0\sigma_0)^{1/2}$ is the transfer function for a half space of conductivity σ_0, μ_0 is the magnetic permeability of free space and $r_n = 0.5 \ln(\sigma_{n-1}/\sigma_n)$ is the logarithm of the conductivity ratio between layers. The frequency dependent function is

$$S_n(\omega) = \exp[-(n-1)(-i\omega)^{1/2}\beta], \tag{21c}$$

where $\beta = (4\mu_0\sigma_n)^{1/2} h_n$ is the same constant for each layer. The approximation improves in accuracy as r_n or the frequency approaches zero. The low frequency limit is $\bar{Z}(\omega) \approx (-i\omega/\mu_0\sigma_N)^{1/2}$ and the high frequency limit is $\bar{Z}(\omega) \approx (-i\omega/\mu_0\sigma_1)^{1/2}$. These are transfer function for half spaces of, respectively σ_N and σ_1 conductivities.

The ratio of the transfer functions $\bar{Z}'$ and $\bar{Z}''$ for two

nearly equal H^+ models can then be approximated by

$$\bar{Z}'/\bar{Z}'' \approx 1 + \sum_{n=1}^{N-1} (r'_n - r''_n) S_n(\omega).$$

Hence (21c) can be used as the frequency dependent function for the distortion function. I have found good results using (21c). Its dependency on the square root of frequency suggests an approximation by a power series in square root of frequency. The frequency dependent term can then be written as

$$S_n(\omega) = [2(\omega/\omega_J)^{1/2} - 1]^{n-1}. \tag{21d}$$

This representation gives good results and avoids the problem of specifying β.

The use of (21b) implies that the conductivity models are terminated at depth by a perfect conductor and the use of (21c) and (21d) implies that the conductivity models merge, at low frequencies, to a uniformly conducting half space. There is therefore a fundamental difference between the representations (21b) and (21d) and these representations may lead to different low frequency estimates of the transfer function. In general the better approximation to the low frequency part of the transfer function is probably given by a perfect conductor at depth because the conductivity generally increases with depth. The series expansion in frequency (21b) is therefore used for the examples although (21d) is found to give comparable results.

4.3 Smooth MT relationship

Smooth continuous transfer functions are found by representing the perturbation function in (17a,b) by a series expansion

$$U(\omega) = \sum_{n=1}^{N} u_n S_n(\omega) \tag{22}$$

using one of the frequency dependent functions (21b–d). The MT relationship (13) can then be written, without any approximation, as

$$E'(\omega) = \sum_{n=1}^{N} u_n S'_n(\omega) + R'(\omega), \tag{23}$$

where $E'(\omega) = E_{Fwgv}(\omega)$ and $S'_n(\omega) = V(\omega)\{G(\omega - \hat{\omega}) * [B_{Zfw}(\hat{\omega}) S_n(\hat{\omega})]\}$ can be computed directly from the time series using the previously estimated weights and transfer function Z. The preferred solution of the MT relationships would be to solve (23) for the u's by generating $S'_n(\omega)$ but this requires KNJ storage that can easily exceed the capacity of a normal computer for long time series. Another approach is to generate the cross spectral terms $\sum_\omega S'_m(\omega) S'_n(\omega)^*$ as they are needed, but this requires a vast increase in computing time because they are convolved with G.

The approximate weighted MT relationships (17a,b) are therefore used and solved by successive iterations. It was found that the solutions sometimes converged to oscillating values if one assumed that the next estimate was given by $Z_i = U_i Z_{i-1}$. This problem was avoided by assuming $Z_i = U_i \bar{Z}_{i-1} D_{i-1}$. Thus the next estimate of the transfer function is always restarted from the previous best fitting 1-D transfer function $\bar{Z}$ and distortion function D smoothed to eliminate any oscillations. These are then used to compute the predicted data $P(\omega) = Z(\omega) D(\omega) B(\omega)$. The solutions for the coefficients u_n of the perturbation U are then found from the weighted electric and predicted data using the MT relationships. These are given, in matrix form, by

$$\mathbf{E}' \approx \mathbf{p}' \cdot \mathbf{u} + \mathbf{R}' \tag{24a}$$

and

$$\mathbf{E}'' \approx \underline{\mathbf{p}}'' \cdot \underline{\mathbf{u}} + \underline{\mathbf{R}}'', \tag{24b}$$

where the $\mathbf{E}$'s and $\mathbf{R}$'s are J-vectors having, respectively, elements e_j and r_j for $j = 1, \ldots, J$ frequencies; the $\mathbf{u}$'s are M-vectors of dimension $M = KN$ having elements u_m with $m = n + N(k-1)$, for $n = 1, \ldots, N$ terms and $k = 1, \ldots, K$ magnetic components; and the $\mathbf{r}$'s are $J \times M$ matrices having elements $p'_{jm} = P'_k(\omega_j) S_n(\omega_j)$ or $\underline{p}''_{jm} = \underline{P}''_k(\omega_j) S_n(\omega_j)$ for the kth magnetic component of, respectively, the weighted locally or remotely predicted data where $P'(\omega) = P_{Fwgv}(\omega)$ and $\underline{P}''(\omega) = \underline{P}''_{Fwgv}(\omega)$ for $P(\omega) = \bar{Z}(\omega) D(\omega) B(\omega)$ and $\underline{P}(\omega) = \underline{\bar{Z}}(\omega) \underline{D}(\omega) \underline{B}(\omega)$.

4.4 Noise variance estimates

The examples studied show that the distribution of the weighted residuals are consistent with the assumption that the weighted residuals are identically distributed. Furthermore their white spectrum shows they are indepenent. In order to obtain a tractable solution, the assumption is therefore made here that the noise resides mainly in the weighted electric field. This is consistent with the oceanographic observation that the fluid induced variations in the Florida Straits, here the noise, resides in the electric data. The covariance of the noise is then given by

$$\text{Cov}\,(\mathbf{E}' \cdot \mathbf{E}'^{\mathrm{H}}) = \text{Var}\,|R'|\,\mathbf{I} \tag{25a}$$

and

$$\text{Cov}\,(\mathbf{E}'' \cdot \mathbf{E}''^{\mathrm{H}}) = \text{Var}\,|\underline{R}''|\,\mathbf{I} \tag{25b}$$

where $\mathbf{I}$ is the identity matrix. The variances of the residuals are

$$\text{Var}\,|R'| = |\mathbf{R}'^{\mathrm{H}} \cdot \mathbf{R}'|/\nu' \tag{26a}$$

and

$$\text{Var}\,|\underline{R}''| = |\underline{\mathbf{R}}''^{\mathrm{H}} \cdot \underline{\mathbf{R}}''|/\nu'', \tag{26b}$$

and ν' and ν'' are the effective degrees of freedom for, respectively, the local and remote weighted data.

The effective degrees of freedom are based on the following considerations. The maximum possible degrees of freedom for unweighted residuals that are assumed uncorrelated is $\nu_{\max} = 2J - 2KN$ for J frequencies, K magnetic components and N estimated complex terms. The weights will reduce any correlation that may exist between the residuals but the degrees of freedom will also be reduced because the residual spectrum is changed by the continuum weights and values are eliminated, i.e. one degree of freedom is lost for each $g(t) = 0$ and two degrees are lost for each $F(\omega) = 0$. The degrees of freedom will also be reduced by the time and frequency weights that are less than unity but greater than zero. The effective degrees of freedom are

therefore approximated here by

$$\nu = J^{-1} \sum_t g(t)^2 \sum_\omega F(\omega)^2 |\hat{W}(\omega)\hat{V}(\omega)|^2, \tag{27}$$

where the continuum weights have been rescaled to be less than unity, $|\hat{W}(\omega)\hat{V}(\omega)| \leq 1$.

4.5 Transfer estimates

The least-squares solution minimizes $|\mathbf{R}'^{\mathrm{H}} \cdot \mathbf{R}'|$ and $|\mathbf{R}''^{\mathrm{H}} \cdot \mathbf{R}''|$: the variance of the weighted residuals over the entire frequency range $j = 1, \ldots, J$. The solutions of (24a,b) are

$$\mathbf{u} = \mathbf{A} \cdot (\mathbf{p}'^{\mathrm{H}} \cdot \mathbf{E}') \tag{28a}$$

and

$$\underline{\mathbf{u}} = \underline{\mathbf{A}} \cdot (\underline{\mathbf{p}}''^{\mathrm{H}} \cdot \mathbf{E}''), \tag{28b}$$

where the $\mathbf{A}$'s are the matrix inverses given by

$$\mathbf{A} = (\mathbf{p}'^{\mathrm{H}} \cdot \mathbf{p}' + \lambda \mathbf{I})^{-1}$$

and

$$\underline{\mathbf{A}} = (\underline{\mathbf{p}}''^{\mathrm{H}} \cdot \underline{\mathbf{p}}'' + \underline{\lambda} \mathbf{I})^{-1}$$

that are Hermitian and λ and $\underline{\lambda}$ are damping factors.

These are the least squares solutions using the normal equations. They usually give inferior solutions due to roundoff errors when compared to other more accurate methods. Here, however, smooth solutions rather than the most accurate solutions are sought in order to prevent any possible oscillations. This is accomplished by a damping factor λ equal to 10^{-q} times the leading term of $\mathbf{p}'^{\mathrm{H}} \cdot \mathbf{p}'$ or $\underline{\mathbf{p}}''^{\mathrm{H}} \cdot \underline{\mathbf{p}}''$ where the smallest q from $q = 2, \ldots, 6$, is chosen so that the coherence between the weighted electric and predicted data is slightly less than 1 per cent of the coherence for the undamped solution. The motivation for using least squares with constraints is discussed by Twomey (1977, chapter 6).

The covariance of the solutions using (25a,b) yields, respectively,

$$\mathrm{Cov}\,(\mathbf{u} \cdot \mathbf{u}^{\mathrm{H}}) = \mathrm{Var}\,|R'|\, \mathbf{A} \cdot (\mathbf{p}'^{\mathrm{H}} \cdot \mathbf{p}') \cdot \mathbf{A} \tag{29a}$$

and

$$\mathrm{Cov}\,(\underline{\mathbf{u}} \cdot \underline{\mathbf{u}}^{\mathrm{H}}) = \mathrm{Var}\,|\underline{R}''|\, \underline{\mathbf{A}} \cdot (\underline{\mathbf{p}}''^{\mathrm{H}} \cdot \underline{\mathbf{p}}'') \cdot \underline{\mathbf{A}} \tag{29b}$$

Since the damping factor is small the approximations $\mathbf{A} \cdot (\mathbf{p}'^{\mathrm{H}} \cdot \mathbf{p}') \cdot \mathbf{A} \approx \mathbf{A}$ and $\underline{\mathbf{A}} \cdot (\underline{\mathbf{p}}''^{\mathrm{H}} \cdot \underline{\mathbf{p}}'') \cdot \underline{\mathbf{A}} \approx \underline{\mathbf{A}}$ can be used.

These least-squares estimates are biased by the noise in $\mathbf{p}'$ and $\underline{\mathbf{p}}''$. If remote reference data exist, this noise can be partially reduced by constructing averages between the local magnetic data and the projected magnetic data from the remote magnetic data (Appendix B). This average magnetic data is then used instead of the local magnetic data.

