Geophysical Journal International

Geophys. J. Int. (2014) **196,** 1365–1374
Advance Access publication 2014 January 2

doi: 10.1093/gji/ggt484

Robust magnetotelluric inversion

Tetsuo Matsuno,[1] Alan D. Chave,[2] Alan G. Jones,[3] Mark R. Muller[3] and Rob L. Evans[4]

[1]*National Institute of Polar Research, 10–3 Midoricho, Tachikawa, Tokyo 190–8518, Japan. E-mail: matsuno.tetsuo@nipr.ac.jp*
[2]*Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA*
[3]*School of Cosmic Physics, Dublin Institute for Advanced Studies, 5 Merrion Square, Dublin 2, Ireland*
[4]*Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA*

Accepted 2013 November 27. Received 2013 November 27; in original form 2013 June 28

SUMMARY

A robust magnetotelluric (MT) inversion algorithm has been developed on the basis of quantile-quantile (q-q) plotting with confidence band and statistical modelling of inversion residuals for the MT response function (apparent resistivity and phase). Once outliers in the inversion residuals are detected in the q-q plot with the confidence band and the statistical modelling with the Akaike information criterion, they are excluded from the inversion data set and a subsequent inversion is implemented with the culled data set. The exclusion of outliers and the subsequent inversion is repeated until the q-q plot is substantially linear within the confidence band, outliers predicted by the statistical modelling are unchanged from the prior inversion, and the misfit statistic is unchanged at a target level. The robust inversion algorithm was applied to synthetic data generated from a simple 2-D model and observational data from a 2-D transect in southern Africa. Outliers in the synthetic data, which come from extreme values added to the synthetic responses, produced spurious features in inversion models, but were detected by the robust algorithm and excluded to retrieve the true model. An application of the robust inversion algorithm to the field data demonstrates that the method is useful for data clean-up of outliers, which could include model as well as data inconsistency (for example, inability to fit a 2-D model to a 3-D data set), during inversion and for objectively obtaining a robust and optimal model. The present statistical method is available irrespective of the dimensionality of target structures (hence 2-D and 3-D structures) and of isotropy or anisotropy, and can operate as an external process to any inversion algorithm without modifications to the inversion program.

Key words: Inverse theory; Probability distributions; Magnetotellurics.

1 INTRODUCTION

The magnetotelluric (MT) technique is widely used for electromagnetic investigation of Earth's interior. The electrical parameter revealed by the MT technique, electrical resistivity (or its inverse conductivity), provides useful geophysical constraints on the structure and dynamics of Earth independent of other geophysical methods such as seismology. In recent times, MT experiments utilizing large numbers of sites to reveal 2-D and 3-D electrical resistivity structures have become the norm in academia and industry, posing challenges in data processing and interpretation.

The fundamental MT datum is the site-specific, frequency-dependent tensor relationship between the measured horizontal electric and magnetic fields

$$\mathbf{E} = \mathbf{Z} \bullet \mathbf{B}, \tag{1}$$

where **E** and **B** are two-vectors of the horizontal electric and magnetic field components at a specific site and frequency, **Z** is the second rank 2 × 2 MT response tensor connecting them, and • denotes the inner product. The MT response tensor is usually obtained through row-by-row robust or bounded influence least squares regression applied to the above equation.

The least squares solution is optimal if the theoretical conditions underlying the method are valid. However, the least squares solution is well known to be sensitive to small amounts of unusual data and uncorrelated noise in the electric and magnetic field variables. The development of remote reference and robust processing methods (e.g. Gamble *et al*. 1979; Jones & Jödicke 1984; Jones *et al*. 1989; Egbert 1997; Chave & Thomson 2004; Chave 2012) has solved this problem in most circumstances. These involve data-adaptive robust weighting schemes that eliminate the influence of data corresponding to large residuals in a statistically defensible manner.

It is not unusual to find one or more sites in a large MT survey that remain anomalous even after robust processing, usually due to correlated noise in the electric and magnetic channels, but sometimes due to unknown causes. However, least squares principles are the basis for MT inversion algorithms, and hence outlying sites (or frequencies at a given site) can produce bias in the resulting inversion models that can lead to incorrect interpretations.

Other than subjective exclusion of selected MT responses during inversion, the oldest approach to excluding outliers is based on the dispersion relations governing the MT response function (Weidelt & Chave 2012). This is most easily implemented using the D+ (Parker & Whaler 1981) and $\rho+$ (Parker & Booker 1996) algorithms to detect outliers. These methods are defined for a 1-D resistivity structure, but apply equally to the 2-D TM (transverse magnetic) mode based on theory (Weidelt & Kaikkonen 1994), and to 2-D TE (transverse electric) mode data based on empirical evidence (Weidelt & Chave 2012). However, Parker (2010) has cast doubt on the latter using a simple but geologically implausible model, hence caution is required in the event of a perverse and unlikely subsurface geometry.

Robust algorithms can be equally applied to MT inversion. One approach is to use an L_1 rather than L_2 statistical norm applied to the fit between data and calculated responses. Although superior to the L_2 norm in its treatment of outliers, the choice of the L_1 norm is not optimal because the distribution of inversion residuals are typically described by a Gaussian probability density function (PDF) contaminated by a fraction of outliers. In addition, the L_1 norm is not efficient; the variance of the median is $\pi/2$ times the variance of the mean for Gaussian variates. Porsani *et al*. (2001) introduced a multiple re-weighted least squares method (MRLS) that is based on a variable L_p norm approach for robust inversion of vertical electrical sounding data, and Santos *et al*. (2005) applied this algorithm to MT data. However, MRLS has the same statistical limitations as L_1 norm approaches with its implicit assumption about the distribution of residuals. Varentsov (2002) developed a robust inversion by introducing the robust measure of Huber into the calculation of the misfit and a weighting matrix into the model updates corresponding to the size of the residuals.

Robust processing in inversion has not been widely applied, mainly because it adds an additional nested iteration to the inversion process that typically requires substantial program modification and increases the computation time by a factor of 10 or more. Perhaps more importantly, some widely used inversion algorithms are proprietary, and hence are provided without source program. In this instance, a robust approach that does not require changes to the inversion program is preferred.

In this paper, we propose a robust MT inversion algorithm that is based on the statistics of the residuals obtained in inversion. Anomalous data that have extreme residuals are checked by quantile-quantile (hereafter q-q) plotting with confidence band at a specified significance level, and are objectively detected and removed from the inversion data set by statistical modelling of expected outliers. The principle is nearly identical to the widely used M-estimator for MT data processing. The data culling is repeated until outliers are excluded from the data set through the statistical modelling by confirming in the q-q plot with the confidence band, the inversion residuals are not changed substantially, and summary statistics like the rms misfit does not vary statistically. The method does not depend on model dimensionality or the use of a specific inversion algorithm, hence can be generally applied to either 1-D, 2-D or 3-D inversion, with or without anisotropy. In addition, the method operates as an external process to an inversion program and does not require any modifications to the inversion program.

2 METHOD

The robust inversion algorithm is a simple implementation of an M-estimator that is guided by q-q plotting of inversion residuals and the exclusion of data used in inversion based on the size of the residuals. The q-q plot with the confidence band serves as a visual aid to check the presence of outliers, and to examine the compatibility of the distribution of the residuals with the theoretical one. The detection and exclusion of extreme residuals from the inversion data set is objectively implemented by statistical modelling of the inversion residuals.

For the q-q plotting, the data component is obtained from the order statistics of the inversion residuals. An electrical resistivity model is first obtained using an inversion program in a standard manner. The inversion residuals are calculated at each frequency and at each site using the observed response y and the counterpart $G(\mathbf{m})$ predicted from the derived inversion model $\mathbf{m}$ through a forward modelling operator G, and are then normalized by the observed standard deviation s, as $[y - G(\mathbf{m})]/s$. They are finally sorted into ascending values for each MT element.

The quantile for the q-q plotting is drawn from a theoretical (or expected) distribution that the inversion residuals should follow. The theoretical quantile can be calculated by the formulation for the quantile function for N residuals,

$$q_i = F^{-1}\left[(i - 1/2)/N\right] \quad i = 1, \ldots, N, \tag{2}$$

where F is the appropriate cumulative distribution function (CDF). The quantiles q_i divide the area under the PDF $f(x)$ that corresponds to $F(x)$ into $N + 1$ equal area pieces; hence, the quantiles define equal probability increments. The normal (or Gaussian) distribution is used for F in this study. The theoretical distributions for apparent resistivity (or the magnitude squared response function) and phase introduced in Chave & Lezaeta (2007) could also be used, but with typical observational uncertainty for the MT response function of a few per cent or less, these distributions are well approximated by a Gaussian model.

The q-q plot is then constructed to compare the statistical distribution of the inversion residuals with the theoretical distribution (it is usual that the quantiles are plotted in the x-axis and the order statistics are plotted in the y-axis). If the inversion residuals are drawn from $f(x)$, then the q-q plot will approximate a straight line, with the slope and intercept defining the scale and location parameters for the distribution. With real data, it is more typical to observe a straight line at the centre (presuming that the statistical model is correct) with a small fraction of outliers that appear as sharp vertical excursions at the distribution extremes. Sometimes the q-q plot is not straight at the centre due to the presence of outliers and a systematic difference between the target distribution and the real distribution.

When some of the data are excluded as outliers from the data set used in the inversion, the truncated distribution must be used to compute quantiles for further q-q plotting. The computation is implemented in the standard manner as

$$F(q_i) = [F(b) - F(a)] \times \frac{i - 1/2}{M} + F(a) \quad i = 1, \ldots, N, \tag{3}$$

where m_1 and m_2 are, respectively, the number of data censored from the bottom and top of the distribution, and M is the number of data after excluding the censored data as outliers (i.e. $M = N - m_1 - m_2$; Chave & Thomson 2004). Reasonable choices of a and b are the m_1th and $(N - m_2 + 1)$th quantiles of the original distribution.

The confidence band on the q-q plot is useful to examine the normality of the distribution of the inversion residuals at a specified

significance level. Confidence upper and lower bands, q_+ and q_-, are calculated by

$$q_+ = \bar{y} + s \times F^{-1}\left(\frac{i-1}{N} + d_c\right),$$
$$q_- = \bar{y} + s \times F^{-1}\left(\frac{i}{N} - d_c\right), \qquad i = 1, \ldots, N \tag{4}$$

where $\bar{y}$ is a sample mean, s is a sample standard deviation for the Gaussian CDF F and d_c is a critical point of the Kolmogorov–Smirnov (K–S) statistic (Stirling 1982; Michael 1983). The critical point of the K–S statistic is asymptotically calculated by $d_c = \text{const.}/\sqrt{N}$. The constant may be changed by the significance level tested and the knowledge of the target distribution (Stephens 1974). For 95 per cent significance level in a double sided test for the elements of the MT response function, the 5 per cent tail probability should be equally allocated to the upper and lower tails of all the elements. For the elements of the 2-D MT response function, for example, the percentage point for the critical point of the K–S statistic would be 99.375 per cent for each distribution end. The mean of the inversion residuals is known to be zero and the standard deviation of them is estimated from the sample population, and hence case 2 in Stephens (1974) may be proper for an application to the inversion residual. Interpolation and extrapolation of the values listed in case 2 of Stephens (1974) by considering the above percentage point and the number of inversion residuals tested in the following sections (~500 for the synthetic test and ~5000 for the real data application) yielded the constant of 1.65 for 95 per cent significance level of the double sided test for all four elements of the 2-D MT response function. If some data corresponding to extreme residuals are culled from the inversion data set, the reduced number of residuals and the truncated distribution form are used in the calculation of the critical point and the confidence band.