The remote reference solution minimizes $|\underline{\mathbf{R}}''^{\mathrm{H}} \cdot \mathbf{R}'|$, the modulus of the covariance between the weighted residuals, over the entire frequency range, $j = 1, \ldots, J$. The solutions of (24a,b) are

$$\mathbf{u} = \mathbf{C} \cdot (\underline{\mathbf{p}}''^{\mathrm{H}} \cdot \mathbf{E}') \tag{30a}$$

and

$$\underline{\mathbf{u}} = \mathbf{C}^{\mathrm{H}} \cdot (\mathbf{p}'^{\mathrm{H}} \cdot \mathbf{E}''), \tag{30b}$$

where $\mathbf{C}$ is the matrix inverse given by

$$\mathbf{C} = (\underline{\mathbf{p}}''^{\mathrm{H}} \cdot \mathbf{p}' + \lambda \mathbf{I})^{-1}$$

that is not Hermitian. The λ is a damping factor equal to 10^{-q} times the leading term $|\underline{\mathbf{p}}''^{\mathrm{H}} \cdot \mathbf{p}'|$ where q is found in the same way as for the least squares estimates. It is assumed here that this sort of damping can be applied even though the remote reference method is not least squares.

The covariance of the solutions using (25a,b) yields, respectively,

$$\mathrm{Cov}\,(\mathbf{u} \cdot \mathbf{u}^{\mathrm{H}}) = \mathrm{Var}\,|R'|\, \mathbf{C} \cdot (\underline{\mathbf{p}}''^{\mathrm{H}} \cdot \underline{\mathbf{p}}'') \cdot \mathbf{C}^{\mathrm{H}} \tag{31a}$$

and

$$\mathrm{Cov}\,(\underline{\mathbf{u}} \cdot \underline{\mathbf{u}}^{\mathrm{H}}) = \mathrm{Var}\,|\underline{R}''|\, \mathbf{C}^{\mathrm{H}} \cdot (\mathbf{p}'^{\mathrm{H}} \cdot \mathbf{p}') \cdot \mathbf{C}. \tag{31b}$$

The method of successive iterations used to obtain the transfer functions is outlined in Appendix E using frequency weights, time weights, 1-D transfer functions and smooth distortion functions. The ith estimate of the transfer function is then given by

$$Z_i(\omega) = \tilde{Z}_{i-1}(\omega) D_{i-1}(\omega) U_i(\omega) \tag{32}$$

where the $U_i(\omega)$ is generated by (22) using the least-squares solutions (28a,b) or the remote reference solutions (30a,b). The new 1-D model transfer function $\tilde{Z}_i(\omega)$ is generated by the D^+ inversion routine and the new distortion $D_i(\omega)$ is generated (Appendix D) from $\tilde{Z}_i(\omega)$, $Z_i(\omega)$ and its variance. The final transfer estimates are found when the difference between the perturbation function and unity is less than the confidence interval, the improvement in the coherence is less than a few per cent, and the weighted residuals in the frequency and time domain have an approximately normal distribution.

4.6 Confidence intervals

The perturbation function $U(\omega)$ represented by (22) is basically a regression curve and the estimates using (22) will not be indepenent. Various representations of the confidence interval for $U(\omega_j)$, $j = 1, \ldots, J$ estimates can be used that depend on the number of terms N, the number of estimates J, the degrees of freedom ν (27) and the $100(1-p)$ percentage level. In general, one uses a representation that gives the smallest intervals (Seber 1977). For the examples $J = 30$ and $N = 10$ and the appropriate representation suggested by Seber (1977) is the Bonferroni t-interval $\alpha_{\mathrm{B}}(J, \nu, p)$ that is based on the student t distribution. Some values for $\nu \gg 1$ and $p = 0.05$ are: $\alpha_{\mathrm{B}} = 1.96$ for $J = 1$; $\alpha_{\mathrm{B}} = 2.81$ for $J = 10$; $\alpha_{\mathrm{B}} = 3.02$ for $J = 20$; and $\alpha_{\mathrm{B}} = 3.15$ for $J = 30$. Thus the frequency resolution cannot be increased by merely generating more values of U because increasing J merely causes the confidence intervals to increase.

The confidence interval for the modulus of the perturbation function is then given by

$$\delta U(\omega) = \alpha_{\mathrm{B}}(J, \nu, p)(\mathrm{Var}\,|U(\omega)|)^{1/2} \tag{33}$$

for the $j = 1, \ldots, J$ estimates where the variance of the

modulus of the perturbation function is given by

$$\text{Var}\,|U(\omega)| = \left| \sum_{m=1}^{N} \sum_{n=1}^{n} \text{Cov}\,(u_n u_m) S_n(\omega) S_m(\omega)^* \right| \qquad (34)$$

using estimates of $\text{Cov}\,(u_n u_m)$ for the least-squares solutions (29a,b) or the remote reference solutions (31a,b). Since the factor α_B is independent of frequency any rescaling of the confidence interval for changes in J or p can easily be carried out at a later stage.

The electric data is expected to be noisier than the magnetic data. The assumption was made, therefore, in deriving the variance of the solutions, that the weighted electric field is noisy and the weighted predicted field is noise free. This makes the estimates of the variance tractable and implies that $\tilde{Z}_{i-1}$, D_{i-1} and the magnetic field are noise free. The effects of all the noise is therefore contained in $\text{Var}\,|U_i|$. The variance of the modulus of the ith transfer estimate (32) is then given by

$$\text{Var}\,|Z_i(\omega)| = |\tilde{Z}_{i-1}(\omega) D_{i-1}(\omega)|^2\, \text{Var}\,|U_i(\omega)| \qquad (35a)$$

The confidence interval is then

$$\delta Z_i(\omega) = |\tilde{Z}_{i-1}(\omega) D_{i-1}(\omega)|\, \delta U_i(\omega). \qquad (35b)$$

The validity of the assumption of a noise free magnetic field was examined using artificial data where a section of noise (8 per cent of the data) was added to the magnetic field. It was found for this example that the weighted predicted field was essentially free of this noise.

5 EVALUATION OF TRANSFER ESTIMATES

Let the noise contributed by the weighted electric data be $\boldsymbol{\Delta}_E$, the noise contributed by the weighted locally predicted data be $\boldsymbol{\Delta}_L$, and the noise contributed by the weighted remotely predicted data be $\boldsymbol{\Delta}_R$. The locally and remotely derived weighted residuals are derived from (17a,b) and are related to the noise by, respectively, $\mathbf{R}' = \boldsymbol{\Delta}_E + \boldsymbol{\Delta}_L$ and $\underline{\mathbf{R}}'' = \boldsymbol{\Delta}_E + \boldsymbol{\Delta}_R$, where it is assumed that the estimated transfer functions are relatively independent of the noise and that the solutions have converged: $\mathbf{U} \approx 1$ and $\underline{\mathbf{U}} \approx 1$.

The weighted residuals can be summarized by two parameters: one is the complex coherency between the locally and remotely derived weighted residuals

$$\Gamma = (\mathbf{R}'^{\mathrm{H}} \cdot \underline{\mathbf{R}}'')/(|\mathbf{R}'^{\mathrm{H}} \cdot \mathbf{R}'|\, |\underline{\mathbf{R}}''^{\mathrm{H}} \cdot \underline{\mathbf{R}}''|)^{1/2} \qquad (36a)$$

and the other is the ratio of the residual variance of the weighted residuals

$$\Lambda = |\underline{\mathbf{R}}''^{\mathrm{H}} \cdot \underline{\mathbf{R}}''|/|\mathbf{R}'^{\mathrm{H}} \cdot \mathbf{R}'| \qquad (36b)$$

If the noise in the weighted residual is due mainly to the electric data, then $\mathbf{R}' \approx \underline{\mathbf{R}}'' \approx \boldsymbol{\Delta}_E$ and one has $\mathcal{R}e\,\Gamma \approx 1$, $\mathcal{I}m\,\Gamma \approx 0$ and $\Lambda \approx 1$. If the noise is due mainly to the locally predicted data, then $\mathbf{R}' \approx \boldsymbol{\Delta}_L$ and $\underline{\mathbf{R}}'' \approx 0$ and one has $\mathcal{R}e\,\Gamma \approx 0$, $\mathcal{I}m\,\Gamma \approx 0$ and $\Lambda \approx 0$. If the noise is due mainly to the remotely predicted data, then $\mathbf{R}' \approx 0$ and $\underline{\mathbf{R}}'' \approx \boldsymbol{\Delta}_R$ and one has $\mathcal{R}e\,\Gamma \approx 0$, $\mathcal{I}m\,\Gamma \approx 0$ and $\Lambda \gg 1$. Thus the residual coherency Γ approaches unity for mainly noisy electric data and zero for mainly noisy predicted data and the residual variance ratio Λ approaches a large value for mainly noisy remotely predicted data and zero for mainly noisy locally predicted data. Note that the noise in the predicted data can be small because the magnetic noise is small or because the noise is suppressed by a small transfer function.

5.1 Necessary conditions on residuals for uncorrelated noise

The remote reference method is justified on the assumption that the noise between the data are uncorrelated. When this assumption is valid the weighted residuals in terms of the noise reduce to $|\mathbf{R}'^{\mathrm{H}} \cdot \mathbf{R}'| = |\boldsymbol{\Delta}_E^{\mathrm{H}} \cdot \boldsymbol{\Delta}_E| + |\boldsymbol{\Delta}_L^{\mathrm{H}} \cdot \boldsymbol{\Delta}_L|$, $|\underline{\mathbf{R}}''^{\mathrm{H}} \cdot \underline{\mathbf{R}}''| = |\boldsymbol{\Delta}_E^{\mathrm{H}} \cdot \boldsymbol{\Delta}_E| + |\boldsymbol{\Delta}_R^{\mathrm{H}} \cdot \boldsymbol{\Delta}_R|$ and $(\mathbf{R}'^{\mathrm{H}} \cdot \underline{\mathbf{R}}'') = |\boldsymbol{\Delta}_E^{\mathrm{H}} \cdot \boldsymbol{\Delta}_E|$. Then the residual coherency becomes

$$\mathcal{R}e\,\Gamma = (1 + \Lambda_L)^{-1/2}(1 + \Lambda_R)^{-1/2} \qquad (37a)$$

$$\mathcal{I}m\,\Gamma = 0 \qquad (37b)$$

and the variance ratio becomes

$$\Lambda = (1 + \Lambda_R)/(1 + \Lambda_L), \qquad (37c)$$

where

$$\Lambda_L = |\boldsymbol{\Delta}_L^{\mathrm{H}} \cdot \boldsymbol{\Delta}_L|/|\boldsymbol{\Delta}_E^{\mathrm{H}} \cdot \boldsymbol{\Delta}_E| \qquad (38a)$$

and

$$\Lambda_R = |\boldsymbol{\Delta}_R^{\mathrm{H}} \cdot \boldsymbol{\Delta}_R|/|\boldsymbol{\Delta}_E^{\mathrm{H}} \cdot \boldsymbol{\Delta}_E| \qquad (38b)$$

are the variance ratios of the noise contributed by, respectively, the locally and remotely predicted data relative to the noise contributed by the electric data.

Since the noise coherency and variance ratios are estimates, there will be an allowable range of estimates for which the assumption of uncorrelated noise will be valid. The allowable range for the complex residual coherency are approximated by

$$-\Gamma_N \le \mathcal{R}e\,\Gamma \qquad (39a)$$

$$-\Gamma_N \le \mathcal{I}m\,\Gamma \le \Gamma_N, \qquad (39b)$$

where

$$\Gamma_N(\nu, p) = (1 - p^{2/(\nu-2)})^{1/2}$$

is the $100(1-p)$ per cent level for the coherency of uncorrelated noise for ν degrees of freedom (Chave & Filloux 1985). The allowable range for the residual variance ratio, combining (37a) and (37c), and using the fact that $\Lambda_L \ge 0$ and $\Lambda_R \ge 0$, are

$$\mathcal{R}e\,\Gamma \le \Lambda^{1/2} \le 1/\mathcal{R}e\,\Gamma \qquad (39c)$$

The conditions (39a–c), using band averaged estimates of the variance and covariance of the residuals, establish the frequency range for which the noise can be treated as uncorrelated. When this occurs, the locally and remotely predicted variance ratios using equations (37a) and (37c) are given, respectively, by

$$\Lambda_L = (\Lambda^{-1/2}/\mathcal{R}e\,\Gamma) - 1 \qquad (40a)$$

and

$$\Lambda_R = (\Lambda^{1/2}/\mathcal{R}e\,\Gamma) - 1 \qquad (40b)$$

These are used to evaluate the least-squares and remote reference solutions. For example, the least-squares and remote reference methods will yield comparable results if

the noise in the predicted data is small, $\Lambda_L \approx \Lambda_R \ll 1$. If the noise in the locally predicted data is large, $\Lambda_L \gg \Lambda_R$, and the remote reference method will be superior. If the noise in the remotely predicted data is large, $\Lambda_L \ll \Lambda_R$, and the least-squares method will be superior. If the conditions for uncorrelated noise are not met, the remote reference method may yield inferior results.

5.2 Coherency

Another method for evaluating the transfer function is to compute the coherency between the weighted electric $\mathbf{E}'$ and predicted values $\mathbf{P}' = \mathbf{p}' \cdot \mathbf{u}$. It is given by

$$\gamma = (\mathbf{P}'^{H} \cdot \mathbf{E}')/(|\mathbf{P}'^{H} \cdot \mathbf{P}'|\,|\mathbf{E}'^{H} \cdot \mathbf{E}'|)^{1/2}$$

The least-squares solution requires that $\mathbf{P}'^{H} \cdot \mathbf{R}' = 0$ where $\mathbf{E}' = \mathbf{P}' + \mathbf{R}'$. The coherency therefore becomes the real valued term

$$\gamma = [|\mathbf{P}'^{H} \cdot \mathbf{P}'|/|\mathbf{E}'^{H} \cdot \mathbf{E}'|]^{1/2}$$

The remote reference solution requires that $\mathbf{P}'^{H} \cdot \underline{\mathbf{R}}'' = \underline{\mathbf{P}}''^{H} \cdot \mathbf{R}'^{H} = 0$. The coherency is therefore a complex valued term since $\mathbf{P}'^{H} \cdot \mathbf{R}'$ does not, in general, vanish. If, for example, $\mathbf{E}' = \mathbf{E}''$, then the coherency becomes $\underline{\gamma} = (\mathbf{P}'^{H} \cdot \underline{\mathbf{P}}'')/(|\mathbf{P}'^{H} \cdot \mathbf{P}'|\,|\mathbf{E}'^{H} \cdot \mathbf{E}'|)^{1/2}$. Then, since $|\mathbf{P}'^{H} \cdot \underline{\mathbf{P}}''| \leq (|\mathbf{P}'^{H} \cdot \mathbf{P}'|\,|\underline{\mathbf{P}}''^{H} \cdot \underline{\mathbf{P}}''|)^{1/2}$, one has $|\underline{\gamma}| < (|\underline{\mathbf{P}}''^{H} \cdot \underline{\mathbf{P}}''|/|\mathbf{E}'^{H} \cdot \mathbf{E}'|)^{1/2}$. If the remote data is very noisy, then $|\underline{\mathbf{R}}''^{H} \cdot \underline{\mathbf{R}}''| \gg |\mathbf{R}'^{H} \cdot \mathbf{R}'|$ and $|\underline{\mathbf{P}}''^{H} \cdot \underline{\mathbf{P}}''| \ll |\mathbf{P}'^{H} \cdot \mathbf{P}'|$. Thus $|\underline{\gamma}|$ can be appreciably smaller than γ and the remote reference method will then be inferior to the least-squares method.

6 EXAMPLES

The examples illustrate various situations that can arise. The first example uses an artificial electric field data generated from known magnetic data and transfer functions with normally distributed noise added to the electric and magnetic time series. The transfer estimates are then compared with the known using the various representations (21b–d) for the distortion and perturbation function. The second example uses cable voltages across the Florida Straits. These voltages, divided by the cable length, are compared with the distant continental magnetic stations at San Juan, Puerto Rico (1400 km eastwards) and Fredericksburg, Virginia (1000 km northwards). This example shows that remote magnetic stations can be used to remove most of the hourly geomagnetic induced variations in the cable voltages. The third example uses remote reference MT data from a site in the Tiwi geothermal field located on the east coast of Luzon, The Phillipines, with a remote site approximately 3 km distance. The example shows that the robust method works for data that has large narrow-band frequency noise. The fourth example uses magnetic data from distant sites, Tucson and Honolulu (5000 km separation) and shows that the remote reference method can be used on global scales.

For convenience, the transfer estimates are plotted as logarithms $L(\omega) = \ln[(i\mu_0/\omega)^{1/2} Z(\omega)]$ where $(-i\omega/\mu_0)^{1/2}$ is the transfer function for a $1\,S/m$ conducting half space. The confidence intervals are $\delta L(\omega) = \delta Z(\omega)/|Z(\omega)|$. The continuum weights W and V use $M = 10$ for all examples.

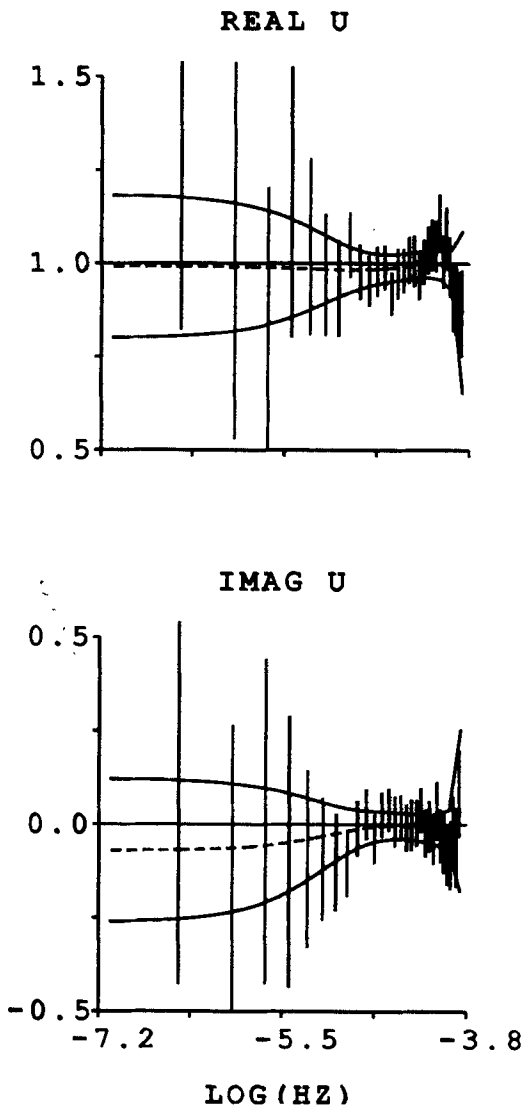

Figure 2. Perturbation function (dashed curves) for the Florida Straits hourly cable data by robust least squares. Pairs of smooth solid curves are the 95 per cent confidence intervals about the smooth solution for real (top panel) and imaginary (bottom panel) parts of the perturbation. Vertical lines are the 95 per cent confidence intervals for the robust band estimates. The figures show that the solution has converged and that the confidence intervals for the band averaged and smooth representation are in reasonably good agreement.

An example of the conversion of the solutions is given in Fig. 2 for the Florida Straits cable hourly data. It shows that the solution has converged after six iterations because the perturbation function $U(\omega) - 1$ lies well within the confidence interval for the band averaged and smooth estimates. The remaining examples discussed here give equally valid conversions within six to eight iterations.

6.1 Tiwi data for given transfer functions and noise

This data (12 288 values per series with 1/32 s between values) was chosen to illustrate the situation when the

transfer functions and noise are known. The electric data is generated from the observed Tiwi 16 hz magnetic data using known 1-D transfer functions. The noise (Fig. 3) is normally distributed and added to the electric and magnetic data with a shorter section (1000 values, about 8 per cent of series) having variance 10 times larger. The time and frequency weights are shown in Fig. 4 where the time weights clearly show the short noisy sections have been downweighted except for the noisy section in the magnetic north component. This component has a partial coherence squared with the electric data of only 0.21, and therefore plays a minor role compared to the dominant east magnetic component that has a partial coherence squared of 0.78. Thus the weights have the ability to reduce the noise in both the electric and magnetic data. This increases the coherence from 0.80 to 0.88 for the final iteration.

The least-squares robust transfer estimates are compared with the known in Fig. 5 for the dominant east magnetic component using the distortion represented by square root frequency (21d) and by exponentials (21c) and in Fig. 5 for the north and east magnetic components using the distortion represented by frequency (21b). The figures clearly show that the distortion represented by (21b) and (21d) give equally good fits to the known transfer function. The (east, north) transfer estimates in Fig. 6 have biases of $(-4.9 \pm 2.0,\ -6.0 \pm 0.8)$ per cent due to the noise in the magnetic field and phase shifts of $(-0.4 \pm 1.2,\ -0.7 \pm 0.5)$ degrees relative to the known transfer function. The main differences are due to the estimates of the confidence intervals for the low frequency portion of the spectra where the data are poorly correlated. The factor $\beta = 1.5$ is probably not the optimum value for the exponential representation.

The favorable comparison with the known transfer functions shows that the least-squares estimates are only slightly biased downwards and that they are therefore good estimates of the known transfer functions. The bias is found

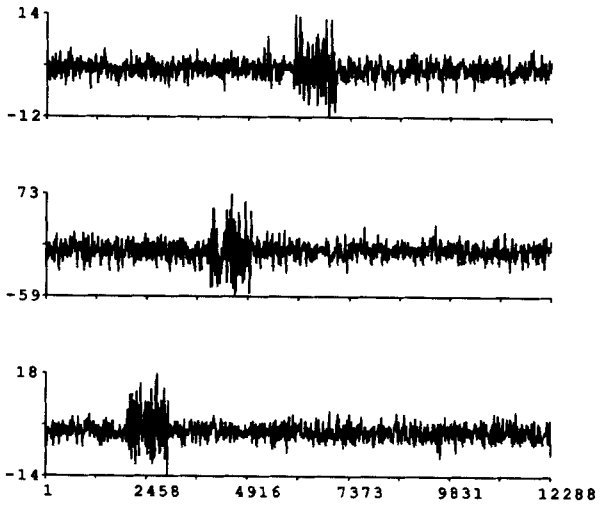

Figure 3. Gaussian noise in the electric (bottom panel), north magnetic component (middle panel) and east magnetic component (top panel) for the artificial example. The series are plotted as averages over 12 values. The vertical scale is arbitrary.