Outliers that are recognized in the q-q plot can be objectively detected by introducing a statistical model for the inversion residuals. The statistical model used in this study is that sample populations except outliers follow a single normal distribution while outliers separately follow different normal distributions, and that the means of all the normal distributions are different from each other while their standard deviations are common (Kitagawa 1979, 1981). The likelihood of the statistical model can be derived from the joint PDF of the order statistics (Pynnönen 1992; David & Nagaraja 2003),

$$L(x) = (N - m_1 - m_2)! \sum_{i,j} \left[\prod_{i=1}^{m_1} f(x_i; \mu_j, \sigma^2) \times \prod_{i=m_1+1}^{N-m_2} f(x_i; \mu_j, \sigma^2) \prod_{i=N-m_2+1}^{n} f(x_i; \mu_j, \sigma^2) \right],$$
$$i = 1, \ldots, N, \quad j = 1, \ldots, m_1 + m_2 + 1$$

with

$$f(x; \mu_i, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu_i)^2}{\sigma^2}\right] \quad i = 1, \ldots, N \tag{5}$$

for the normal distribution. The summation extends over all possible permutations of indices of i and j, but some of them can be omitted by making assumptions on the statistical model; outliers exist only on the upper and lower ends of the distribution, are separated from the mean of the sample population excluding outliers, and are not clustered within each distribution end (Kitagawa 1981; Kitagawa & Akaike 1982; Pynnönen 1992). Under these conditions, the likelihood function is approximated by

$$L(x) = (N - m_1 - m_2)! \prod_{i=1}^{m_1} f(x_i; \mu_j, \sigma^2) \prod_{i=m_1+1}^{N-m_2} f(x_i; \mu_j, \sigma^2) \times \prod_{i=N-m_2+1}^{n} f(x_i; \mu_j, \sigma^2). \tag{6}$$

This formula is a good approximation to the original one in practical calculations, although it could produce underestimation of the original likelihood function when outliers are clustered within each distribution end (Kitagawa 1981; Kitagawa & Akaike 1982; Tango 1986). The log likelihood function for the formula is then given by

$$\log L(x) = \log(N - m_1 - m_2)! - \frac{N}{2}\log(2\pi\sigma^2) - \frac{\sum_{i=m_1+1}^{N-m_2}(x_i - \mu)^2}{2\sigma^2}. \tag{7}$$

A sample mean $\bar{y}$ and a sample variance s^2 are used to evaluate the function instead of μ and σ^2.

Outliers in the sample population may be determined in a condition that the log likelihood function is maximized and the number of the outliers is minimized. Given the condition, an information criterion is useful to objectively determine outliers that are optimal for the data set. The Akaike information criterion (AIC) is a simple and versatile information criterion, and the AIC for the statistical model of the outlier is introduced by

$$\text{AIC} = -2\log L(x) + 2(m_1 + m_2 + 2) \tag{8}$$

(Kitagawa 1981). For candidate outliers on the upper and lower ends of the distribution, the AIC is systematically calculated for all possible combinations of outliers up to a specified limit (for example, 10–20 per cent of the upper and lower ends of the distribution of the ordered inversion residuals). The combination with a minimum AIC is considered to be optimal, and data corresponding to the predicted outliers are excluded from the inversion data set.

After the exclusion of data outliers, the subsequent inversion is carried out in the usual manner using the culled data set. After that inversion is completed, the inversion residuals are recalculated using the MT response functions predicted from the inversion model and the initial data set. By comparing the predictions calculated from the latest inversion model with the initial data set, the excluded data can change at any step due to updating of the inversion model. This iteration of data culling and the subsequent inversion (i.e. robust run) is stopped when the q-q plot does not show significant (compared to the confidence bound) deviation over the distribution, particularly at the distribution ends, outliers predicted by the statistical model are not changed, and summary statistics such as the rms misfit and the objective function are not significantly changed from the prior inversion.

3 SYNTHETIC TEST

The robust inversion algorithm is first tested on a synthetic 2-D model. Consequently, the data elements for analysis are apparent resistivity and phase of the TE and TM modes. These four elements are dealt with independently.

Fig. 1(a) shows a synthetic 2-D model consisting of a homogeneous background of 100-Ω-m resistivity with a pair of lower (10 Ω-m) and higher (1000 Ω-m) resistivity square anomalies embedded within it. The two square anomalies are 10 × 10 km in