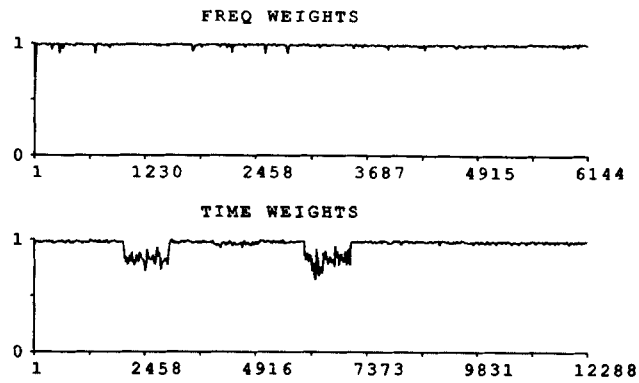

Figure 4. Frequency weights (top panels) and time weights (bottom panel) squared and averaged over 25 values for the artificial example. The horizontal axes represents Fourier frequency terms (top panel) and time values (bottom panel).

to be frequency independent which means that the bias effect ultimately ends up as a small bias in the static distortion effect. The removal of this constant bias shows that the shape of the transfer functions for the various estimates are in complete agreement with the known transfer function. The representation of the distortion in terms of a power series in frequency is used for all of the remaining examples.

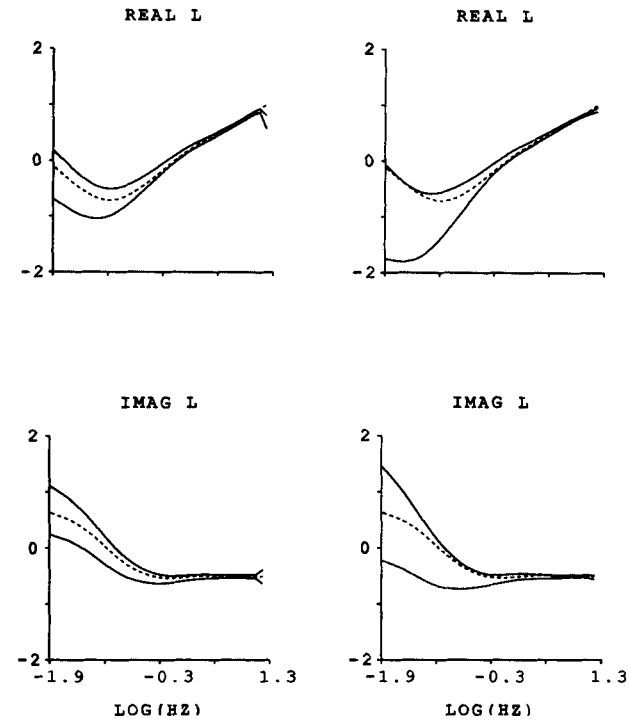

Figure 5. Logarithm of the least-squares transfer estimates for the east magnetic component of the artificial data using an expansion in square root frequency (left panels) and in exponentials with $\beta = 1.5$ (right panels) where pairs of curves are 95 per cent confidence intervals for real (top panels) and imaginary (bottom panels) parts. The dashed curves are the known transfer function.

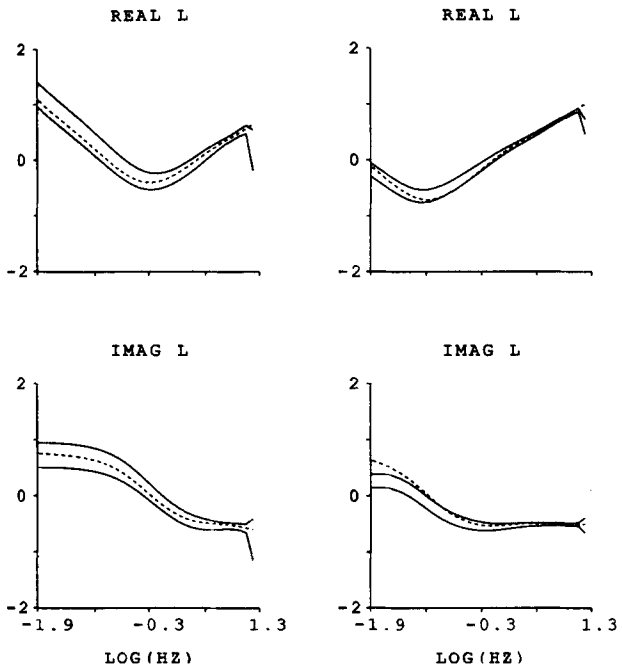

Figure 6. Logarithm of the least-squares estimates for the north (left panels) and the east (right panels) magnetic components of the artificial data using an expansion in frequency where pairs of curves are 95 per cent confidence intervals for real (top panels) and imaginary (bottom panels) parts. Dashed curves are the known transfer function. The figures show that the estimates are in good agreement with the known transfer functions.

6.2 Florida Straits cable data

This data (17 520 hourly mean values per series) consists of voltages recorded across the Florida Straits using an abandoned telephone cable and illustrates how measurements of the motionally induced voltages by the Florida Current can be improved by removing the geomagnetic induced voltages. These are determined from the estimates of the transfer functions and the simultaneous recordings of the north magnetic data from the magnetic observatories at San Juan, Puerto Rico and Fredericksburg, Virginia. The data were preprocessed by removing the trend and its variation by a least-squares cubic spline having five equally spaced knots and removing the tidal and solar diurnal variation and its harmonics by least squares estimates of their amplitude and phase.

The time and frequency weights are shown in Fig. 7. The per cent reduction in degrees of freedom for (continuum, frequency, time) weights are (39, 4.9, 9.3). There is an additional 3.6 per cent reduction due to missing data. The principal reduction in degrees of freedom are therefore caused by the continuum weights that flatten the low frequency noise in the residual spectrum followed by time weights. The weights caused the coherence to increase from 0.73 to 0.86 for the final iteration. Thus the time and frequency weights have substantially increased the coherence and the reliability of the transfer estimates.

The distributions of the residuals divided by robust estimates of the standard deviation are plotted in Fig. 8 for

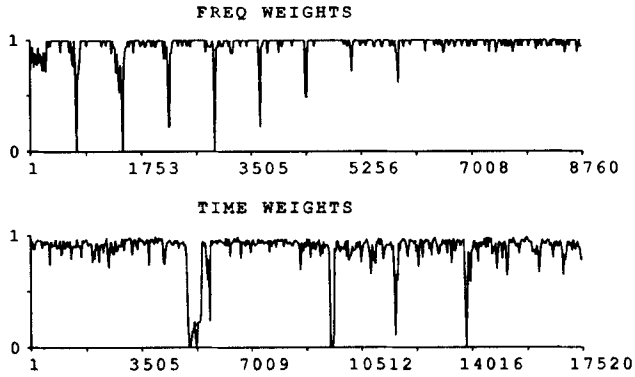

Figure 7. Frequency weights (top panel) and time weights (bottom panel) squared and averaged over 35 values for the Florida Straits hourly cable data. The horizontal axes are Fourier frequency terms (top panel) and time values (bottom panel).

the unweighted and weighted frequency and time residuals where the distribution of the real and imaginary parts of the frequency residuals are combined. The figure shows that the weighted residuals have a nearly normal distribution while the unweighted residuals, especially the frequency residuals, do not. The χ^2 misfit to a normal distribution divided by the 95 per cent confidence level are (497, 1.0) for the (unweighted, weighted) frequency residuals and are (11, 1.0) for the time residuals. The weights have therefore accomplished the desired effect of making the weighted residuals normally distributed.

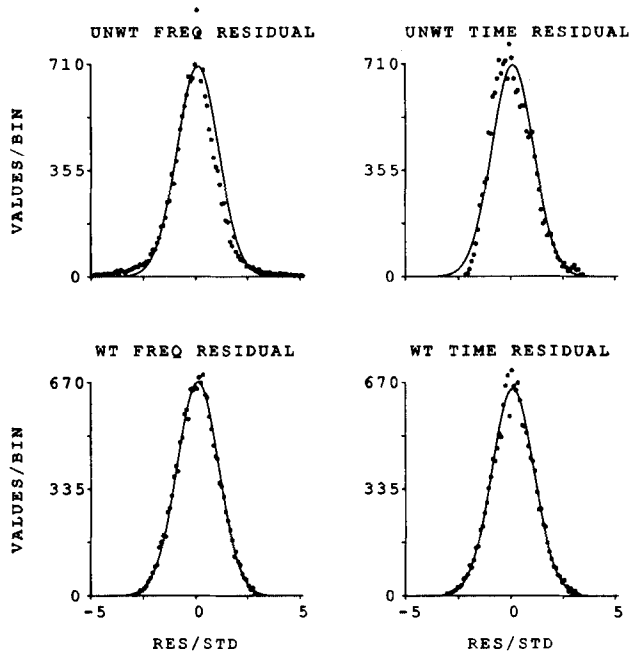

Figure 8. Distribution of residuals for the Florida Straits hourly cable data. Horizontal axes are the residuals divided by robust estimates of the standard deviation. Dots are the number of values per 0.1 binwidths and solid curves are the normal distributions for: unweighted frequency distribution (top left panel), unweighted time distribution (top right panel), weighted frequency distribution (bottom left panel) and weighted time distribution (bottom right panel). The figures show that the weighted residuals are normally distributed.

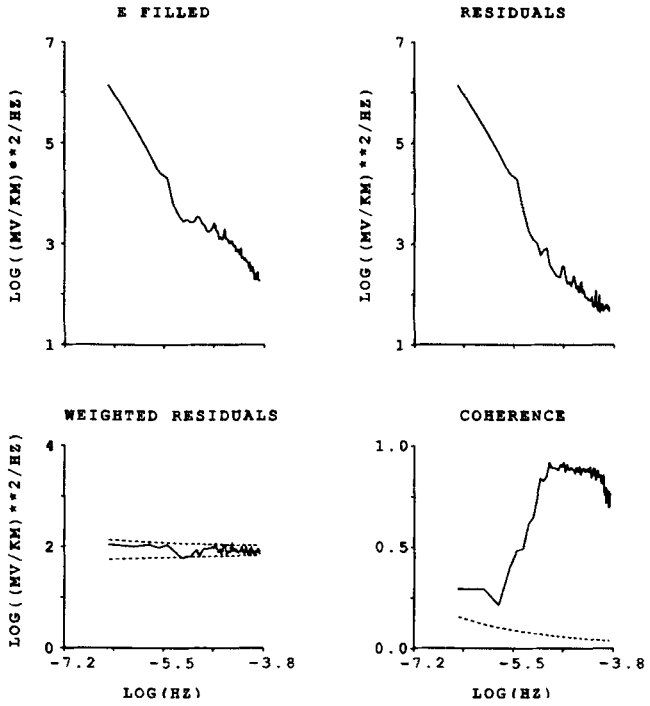

Figure 9. Spectral estimates of the electric data (top left panel), residuals (top right panel), weighted residuals (bottom left panel) and coherence (bottom right panel) for the Florida Straits hourly cable data. The 95 per cent confidence intervals for all spectra are plotted (dashed curves) in the lower left panel. The coherence at the 95 per cent level for uncorrelated noise is plotted (dashed curve) in the lower right panel.

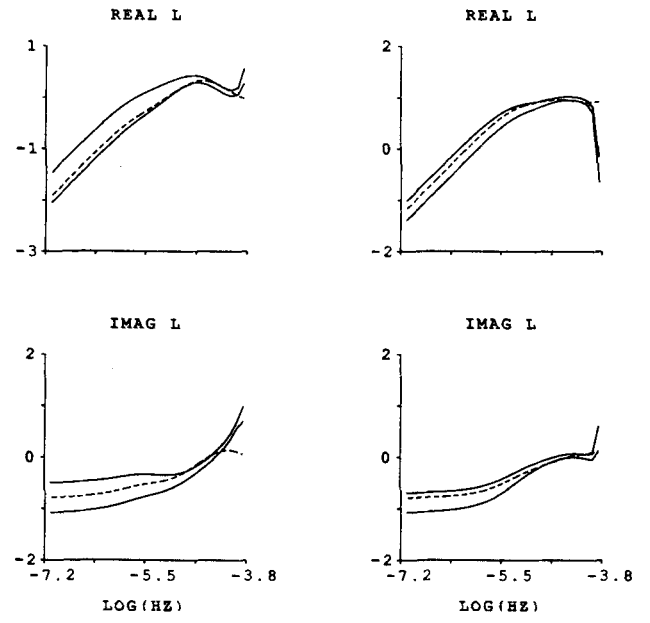

Figure 10. Logarithm of transfer estimates for the Florida Straits hourly cable voltages divided by cable length using the north magnetic components from San Juan (left panels) and Fredericksburg (right panels) where pairs of curves are the 95 per cent confidence intervals for the real (top panels) and the imaginary (bottom panels) parts. Dashed curves are the estimated 1-D model transfer functions.

Spectra of the electric data, the residuals, the weighted residuals and the coherence are given in Fig. 9 where the voltages have been converted into electric field data by dividing the voltage by the cable length of 99 km. The residual spectrum is dominated by a large increase in the continuum towards the low frequencies. This is caused by variations in the transport of the Florida Current so that the coherence only exceeds 0.6 for frequencies greater than 0.4 cpd. The weighted residual spectrum is, however, essentially flat with variations lying well within the 95 per cent confidence interval. This shows that the weighting has made the residuals independent. A comparison between the residual and electric spectra in Fig. 9 shows that the removal of the geomagnetic induced variations has caused a substantial reduction in the variance of the electric data for frequencies greater than 0.4 cpd. This means that diurnal and semi-diurnal tidal induced signals can be accurately determined.

The least-squares robust estimates of the transfer functions for the San Juan and Fredericksburg north magnetic components are shown in Fig. 10 and are found to approximate a 1-D transfer functions for the low frequencies.

6.3 Tiwi MT data

This data (12 288 values per series with 1/256 s between values) illustrates the situation of remote reference MT data where the frequency domain noise exceeds the time domain noise. The frequency noise is found to be narrow-band and appears to be caused by the 60 Hz power grid and harmonics and subharmonics of the approximately 48 Hz power generator. The largest peak is centered at 47.9 Hz followed by a lesser peak at 60 Hz and smaller peaks at 23.9, 95.8 and 112.1 Hz. The remote reference data shows the same dominant peaks with an extra peak at 53.8 Hz and no peak at 23.9 Hz.

The time and frequency weights, shown in Fig. 11 for the least-squares estimates, clearly show the large amount of noise in a few narrow frequency bands. The per cent reduction in degrees of freedom for (continuum, frequency,

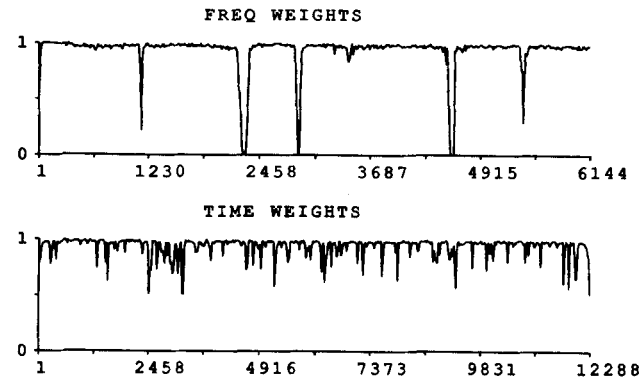

Figure 11. Frequency weights (top panel) and time weights (bottom panel) squared and averaged over 25 values for the Tiwi 128 Hz data using least squares. The horizontal axes are Fourier frequency terms (top panel) and time values (bottom panel). The figures show that the residuals are dominated by large noise in the frequency domain.

time) weights are (60, 7.5, 6.5) for least squares and (67, 6.1, 6.7) for remote reference. The principal reduction in degrees of freedom are caused by the continuum weights that flatten the low and high frequency noise of the residual spectrum and by the frequency weights that eliminate the large spectral peaks. The weights for the (least-squares, remote reference) estimates increased the coherence from the initial values of (0.77, 0.55) to the remarkably large values of (0.98, 0.98) for the final iteration.

The distributions for the least-squares estimates of the residuals divided by robust estimates of the standard deviation are plotted in Fig. 12 for the unweighted and weighted frequency and time residuals. They show that the weighted residuals have a nearly normal distribution whereas the unweighted residuals clearly do not. The χ^2 misfit to a normal distribution divided by the 95 per cent confidence level are (2251, 1.0) for the (unweighted, weighted) least squares frequency residuals, (28, 0.9) for the least-squares time residuals, (3225, 1.1) for the remote reference frequency residuals and (29, 0.8) for the remote reference time residuals. The small misfits for the weighted residuals show that the weighting has accomplished the desired effect of making the residuals normally distributed.

Spectra of the electric data, the residuals, the weighted residuals and the coherence are given in Fig. 13. The residual spectrum is dominated by large spectral peaks but the weighted residual spectra show that these have been removed and that the variations lie mostly within the 95 per cent confidence interval. The coherence is close to unity over most of the frequencies and this shows that the

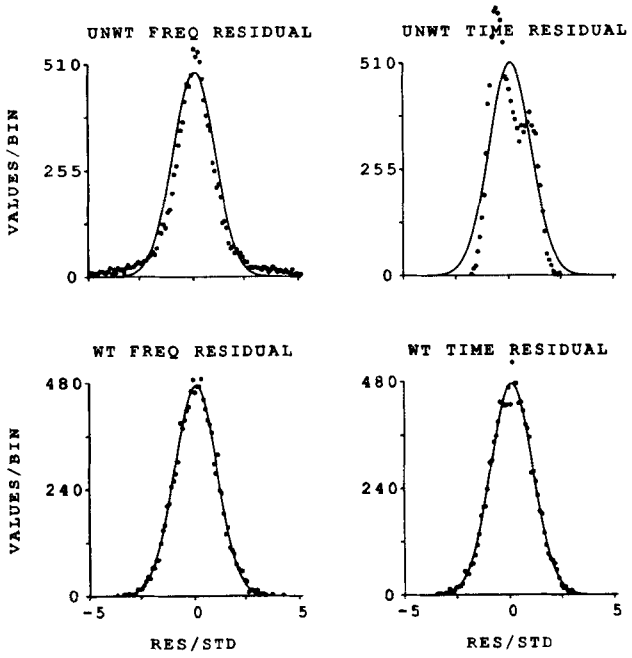

Figure 12. Distribution of residuals for the Tiwi 128 Hz data. Horizontal axes are the residuals divided by robust estimates of the standard deviation. Dots are the number of values per 0.1 binwidths and solid curves are the normal distributions for: unweighted frequency residuals (top left panel), unweighted time residuals (top right panel), weighted frequency residuals (bottom left panel) and weighted time residuals (bottom right panel). The figures show that the weighted residuals are normally distributed.

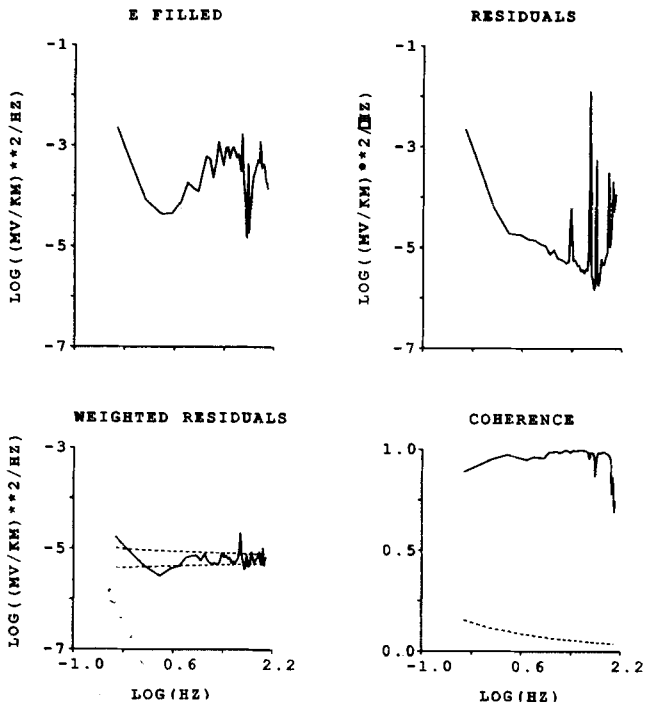

Figure 13. Spectral estimates of the electric data (top left panel), residuals (top right panel), weighted residuals (bottom left panel) and coherence (bottom right panel) for the Tiwi 128 Hz data. The 95 per cent confidence intervals for all spectra are plotted (dashed curves) in the lower left panel. The coherence at the 95 per cent level for uncorrelated noise is plotted (dashed curve) in the lower right panel. The residual spectrum shows the large spectral lines due to the power generator and power grid.