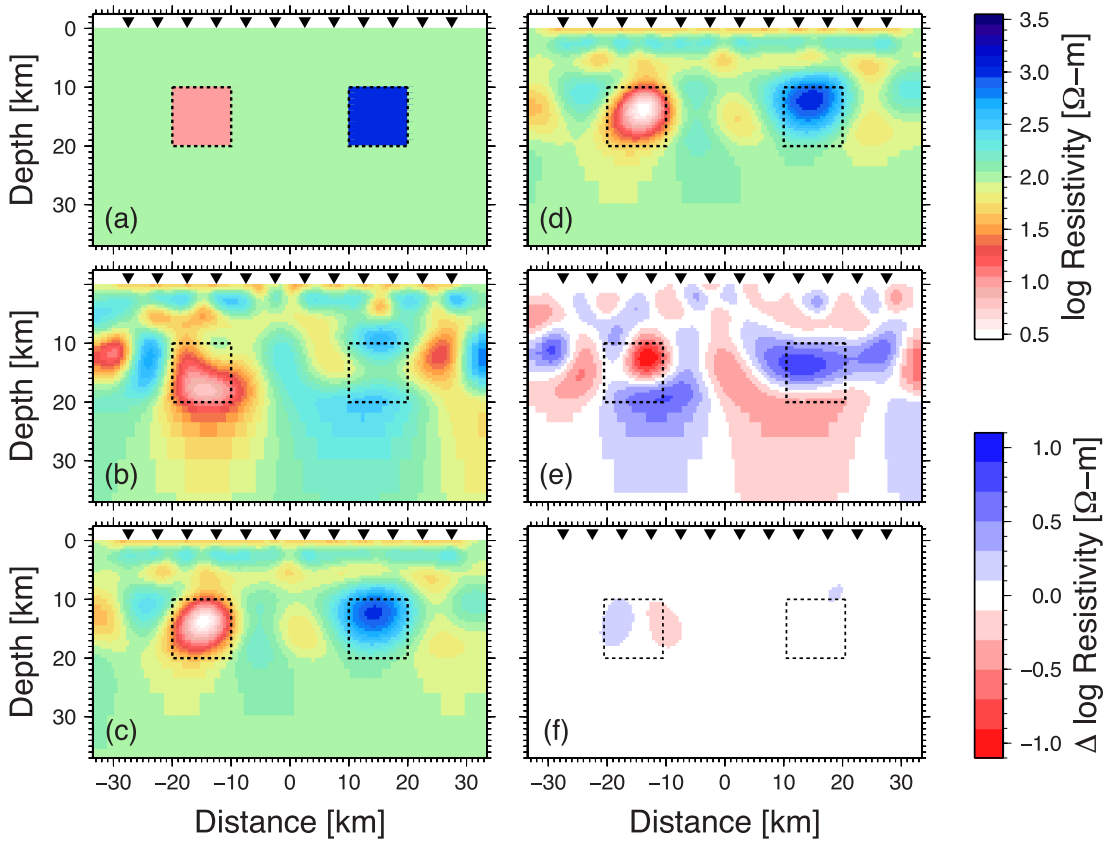

Figure 1. (a) A synthetic 2-D model comprising two square regions of 10 and 1000 Ω-m resistivity embedded within a 100 Ω-m halfspace, (b) the model obtained by inverting the synthetic data with 3 per cent Gaussian noise and 10 per cent additive Cauchy-distributed noise for contamination, as described in the text, (c) the model obtained by inverting the culled data set at the second (final) robust run, (d) the model obtained by inverting the synthetic data with only the 3 per cent Gaussian noise, (e) the difference in logarithmic resistivity between the two models in (b) and (d), and (f) the difference in logarithmic resistivity between the two models in (c) and (d). The dashed squares in the figures indicate the location of the two square regions in the synthetic model of (a). The inverted triangles at the top of the models denote the site locations. In all of the inversions, the regularization parameter for model smoothness τ_s was 1, and the weighting factors for model smoothness α and β were 1 and 3, respectively.

cross-section, buried at depths of 10 to 20 km, and separated horizontally by 20 km.

Synthetic data were generated from the 2-D model of Fig. 1(a) using the forward modelling program of Baba & Seama (2002). The synthetic data were calculated at 11 periods from 0.32 to 100 s spaced equally on a logarithmic scale, and at 12 sites at 10 km intervals spanning the square-shaped anomalies, as indicated by the inverted triangles in Fig. 1(a). The total number of data for each MT element is 132. Three per cent Gaussian noise was added to all of the MT response functions to simulate scatter and the standard deviation of actual data.

To simulate contamination of the MT response function, one datum at each site in each MT element (comprising in total 9 per cent of the data) was more scattered with 10 per cent Cauchy-distributed noise. The Cauchy distribution (also Student's t distribution with one degree of freedom) possesses algebraic rather than exponential tails, and hence exhibits a much greater tendency to produce extreme values (outliers) by comparison to the Gaussian distribution. During this process, the standard deviations of the MT response functions, which were derived from the 3 per cent Gaussian noise, were not changed.

The non-linear conjugate gradient algorithm of Rodi & Mackie (2001) that can handle transverse electrical anisotropy was used to invert the synthetic data. The regularization parameter for the degree of anisotropy (Baba *et al*. 2006) was set to 100 to force isotropy on the result. This value yielded isotropic models in prior studies (e.g. Baba *et al*. 2006; Matsuno *et al*. 2010). The starting model for all of the inversions is a uniform 100 Ω-m structure. The regularization parameter for model smoothness (τ_s) to determine an optimal model can be varied at each step of the robust run. Fig. 2 shows the rms misfit and the model roughness during the robust runs for several τ_s values (30, 10, 3, 1, 0.3, 0.1 and 0.03). Model roughness at later runs is smaller because the culling of outliers does not generate spurious heterogeneities that overfit the data. The curves of rms misfit and model roughness display an L shape. Based on the usual L-curve criterion, the inversion models with $\tau_s = 1$ are considered to be optimal for all of the robust runs.

There are other free parameters for model smoothness in the inversion program used, α and β factors in weighting functions that control horizontal and vertical model smoothness. These factors can have large impacts on the resultant inversion model (e.g. Matsuno *et al*. 2010), and optimal values could also change during the robust runs. A plausible selection of the optimal factors is ones yielding a model with the smallest misfit and the smallest model roughness. However, this selection does not necessarily give a model compatible with other geophysical and geological constraints, and may be beyond the ability of the robust algorithm to address. For the contaminated synthetic data, a search for optimal α and β values showed that differences in the misfit and the model roughness are insignificantly small and did not show a clear trend for selecting

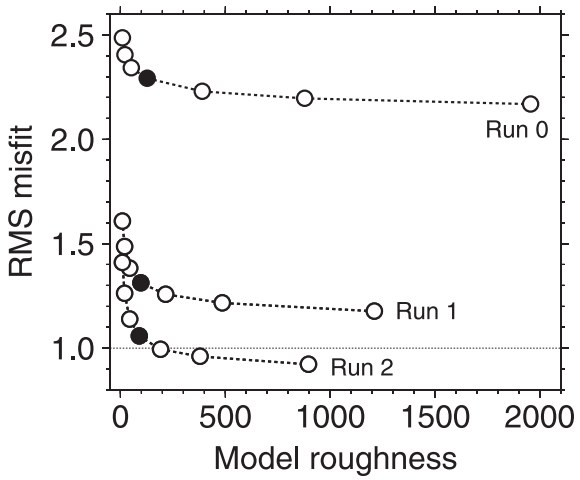

Figure 2. The rms misfit and model roughness for the inversion models with the contaminated synthetic data at all of the robust runs with several regularization parameters for model smoothness τ_s. The τ_s values examined are 30, 10, 3, 1, 0.3, 0.1 and 0.03. The larger τ_s consistently produces the model with higher rms misfit and smaller model roughness. The filled circle indicates the inversion model for $\tau_s = 1$ at each robust run, which was selected as optimal by the L-curve criterion.

an optimal solution. The α and β values were, respectively, set to 1 and 3 based on the closeness of a resultant model by the robust algorithm to the true model, and they were fixed during the robust runs.

Fig. 1(b) is the model from inverting the synthetic data contaminated with Cauchy-distributed noise. The two anomalies corresponding to the synthetic squares are quite distorted, and some spurious anomalies appear. Compared with the model from inverting the uncontaminated synthetic data using the same τ_s, α and β values (Fig. 1d), differences in logarithmic resistivity as large as $\sim\pm1.0$ (Fig. 1e). The total rms misfit of the contaminated inversion model is 2.29, which is much larger than 1.00 for the Fig. 1(d) model.

The robust approach reduced the misfit to 1.31 at the first run, and to 1.06 at the second run (Fig. 2). The third run does not produce a significant difference from the second run in either the inversion result, the q-q plot, or the AIC detection of outliers, hence the second run is the final step for the data set. The final inversion model retrieved features of the two conductive anomalies, and precluded spurious anomalies seen in the inversion model with the contaminated data (Fig. 1c), as quantitatively shown by differences in logarithmic resistivity that are almost all $<\pm0.1$ (Fig. 1f).

For the contaminated model in Fig. 1(b), the residuals on both ends of the distribution clearly deviate from the straight line characterizing the middle of the distributions in all four response tensor elements (Fig. 3a). In addition, some of the residuals are beyond the 95 per cent confidence bands between the middle and the ends of the distribution (Fig. 3a), suggesting that the normality of the distribution is violated by the presence of outliers. All of the q-q plots for the final run inversion model show no meaningful deviations at both ends of the distribution, are nearly linear, and are within the 95 per cent confidence bands (Fig. 3b), suggesting that the vast bulk of the outliers have been culled such that the residuals are Gaussian. Truncated distributions were used to compute theoretical quantiles based on the number of data culled. The search for outliers using the AIC was done for ~20 per cent of the original data on each distribution end for each MT element (25 for each distribution end, out of the 132 original data), and consequently 12 data points (9 per cent of the total data) are classified as outliers and culled at the final run for each MT element.

The uncontaminated and contaminated synthetic data and the predictions from the inversion models of Figs 1(b)–(d) are plotted in Figs 4 and 5. It is obvious that data deviating from the population are being excluded as outliers. The prediction from the zeroeth run is different from those from the inversion model without contamination and from the final inversion model by no more than 10 per cent for apparent resistivity and no more than 5 degrees for phase. The prediction from the inversion model at the final run is quite similar to that from the inversion model without contamination.

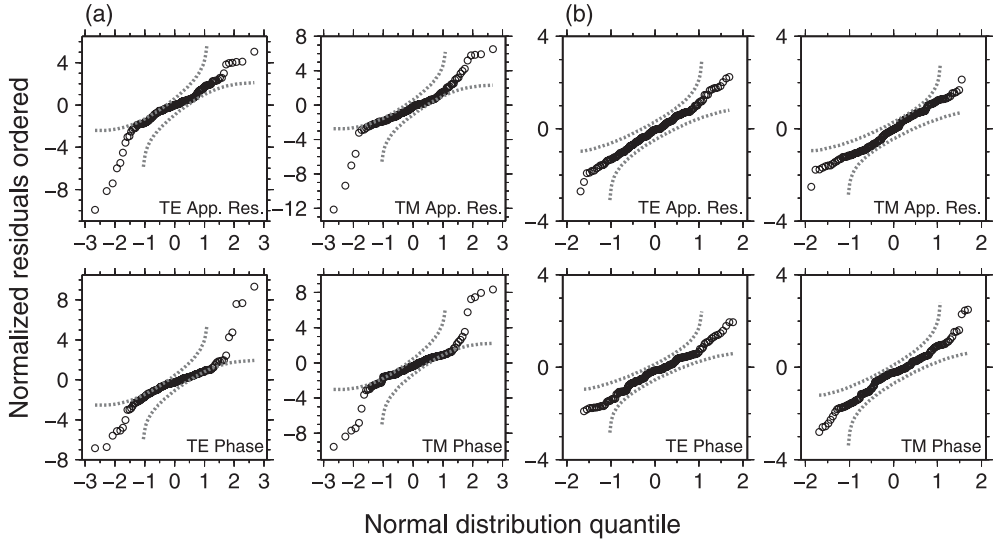

Figure 3. Quantile-quantile plots for the inversion residuals of apparent resistivity and phase for the TE and TM modes for (a) the zeroeth run inversion model of Fig. 1(b), and (b) the second (final) run inversion model of Fig. 1(c). The ordinate is the order statistics obtained by ranking the inversion residuals normalized by their errors (the standard deviation), while the abscissa is the theoretical quantiles derived from the normal distribution. Grey dashed lines represent the 95 per cent confidence bands on the q-q plots. The truncated form of the normal distribution was used to derive the theoretical quantiles for the final inversion model to account for the data excluded in each MT element. The range of the ordinate in Fig. 3(a) is different for the four MT elements, while that in Fig. 3(b) is same for the four MT elements. The range of the abscissa is different for Figs 3(a) and (b).

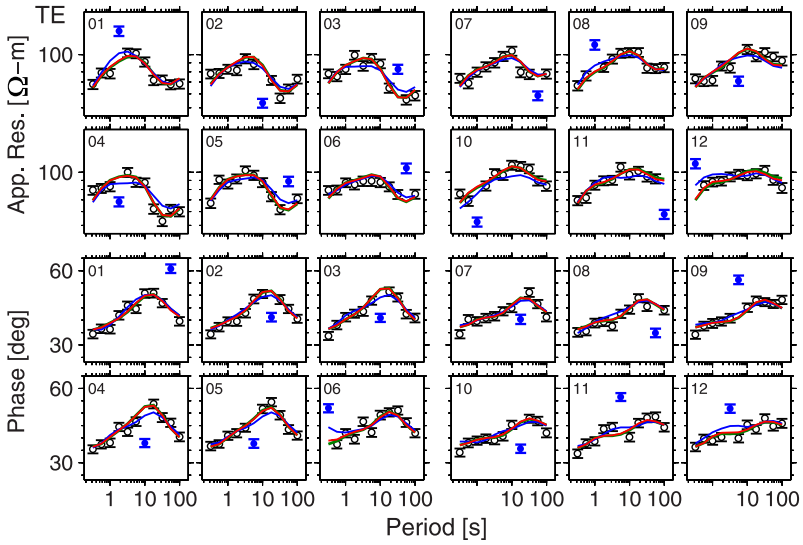

Figure 4. TE mode apparent resistivity (upper two rows) and phase (lower two rows) for the uncontaminated synthetic data (open circles), the contaminated synthetic data (blue filled circles), the prediction from the inversion model at the zeroeth robust run (Fig. 1b; blue curves) and the prediction from the inversion model at the second (final) robust run (Fig. 1c; red curves), the prediction from the inversion model with uncontaminated data (Fig. 1d; green curves, they are nearly identical to the red curves). All of the contaminated data were removed from the data set at the final run. The error bars represent one standard deviation. The site numbers, which are numbered from the left side, are shown at the upper left corner of each figure.