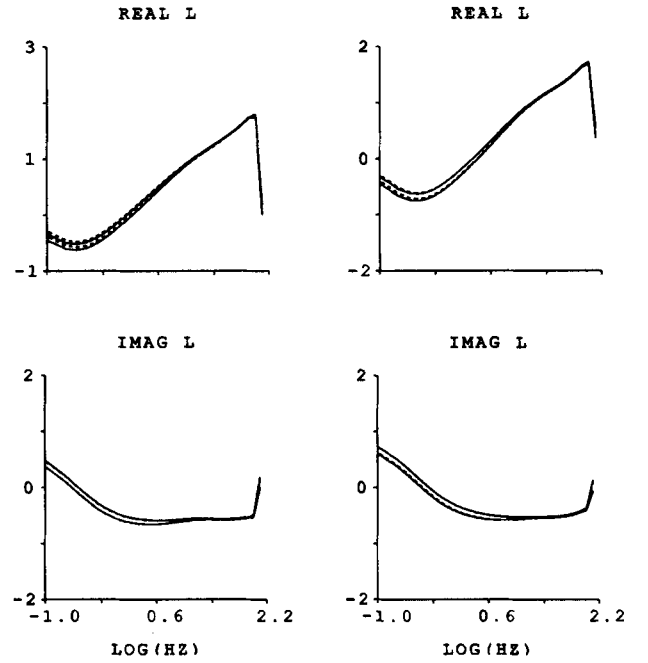

Figure 14. Logarithm of transfer estimates for the Tiwi 128 Hz data using least squares (solid curves) and remote reference (dashed curves) for the local (left panels) and remote (right panels) east magnetic components where pairs of curves are the 95 per cent confidence intervals about for the real (top panels) and the imaginary (bottom panels) parts.

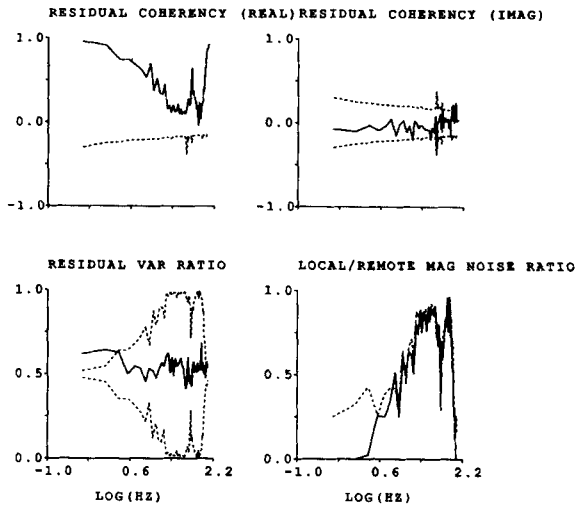

Figure 15. Uncorrelated noise test for the Tiwi 128 Hz data. Top left panel: real part of the residual coherency (solid curve) lying above dashed curve permits an uncorrelated noise interpretation. Upper right panel: imaginary part of residual coherency (solid curve) lying between dashed curves permits an uncorrelated noise interpretation. Lower left panel: residual variance ratio (solid curve) lying between dashed curves permits an uncorrelated noise interpretation. Lower right panel: normalized local magnetic noise ratio $\Lambda_L/(1+\Lambda_L)$ (solid curve) and normalized remote magnetic noise ratio $\Lambda_R/(1+\Lambda_R)$ (dashed curve). The figures show that the noise is essentially uncorrelated for most of the frequencies and is dominated by magnetic noise for frequencies greater than 10 Hz.

elimination of the residual spectral peaks has only reduced the coherence for a few narrow bands.

The least-squares and remote reference robust transfer estimates for the north and east magnetic components are compared in Fig. 14 where the least-squares transfer estimates for the (local, remote) east magnetic components have biases of $(-1.7 \pm 0.4,\ -1.2 \pm 0.5)$ per cent and phase shifts of $(0.0 \pm 0.2,\ -0.1 \pm 0.3)$ relative to the remote reference estimates. The biases are found to be frequency independent and the least-squares and remote reference estimates therefore agree in shape.

The assumption of uncorrelated noise for the remote reference estimates is evaluated in Fig. 15. It shows that the conditions for uncorrelated noise is satisfied over most of the frequency range. The remotely predicted data therefore contributes almost none of the noise for frequencies less than 10 Hz and the locally and remotely predicted data contribute 9 times the variance of the electric data noise for frequencies between 10 and 100 Hz. This violates the assumption that the predicted fields are noise free and makes the estimates of the confidence limits questionable. The remote reference estimates will be superior to the least-squares estimates for the frequencies between 10 and 100 Hz. Since the overall coherence squared is nearly unity, however, the differences in the least-squares and remote reference estimates will only be minor.

6.4 Tuscon magnetic data

The vertical and north magnetic data for this example come from the Tucson magnetic observatory, Arizona with the remote site at the Honolulu magnetic observatory, Hawaii. The data (4018 values per series consisting of 2-day mean values) was chosen to show that remote reference methods can be used for globally separated sites and to examine the situation when the data are dominated by time domain noise. The transfer function for the vertical magnetic data are displayed as a MT transfer function by assuming the source field has a P_1^0 zonal description (Schultz & Larsen 1987). The data were preprocessed by removing the trend and its variation by a least-squares cubic spline having eight equally spaced knots and the annual and semi-annual variation by least-squares estimates of their amplitude and phase.

The time and frequency weights, shown in Fig. 16 for the least-squares estimates, clearly show the large amount of noise in the time domain. The low values of the frequency weights at the lowest frequencies are due to the rejection of the annual and semi-annual variation. The per cent reduction in degrees of freedom for (continuum, frequency, time) weights are: (37, 3.0, 7.2) for least squares; (35, 2.1, 6.1) for least-squares average magnetic data; (3.6, 3.5, 8.3) for remote reference. There is an additional 2.3 per cent reduction due to missing values. The principal reduction in degrees of freedom is caused by the continuum weights that flatten the low frequency noise in the residual spectrum followed by the time weights. The average magnetic data is the average of the Tucson north magnetic data and the Honolulu north magnetic data projected to Tucson. The weights cause the (least squares, least squares average magnetic, remote reference) estimates of the coherence to increase from (0.64, 0.65, 0.64) to (0.78, 0.79, 0.79) for the final iteration. The χ^2 misfit to a normal distribution divided by the 95 per cent confidence level are (221, 0.8) for the (unweighted, weighted) least-squares frequency residuals; (2.1, 0.7) for the least-squares time residuals, (357, 0.7) for the least-squares frequency residuals using the average magnetic data; (2.1, 0.7) for the least-squares time residuals using the average magnetic data; (290, 0.6) for the remote reference frequency residuals, and (2.2, 0.6) for the remote reference time residuals. These misfits show that weighting has

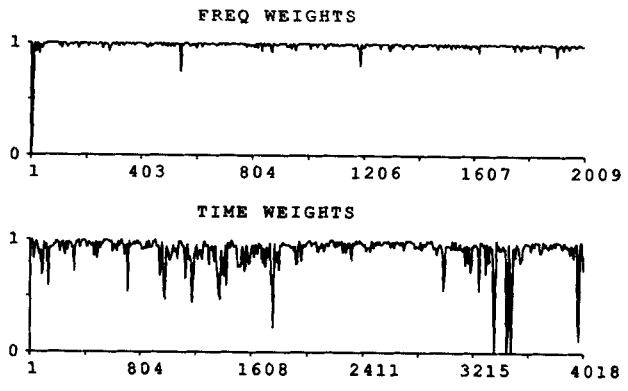

Figure 16. Frequency weights (top panel) and time weights (bottom panel) squared and averaged over eight values for the Tucson 2-day mean magnetic data using least squares. Horizontal axes are Fourier frequency terms (top panel) and time values (bottom panel). The figures show that the residuals are dominated by large noise in the time domain.

accomplished the desired effect of making the weighted residuals normally distributed.

Spectra of the electric data, the residuals, the weighted residuals and the total coherence are given in Fig. 17. The residual spectrum is dominated by a large smooth increase of the continuum towards the low frequencies so that the coherence only exceeds 0.6 for frequencies greater than 1.6 cpy. The weighted residual spectrum, however, is essentially flat with variations well within the 95 per cent confidence interval.

The transfer estimates by least squares using the Tucson magnetic component and the average magnetic data are compared in Fig. 18 with the remote reference estimates. The least squares transfer estimates for (local, average) magnetic component have biases of $(-4.7 \pm 1.8, -0.8 \pm 1.8)$ per cent and phase shifts of $(0.1 \pm 1.0, -0.3 \pm 1.0)$ relative to the remote reference estimates. The least-squares average estimate has no bias showing that averaging the local and remotely projected magnetic fields has eliminated the bias. The bias for the least-squares local estimate is found to be frequency independent. There is, therefore, no difference between the shapes of the three estimates of the transfer function.

The assumption of uncorrelated noise using the remote reference estimates is evaluated in Fig. 19. It shows that the conditions for uncorrelated noise is satisfied for all frequencies. The variance ratios show that the least-squares and remote reference estimates should give comparable results because the noise of the local and remote north

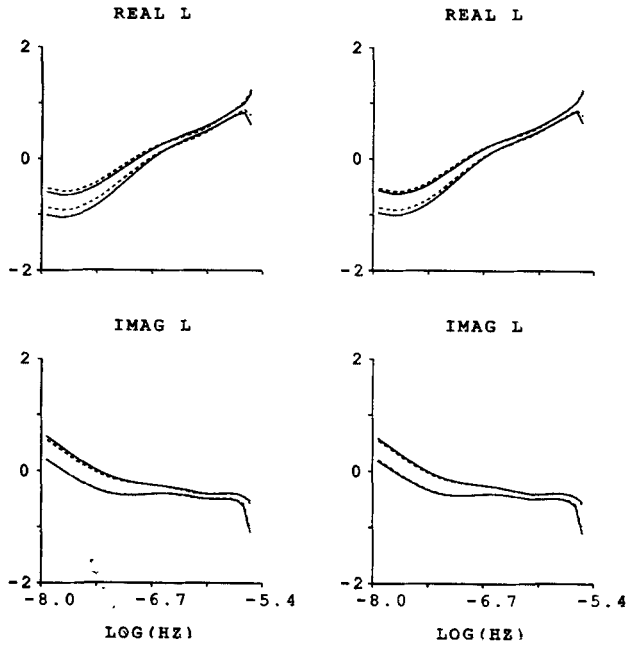

Figure 18. Logarithm of local transfer estimates for the Tucson 2-day mean magnetic data using least squares (solid curves) and remote reference (dashed curves) for the north magnetic component (left panel) and average north magnetic component between Tucson and Honolulu (right panel) where pairs of curves are 95 per cent confidence intervals about estimates for real (upper panels) and imaginary (lower panels) parts.

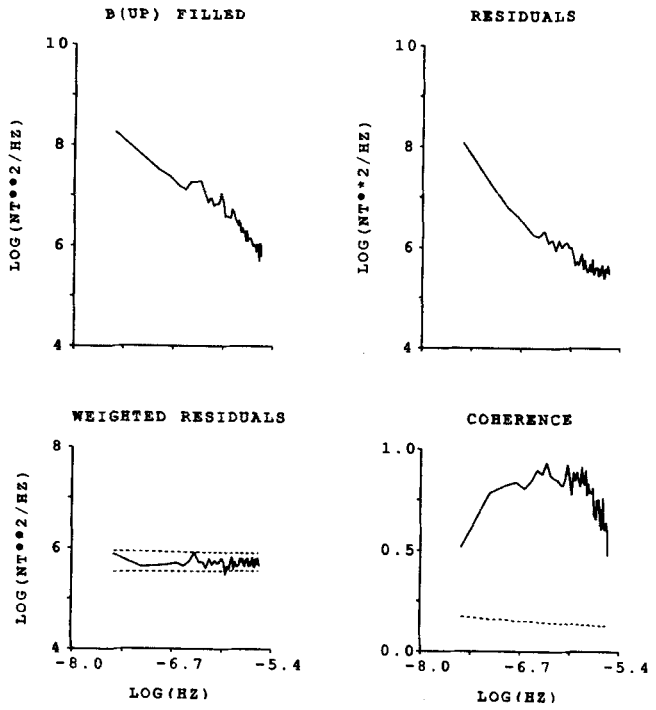

Figure 17. Spectral estimates of the vertical magnetic data (top left panel), residuals (top right panel), weighted residuals (bottom left panel) and coherence (bottom right panel) for Tucson 2-day mean magnetic data. The 95 per cent confidence intervals for all spectra are plotted (dashed curves) in the lower left panel. The coherence at the 95 per cent level for uncorrelated noise is plotted (dashed curve) in the lower right panel.