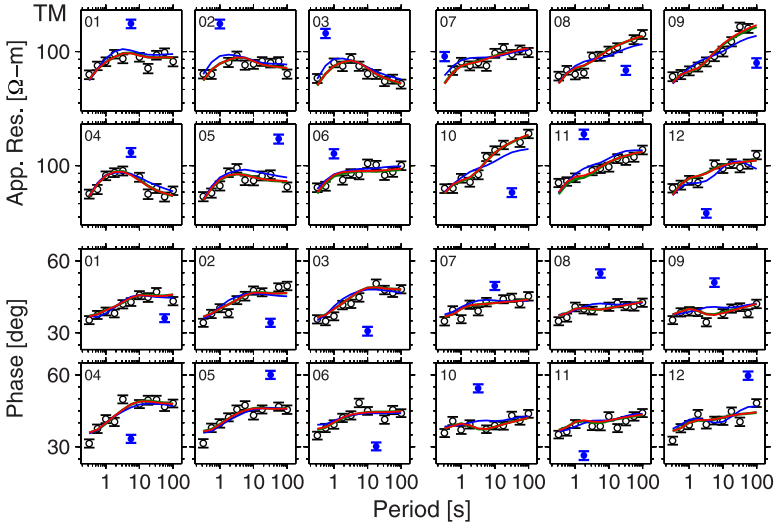

Figure 5. TM mode apparent resistivity and phase for the uncontaminated and contaminated synthetic data, and the predictions from the three types of inversion models (Fig. 1b–d). The form of the figure and the symbols are the same as for Fig. 4. The contaminated data shown by the blue filled circles were removed from the data set at the final run.

4 APPLICATION TO REAL DATA

For a real data application of the robust inversion algorithm, a subset of sites from a 2-D MT transect presented by Muller *et al*. (2009) was employed. The 2-D MT transect crosses the Archaean Kaapvaal Craton, the Proterozoic Rehoboth Terrane and the Late Proterozoic/Early Phanerozoic Ghanzi-Chobe/Damara Belt in southern Africa, and comprises an attempt to image lithospheric structure in order to understand the cratonic evolution as part of the SAMTEX (Southern African MagnetoTelluric Experiment) project (Jones *et al*. 2009). Below, the data subset and the inversion analysis of Muller *et al*. (2009) are reviewed; see Muller *et al*. (2009) for detailed descriptions of the data and their inversion, along with interpretations of the ensuing electrical resistivity models.

The data possess an electrical strike of 25°E of *N* determined using the multi-site and multi-frequency tensor decomposition of Groom & Bailey (1989) presented by McNeice & Jones (2001), and were used to derive an optimal model for the Rehoboth Terrane in Muller *et al*. (2009). The subset comprises 21 sites spaced at about 20 km intervals, and the number of periods initially used from most sites is 65–84, except one where it is 30. The shortest period was 0.003125 s (frequency of 320 Hz) and the longest period was 5814 s. The total number of data for each MT element is 1581.

The anisotropic inversion program of Rodi & Mackie (2001) used for the synthetic test was also used to invert the Rehoboth data. This program is not the isotropic one used by Muller *et al*. (2009), but an isotropic inversion model was forced by setting the regularization parameter for the degree of anisotropy to be 100. Data error floors for apparent resistivity and phase were set to 10 and 5 per cent (1.45 degrees), respectively, in the same manner as in Muller *et al*. (2009). The starting model for all of the inversions was a uniform 100 Ω-m halfspace.

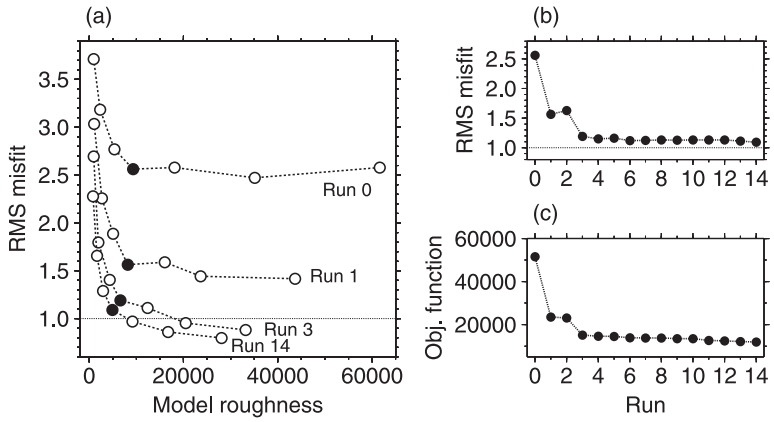

Figure 6. Fig. 6(a) shows rms misfit and model roughness of the inversion models for the Rehoboth Terrane data of Muller *et al*. (2009) at the zeroeth, first, third and fourteenth (final) robust runs with several regularization parameters for model smoothness τ_s. The τ_s vales examined are 30, 10, 3, 1, 0.3, 0.1 and 0.03. The larger τ_s consistently produces the model with higher rms misfit and smaller model roughness. The filled circles indicates the model for $\tau_s = 1$ at each robust run, which was selected as optimal by the L-curve criterion. The total rms misfit and the objective function for the optimal inversion model ($\tau_s = 1$) as a function of run number are shown in Figs 6(b) and (c), respectively.

Optimal regularization parameters for the model smoothness of τ_s were determined by plotting total rms misfit and model roughness for the robust runs with several τ_s values. Fig. 6(a) shows L-curves at four representative (zeroeth, first, third and fourteenth) robust runs. The τ_s values examined were 30, 10, 3, 1, 0.3, 0.1 and 0.03. As the robust run proceeds, the model roughness is reduced and the right end of the L-curve is shortened, yielding smoother inversion models. Based on the L-curve criterion, inversion models with $\tau_s = 1$ are considered to be optimal.

The factors α and β for model smoothness were, respectively, set to 1.0 and 1.3, and were fixed during the robust runs. Their values were selected by examination of the results of trials using different values and through consideration of reasonable geophysical and geological interpretations based on Muller *et al*. (2009).

The choice of final robust run was based on a minimum objective function value. A straightforward criterion for finalization is constancy of the rms misfit and outliers detected at consecutive runs. Fig. 6(b) shows the suite of the total rms misfit for all of the inversion models with $\tau_s = 1$. The rms misfit for the zeroeth inversion model was 2.56, and it was 1.56 for the first run. The misfit did increase slightly to 1.63 at the second run, and then decreased to ∼1.1–1.2 at later runs. The misfit at the later runs was insignificantly changed near the target value of 1.0 but with some variation. Outliers detected in the statistical model were also insignificantly varied (<10 data) at latter runs. Meanwhile, the objective function monotonically decreased until the fourteenth run (Fig. 6c) and increased at the subsequent one, and accordingly the fourteenth run was selected as the final one.