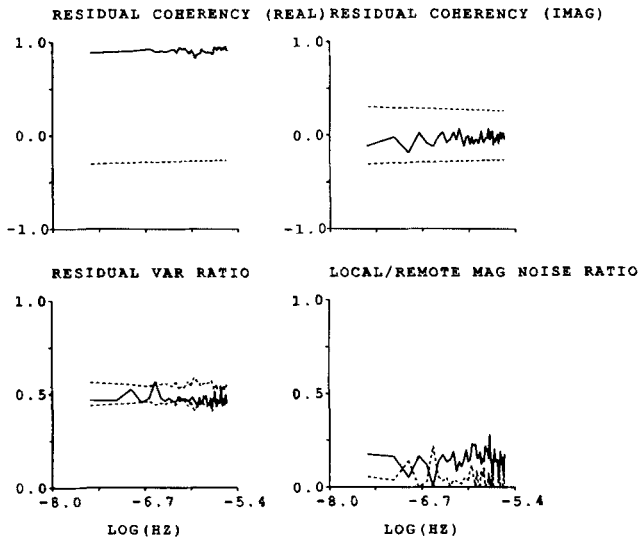

Figure 19. Uncorrelated noise test for the Tucson 2-day mean data. Upper left panel: real part of the residual coherency (solid curve) lying above dashed curve permits an uncorrelated noise interpretation. Upper right panel: imaginary part of the residual coherency (solid curve) lying between dashed curves permits an uncorrelated noise interpretation. Lower left panel: residual variance ratio (solid curve) lying between dashed curves permits an uncorrelated noise interpretation. Lower right panel: normalized local magnetic noise ratio $\Lambda_L/(1+\Lambda_L)$ (solid curve) and normalized remote magnetic noise ratio $\Lambda_R/(1+\Lambda_R)$ (dashed curve). The figures show that the noise is uncorrelated for all frequencies and that the noise is much smaller for the horizontal component than for the vertical magnetic component.

horizontal component are about 0.3 times the variance of the vertical magnetic noise. The remote reference estimates are slightly superior to the least-square estimates because the remote magnetic noise is about 20 per cent smaller than the local magnetic noise. These noise levels confirm the traditional assumption that long period vertical magnetic data are noisier than the horizontal magnetic data.

7 CONCLUSIONS

The improvements of the present method compared to other methods are: (i) Individual outliers in the frequency and time domain are removed using frequency and time weights. The examples used here show that the number of outliers is a small percentage of the total but that there are usually no sections of data free from time outliers. (ii) The use of frequency and time weights substantially increases the coherence for both the least-squares and remote reference methods, makes the tails of the distribution for the time and frequency residuals consistent with the normal distribution found for the bulk of the residuals and makes the residual spectrum white. Confidence limits can then be readily constructed because the weighted residuals are found to be approximately independent and identically and normally distributed. (iii) The analytic representation of the transfer function and its variance makes it possible to easily generate the predicted fields for any length of time series.

The important difference between the least-squares and remote reference methods are: (i) The remote reference method requires the existence of both a local and remote transfer function whereas the least-squares method only requires the existence of a local transfer function. Note that a transfer function will not exist, i.e. be a unique function of frequency if there is, for example, a large variable wavenumber effect. (ii) The remote reference method requires uncorrelated noise whereas the least-squares method only requires the noise to be small.

The problem of noise in both the electric and magnetic fields was not solved but the artificial example indicates that the weighting can reduce the effects of large noise in the magnetic field. The structure of the noise can be examined by frequency bands by comparing the locally and remotely derived residuals. Criterion have therefore been developed to test the assumption of uncorrelated noise. If the noise is deemed uncorrelated, then estimates of the relative amounts of noise contributed by the electric data, local magnetic data and remote magnetic data can be obtained and used to establish whether the remote reference estimates or the least-squares estimates are superior and whether it is justified to assume that the magnetic data are noise free.

ACKNOWLEDGMENTS

I thank Ulrich Schmucker, Peter Weidelt, Alan Chave, Ned Cokelet and Cathy Constable for helpful discussions, Ansley Manke for computing assistance, Jens Pedersen for lending me the remote reference Tiwi data and the Greens Scholarship and UNOCAL for initial support during my tenure as a Greens Scholar at the University of California, San Diego and NOAA for its continuing support.

REFERENCES

Chave, A. D. & Filloux, J. H., 1985. Observation and interpretation of the seafloor vertical electric field in the eastern North Pacific, *Geophys. Res. Lett.*, **12,** 793–796.

Chave, A. D., Thompson, D. J., 1989. Some comments on magnetotelluric response function estimation, *J. Geophys. Res.*, in press.

Chave, A. D., Thompson, D. J., & Ander, M. E., 1987. On the robust estimation of power spectra, coherence and transfer functions. *J. geophys. Res.*, **92,** 633–648.

Clarke, J., Gamble, T. D., Goubau, W. M., Koch, R. H., & Miracky, R. F., 1983. Remote-reference magnetotellurics: equipment and procedures, *Geophys. Prosp.*, **31,** 149–170.

Egbert, G. D. & Booker, J. R., 1986. Robust estimation of geomagnetic transfer functions, *Geophys. J. R. astr. Soc.*, **87,** 173–194.

Fisher, R. A., 1929. Tests of significance in harmonic analysis, *Proc. R. Soc. London A*, **125,** 54–59.

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

Hogg, R. V., 1979. An introduction to robust estimation, in *Robustness in Statistics*, pp. 1–17, eds Launer & Wilkinson, Academic Press, New York.

Holland, P. W. & Welsch, R. E., 1977. Robust regression using iteratively reweighted least-squares, *Commun. Stat. Theor. Meth.*, **A6,** 813–827.

Huber, P. J., 1964. Robust estimation of a location parameter, *Ann. math. Statist.*, **35,** 73–101.

Larsen, J. C., 1975. Low frequency (0.1–6.0 cpd) electromagnetic study of deep mantle electrical conductivity beneath the Hawaiian Islands, *Geophys. J. R. astr. Soc.*, **43,** 17–46.

Larsen, J. C., 1980. Electromagnetic response functions from interrupted and noisy data, *J. Geomag. Geolectr.*, **32 SI,** 89–103.

Larsen, J. C., 1981. A new technique for layered earth magnetotelluric inversion, *Geophysics*, **46,** 1247–1257.

Larsen, J. C. & Sanford, T. B., 1985. Florida Current volume transport from voltage measurements, *Science*, **227,** 302–304.

Munk, W. H. & Cartwright, D. E., 1966. Tidal spectroscopy and prediction, *Phil. Trans. R. Soc. London A*, **259,** 533–581.

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

Parker, R. L. & Whaler, K. A., 1981. Numerical methods for establishing solutions to the inverse problem of electromagnetic induction, *J. geophys. Res.*, **86,** 9574–9584.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T., 1986. *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press, Cambridge.

Schultz, A. & Larsen, J. C., 1987. On the electrical conductivity of the mid-mantle—I. Calculation of equivalent scalar magnetotelluric response functions, *Geophys. J. R. astr. Soc.*, **88,** 733–761.

Seber, G. A. F., 1977. *Linear Regression Analysis*, John Wiley, New York.

Sims, W. E., Bostick, F. X., Jr. & Smith, H. W., 1971. The estimation of magnetotelluric impedance tensor elements from measured data, *Geophysics*, **36,** 938–942.

Thompson, D., 1977. Spectrum estimation techniques for characterization and development of WT4 Wave Guide-1, *Bell System Tech. J.*, **56,** 1769–1815.

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

Vozoff, K., 1972. The magnetotelluric method in the exploration of sedimentary basins, *Geophysics*, **37**, 98–141.

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

Weidelt, P., 1986. Discrete frequency inequalities for magnetotelluric impedances of one-dimensional conductors, *J. Geophys.*, **59**, 171–176.

APPENDIX A: WEIDELT 1-D TRANSFER CONSTRAINTS WITH NOISE

Weidelt (1986) derived inequality constraints $W(\omega) \geq 0$ on the transfer function that are necessary and sufficient conditions for the transfer function to represent a 1-D conductivity model. If these inequalities are not satisfied then the conductivity model must have a 2-D or 3-D structure. The 1-D Weidelt constraints are given as functions of the induction length $c(\omega)$ that is defined in terms of the MT transfer functions by $c(\omega) = g(\omega) + ih(\omega) = iZ(\omega)/\omega$. Since the c's are estimates, the 1-D Weidelt constraints based on these estimates are $\hat{W}(\omega)$. These have confidence interval $\delta\hat{W}(\omega)$ that can be derived from the confidence interval $\delta(\omega)$ for $c(\omega)$ when the noise is assumed uncorrelated between estimates and the same for the real and imaginary parts of $c(\omega)$. Then the Weidelt constraints are bounded by $\hat{W}(\omega) - \delta\hat{W}(\omega) \leq W(\omega) \leq \hat{W}(\omega) + \delta\hat{W}(\omega)$. If $\hat{W}(\omega) + \delta\hat{W}(\omega) < 0$, then $W(\omega) < 0$ and the Weidelt constraints are not satisfied. The frequencies for which $\hat{W}(\omega) \geq -\delta\hat{W}(\omega)$ are therefore used to establish the frequencies for which a 1-D interpretation is possible. Let $c_1 = c(\omega_1)$, $c_2 = c(\omega_2)$, $g_1 = g(\omega_1)$, $g_2 = g(\omega_2)$, $h_1 = h(\omega_1)$, $h_2 = h(\omega_2)$, $\delta_1 = \delta(\omega_1)$, $\delta_2 = \delta(\omega_2)$, $r = \omega_2/\omega_1$ and $s = r(1 - 1/r)^2$. The 1-D Weidelt constraints (1986) for single and adjacent pairs with noise are

$$g_1 \,\&\, h_1 > -\delta_1 \tag{A.1}$$

$$g_2 \,\&\, h_2 > -\delta_2 \tag{A.2}$$

$$r^2 g_2 - g_1 > -(\delta_1^2 + r^4\delta_2^2)^{1/2} \tag{A.3}$$

$$rh_2 - h_1 > -(\delta_1^2 + r^2\delta_2^2)^{1/2} \tag{A.4}$$

$$rsg_1 g_2 - |c_1 - rc_2|^2 > -a \tag{A.5}$$

$$sh_1 h_2 - |c_1 - c_2|^2 > -b \tag{A.6}$$

for

$$a^2 = 4(|c_1 - rc_2|^2 - rsg_1 g_2)(\delta_1^2 + r^2\delta_2^2) + r^2 p(g_2^2\delta_1^2 + g_1^2\delta_2^2)$$

and

$$b^2 = 4(|c_1 - c_2|^2 - sh_1 h_2)(\delta_1^2 + \delta_2^2) + p(h_2^2\delta_2^2 + h_1^2\delta_2^2),$$

where $p = (r - 1/r)^2$.