Fig. 7 shows q-q plots for the zeroeth and final inversion models. Truncated distributions are used to obtain theoretical quantiles for the first to the final runs. Excursions at the lower ends of the distribution for apparent resistivity and at both ends of the distribution for phase are large in the TE and TM modes for the zeroeth inversion model (Fig. 7a). Distortions in the distribution for the zeroeth inversion model are clearly seen as deviations of the residuals from the 95 per cent confidence bands (Fig. 7a). These excursions were almost removed as the robust run proceeded. However, some deviations remain even at the final robust run, especially on the lower end of the TE apparent resistivity and the upper end of the TE phase as shown by residuals lying outside the 95 per cent significance band (Fig. 7b). It is possible with real data that residuals include not only data outliers but also anomalies that result from inconsistencies of the inversion model with the data set. The remaining deviations after the application of the robust algorithm probably show that data include inconsistencies with the 2-D model even though the Groom–Bailey distortion decomposition of McNeice & Jones (2001) was applied, and the robust algorithm likely removed outliers as well as some anomalies due to their inconsistency. A potential cause for the remaining deviations is the existence of 3-D structures off the 2-D profile. The effects of such structures are primarily found in the TE mode data (e.g. Jones 1983; Wannamaker *et al*. 1984; Ledo *et al*. 2002; Ledo 2005), although for some special geometries the TM mode data can be more affected (e.g. Park & Mackie 2000). Removing data that are inconsistent with the assumed dimensionality of the interpretation is a standard approach dating to the 1970s and earlier when non-1-D data were excluded prior to 1-D modelling, and from the 1980s onwards 3-D data have been routinely excluded from 3-D interpretation. Indeed, the whole premise of the Groom–Bailey approach (Groom & Bailey 1989, 1991) and others like it is to identify those data that are statistically inconsistent with a 2-D regional Earth and to remove them prior to inversion.

Searches for outliers using the AIC were done for ∼20 per cent of the original data on each distribution end for each MT element (300 for each distribution end, out of the 1581 original data), and consequently 106 apparent resistivities (6.7 per cent) and 150 phase values (9.5 per cent) in the TE mode, and 24 apparent resistivities (1.5 per cent) and 50 phase values (3.2 per cent) in the TM mode were removed, showing that the number of data culled for the TE data is ∼3.5 times larger than that for the TM mode. The number of sites in which >10 data were excluded is five, that in which 7 data were excluded is one, that in which 3–1 data were excluded is eight and the remaining seven sites did not involve any data culling.

Fig. 8 shows inversion models at the four representative (zeroeth, first, third and final) runs. The zeroeth inversion model has several noticeable features: a shallow resistor of ≥1000 Ω-m resistivity with variable thickness (≤200 km in the middle and on the left, but 20 km on the right), and a conductor (≤1–10 Ω-m) between −80 and −150 km horizontal distance below 80 km depth. After the first run, different features emerged. The conductor seen in the zeroeth model disappeared, and a larger conductor emerged between 0 and ≥250 km horizontal distance below 150 km depth. At the third run, the inversion model retrieved similar features to those seen

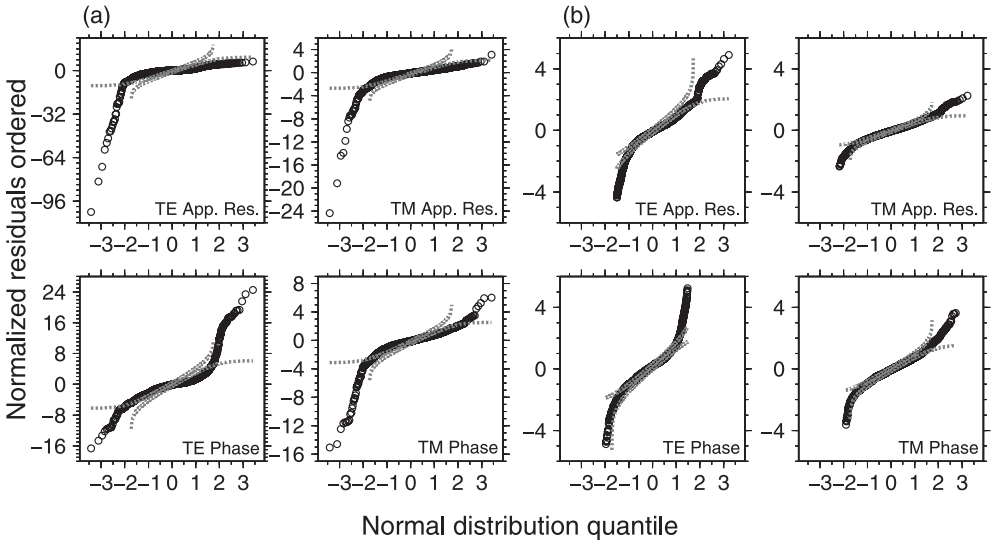

Figure 7. Quantile-quantile plots for inversion residuals of apparent resistivity and phase for the TE and TM modes for inversion models at the zeroeth robust run (a), and at the fourteenth (final) robust run (b) for the Rehoboth data. The ordinate was obtained by ranking the inversion residuals normalized by their errors (the standard deviation), while the theoretical quantiles were derived from the normal distribution. Grey dashed lines represent 95 per cent confidence bands on the q-q plots. The truncated form of the normal distribution was used to derive the theoretical quantiles for the final inversion model to account for the data excluded in each MT element. The range of the ordinate in Fig. 7(a) is different for all four MT elements, while that in Fig. 7(b) is same for all four MT elements. The range of the abscissa is same for Figs 7(a) and (b).