APPENDIX B: AVERAGE MAGNETIC SERIES

A least squares method for reducing the bias in the magnetic data is to compute the average of the local and remote magnetic data where the local magnetic component is related to the remote magnetic components by

$$\mathbf{B} = \underline{\mathbf{B}} \cdot \underline{\mathbf{D}} + \mathbf{R}. \tag{B.1}$$

The $\underline{\mathbf{D}}$ is the transfer function that transforms the remote data into an approximation of the local magnetic component using the power series representation (19) with $S_n(\omega)$ given by (21b) or (21d) for $n = 1, \ldots, N$ terms. This equation is exactly the same as the MT relationship (24a,b) when $\bar{Z}(\omega) = 1$. The matrix representation for the coefficients $\underline{\mathbf{d}}$ can therefore be written as

$$\mathbf{B}' \approx \underline{\mathbf{b}}' \cdot \underline{\mathbf{d}} + \mathbf{R}' \tag{B.2}$$

using weighted data. The $\mathbf{B}' = \mathbf{B}_{Fwgv}$ and $\mathbf{R}' = \mathbf{R}_{Fwgv}$ are J-vectors having, respectively, elements b_j' and r_j' for $j = 1, \ldots, J$ frequencies; $\underline{\mathbf{d}}$ is a M-vector of dimension $M = KN$ having elements $\underline{\mathbf{d}}_m$ with $m = n + N(k-1)$ for $n = 1, \ldots, N$ terms and $k = 1, \ldots, K$ magnetic components; and $\underline{\mathbf{b}}' = \underline{\mathbf{b}}_{DFwgv}$ is a $J \times M$ matrix having elements $\underline{b}'_{jm} = \underline{B}'_k(\omega_j) S_n(\omega_j)$ for the kth component of the weighted remote magnetic data using the previously estimated D and weights. The weights are generated from the residual $\mathbf{R}$. The robust least squares solutions is

$$\underline{\mathbf{d}} = (\underline{\mathbf{b}}'^{\mathrm{H}} \cdot \underline{\mathbf{b}}' + \lambda\mathbf{I})^{-1} \cdot (\underline{\mathbf{b}}'^{\mathrm{H}} \cdot \mathbf{B}') \tag{B.3}$$

for damping factor λ. The projected data are then given by $\mathbf{P} = \underline{\mathbf{b}} \cdot \underline{\mathbf{d}}$ and the average magnetic component is

$$\bar{\mathbf{B}} = (1/2)(\mathbf{B} + \mathbf{P}) \tag{B.4}$$

that replaces the local magnetic data $\mathbf{B}$. Less biased least squares transfer estimates are obtained because the uncorrelated noise will be reduced in $\bar{\mathbf{B}}$.

APPENDIX C: REPLACEMENT OF MISSING AND DELETED VALUES

In order to minimize the effects of the missing and rejected time values on the frequency weights and on the long period variations it was found necessary to replace the missing and rejected time values by the predicted values $p(t)$ plus the periodic part $l(t)$ of the residuals plus a local trend $\tau(t)$ used to merge the replacement values with the adjacent observed values. The predicted values $p(t)$ are computed from the Fourier inverse of $\mathbf{P} = \mathbf{B} \cdot \mathbf{Z}$ where $\mathbf{Z}$ is generated for the full frequency range using the smooth transfer representation. The electric data with replacement vaues are given by

$$\hat{e}(t) = \hat{g}(t)e(t) + [1 - \hat{g}(t)][p(t) + l(t) + \tau(t)], \tag{C.1}$$

where $\hat{g}(t) = 0$ for missing and rejected values and $\hat{g}(t) = 1$ otherwise (Appendix D).

Any periodic variations $l(t)$ in the residuals: $\hat{E}(\omega) - P(\omega)$, are given by the Inverse Fourier transform of

$$L(\omega) = [1 - \hat{F}(\omega)][\hat{E}(\omega) - P(\omega)], \tag{C.2}$$

where $\hat{F}(\omega) = 0$ for large amplitude spectral lines in the residuals and $\hat{F}(\omega) = 1$ otherwise. The background continuum for the filled data, the series minus the predicted and periodic parts, is then given by

$$\hat{c}(t) = \hat{g}(t)[e(t) - p(t) - l(t)] + [1 - \hat{g}(t)]\tau(t) \tag{C.3}$$

where the linear trend $\tau(t)$ is generated, for each gap, by a least squares fit to a number of values $\hat{c}(t)$ adjacent to the missing values. The electric data with replacement values are then given by

$$\hat{e}(t) = p(t) + l(t) + \hat{c}(t) \tag{C.4}$$

and the residuals with replacements are given by

$$r(t) = \hat{e}(t) - p(t) \tag{C.5}$$

The periodic variations are found by successive iterations starting with

$$L_0(\omega) = [1 - \hat{F}(\omega)][E(\omega) - P(\omega)]$$

then generating the continuum $\hat{c}_0(t)$ by (C.3) and refining the estimates of the periodic variations by

$$L_i(\omega) = L_{i-1}(\omega) + [1 - \hat{F}(\omega)]\hat{C}_{i-1}(\omega)$$

for $i > 1$. The iterations are stopped when $L_i(\omega)$ changes by only a small per cent. The electric data with replacement values are then approximated by

$$\hat{e}(t) = p(t) + l_i(t) + \hat{c}_i(t) \tag{C.6}$$

APPENDIX D: DISTORTION ESTIMATES

The coefficients d_n of the distortion (19) using one of the forms of $S_n(\omega)$ (21b–d) are found by least squares using $\hat{J} = 30$ estimates of $\bar{Z}$, Z and its variance distributed over the entire frequency range. The matrix form of (18) with misfit $\mathbf{r}$ for values divided by the estimates of the standard deviation ε (equal to the square root of the variance of Z) is given by

$$\mathbf{z} = \bar{\mathbf{z}} \cdot \mathbf{d} + \mathbf{r} \tag{D.1}$$

where $\mathbf{z}$ and $\mathbf{r}$ are $\hat{J}$-vectors having, respectively, elements $z_j = Z(\omega_j)/\varepsilon(\omega_j)$ and r_j, for $j = 1, \ldots, \hat{J}$ frequencies; $\mathbf{d}$ is a N-vector having elements d_n for $n = 1, \ldots, N$; and $\bar{\mathbf{z}}$ is a $\hat{J} \times N$ matrix having elements $\bar{z}_{jn} = \bar{Z}(\omega_j)S_n(\omega_j)/\varepsilon(\omega_j)$. The smoothed least-square solution is then

$$\mathbf{d} = (\bar{\mathbf{z}}^{\mathrm{H}} \cdot \bar{\mathbf{z}} + \lambda \mathbf{I})^{-1} \cdot (\bar{\mathbf{z}}^{\mathrm{H}} \cdot \mathbf{z}) \tag{D.2}$$

where $\mathbf{I}$ is the identity matrix. The damping factor λ is equal to 10^{-q} times the leading term of $\bar{\mathbf{z}}^{\mathrm{H}} \cdot \bar{\mathbf{z}}$ where the smallest q from $q = 2, \ldots, 6$, is chosen so that the χ^2 misfit, equal to $|\mathbf{r}^{\mathrm{H}} \cdot \mathbf{r}|$, is slightly less than the 95 per cent confidence level.

APPENDIX E: ITERATIVE METHOD

The iterative steps for obtaining transfer functions, weights and smooth representations are summarized in the following.

Transfer function

The iterative method for obtaining estimates of the transfer function consists of the following steps.

The initial $i = 1$ iteration has three steps

$$E \xrightarrow[2]{} W_0 \xrightarrow[3]{} E'_0,\ B'_0 \xrightarrow[4]{} Z_1,\ \delta Z_1,$$

where the weights are generated by step 2; the weighted data E'_0 and B'_0 are generated by step 3; and the transfer function Z_1 and its confidence interval δZ_1 are generated by step 4 using the least squares or remote reference methods.

For $i > 1$ the iteration has four steps

$$Z_{i-1},\ B \xrightarrow[1]{} R_{i-1},\ E_{i-1} \xrightarrow[2]{} wFgv_{i-1} \xrightarrow[3]{} E'_{i-1},\ P'_{i-1} \xrightarrow[4]{} U_i,\ \delta U_i,$$

where the residuals are generated from the data by step 1 using the estimated Z_{i-1} and the magnetic and electric data corrected for missing and large outliers; the weights generated by step 2; the weighted electric data E'_{i-1} and the weighted predicted data P'_{i-1} are generated by step 3; and the solution $U_i(\omega)$ and confidence interval $\delta U_i(\omega)$ are generated by step 4 using least-squares or remote reference estimates of U_i. The transfer function is then $Z_i(\omega) = Z_{i-1}(\omega)U_i(\omega)$ and the confidence interval is $\delta Z_i(\omega) = |Z_{i-1}(\omega)|\,\delta U_i(\omega)$ assuming noise free predicted electric data and $Z_{i-1} = \bar{Z}_{i-1}D_{i-1}$.

The revised 1-D model transfer function $\bar{Z}_i$ is then found from Z_i and its variance for the 1-D frequency range using the D^+ inversion routine. The revised distortion function D_i is then found from $\bar{Z}_i$, Z_i and its variance.

The process is then repeated for at least six cycles and terminated when $|U - 1|$ is everywhere less than δU.

Weights

The iterative method for obtaining the weights consists of the following steps.

The initial $i = 1$ iteration has a single step

$$E \xrightarrow[1]{} W,$$

where the initial continuum weights W for $M = 10$ is generated using E as a proxy for the residuals.

For $i > 1$ the iteration has six steps

$$R_{\hat{F}} \xrightarrow[1]{} W,\ \hat{W},\ R_{w\hat{w}} \xrightarrow[2]{} \hat{F},\ R_{\hat{F}} \xrightarrow[3]{} W,\ R_w \xrightarrow[4]{} F,\ R_{wF} \xrightarrow[5]{} g,\ R_{wFg} \xrightarrow[6]{} V,$$

where preliminary continuum weights W for $M = 10$ and refinement $\hat{W}$ are generated by steps 1 using the residuals $R_{\hat{F}}$ and a previously estimated rejection weights $\hat{F}$; the new rejection frequency weights $\hat{F}$ are generated by step 2 using $R_{w\hat{w}}$; the new continuum weights W are generated by step 3 using $R_{\hat{F}}$; the frequency weights F are generated by step 4 using R_w; the time weights g are generated by step 5 using R_{wF}; and the final continuum weights V are generated by step 6 using R_{wFg}. Steps 1 and 2 are necessary for finding the rejection weights $\hat{F}$ so that W deals with the continuum spectrum and F deals with the line spectrum.