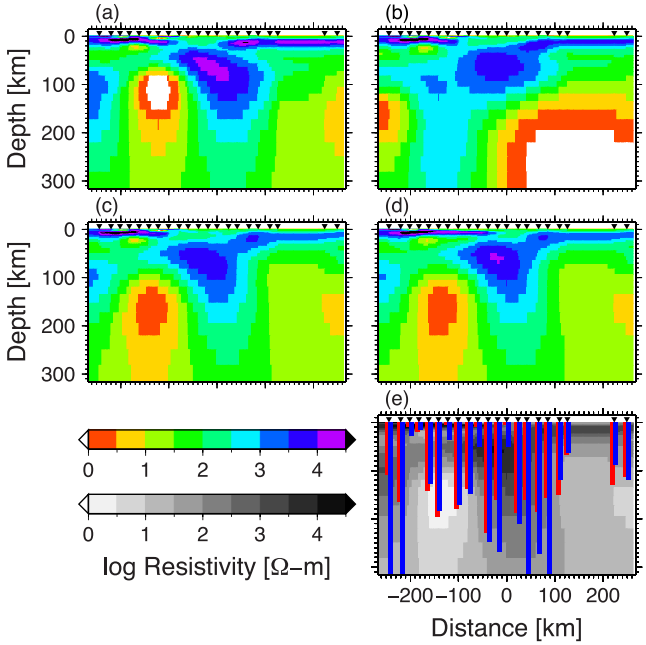

Figure 8. Inversion models at (a) the zeroeth, (b) the first, (c) the third and (d) the fourteenth (final) runs for the Rehoboth data. Fig. 8(e) is a grey scale plot of the final inversion model of Fig. 8(d), on which 1-D Niblett-Bostick transform results (red and blue bars) for the data set at the final run are overlaid. The red bars approximate the penetration depth of the TE mode data at each site, and the blue bars approximate that of the TM mode data. The sites are shown by inverted triangles at the top of each figure.

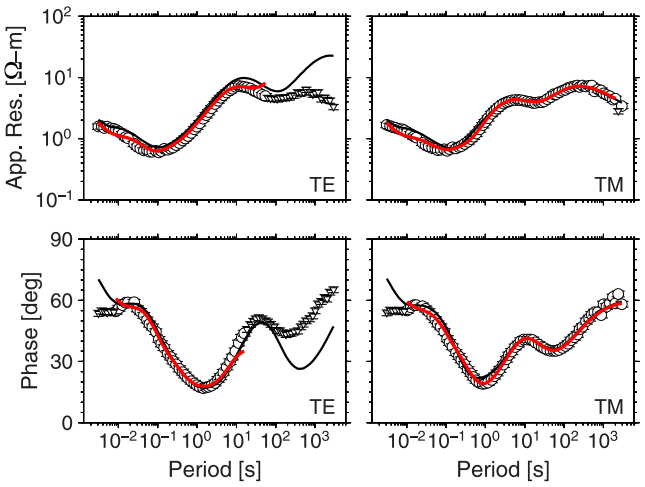

Figure 9. An example of the observed MT response functions (open circles and crosses), and predictions from the zeroeth and fourteenth (final) robust run inversion models of Fig. 8 (curves). Open circles and inverted triangles comprise the initial data set used in the zeroeth inversion, and the inverted triangles indicate the data removed from the data set at the final run. The error bars indicate one standard deviation. The black and red curves are the predictions from the zeroeth and final robust run models, respectively. The site plotted is the seventh from the left in Fig. 8.

in the zeroeth model. The resistivity contrast between the adjacent resistor and conductor becomes weaker in the middle of the model compared to the zeroeth model because overfitting to extreme data has been mitigated by data culling and the model smoothness is enhanced as quantified by the model roughness (Fig. 6a). The final inversion model is similar to the third one but with a smaller model roughness. Penetration depths defined by the 1-D Niblett-Bostick algorithm (Niblett & Sayn-Wittgenstein 1960; Bostick 1977) were also evaluated from the final data set after Muller *et al*. (2009). The results suggest that the final data set is sensitive to many of the observed model features, especially the resistor in the middle of the model.

Fig. 9 shows an exemplar (the site that is seventh from the left) set of the observed MT response functions, without values excluded at the final robust run, and those predicted from inversion models at the zeroeth and final runs. The data excluded are at the high and low

frequency ends, and accordingly many of predictions at the zeroeth and the final runs are different from each other around those periods.

5 CONCLUDING REMARKS

The robust inversion algorithm based on the q-q plotting with the confidence band of the inversion residuals of the MT apparent resistivity and phase and the objective exclusion of inversion data based on the AIC statistical modelling has been introduced. The robust inversion algorithm is an extension to the coarser and more conventional use of summary statistics like the rms misfit to determine a robust and optimal inversion model.

The synthetic testing using a simple 2-D electrical resistivity structure show that outlier contamination can produce a spurious result if the influence of outliers is not considered during inversion, and that the robust algorithm is useful for excluding outliers. The application of the robust algorithm to the real data from the Rehoboth Terrane, South Africa, demonstrates that it achieves objective estimation of a robust inversion model and more efficient data reduction than subjective data culling. While the robust algorithm is implemented as an outer loop to an existing inversion program without any modifications in this study, robust measures can also be directly incorporated into inversion programs. In any case, the q-q plotting with the confidence band and the statistical modelling using the AIC for detection of outliers is a supportive tool to exclude outliers in the data set, resulting in robustness of the inversion model. The robust algorithm can be applied independent of dimensionality and directional dependency (isotropy/anisotropy) of the target electrical resistivity structure. As a final note, the application of the robust algorithm to the inversion of the vertical magnetic field transfer function and other transfer functions is straightforward.

ACKNOWLEDGEMENTS

TM was supported by the scientific program of TAIGA (transcrustal advection and in-situ reaction of global sub-seafloor aquifer) sponsored by the MEXT of Japan, and is supported by the NIPR project KP-7. ADC is supported by US National Science Foundation (NSF) grant EAR1015185. The SAMTEX data were acquired through funding provided by the Continental Dynamics program of NSF (grant EAR0455242 to RLE), the South African Department of Science and Technology (grant to South African Council for Geoscience) and Science Foundation Ireland (grant 05/RGP/GEO001 to AGJ) plus financial and/or logistical support provided by all members of the SAMTEX consortium listed in the acknowledgments of Jones *et al*. (2009). Many people strove hard to acquire the SAMTEX data set, and all are very gratefully thanked (see authors and those acknowledged in Jones *et al*. (2009)). Comments by an anonymous reviewer and the editor, Mark Everett, were valuable. The GMT software (Wessel & Smith 1998) was used to create all of the figures.

REFERENCES

Baba, K. & Seama, N., 2002. A new technique for the incorporation of seafloor topography in electromagnetic modelling, *Geophys. J. Int.,* **150**(2), 392–402.

Baba, K., Chave, A.D., Evans, R.L., Hirth, G. & Mackie, R.L., 2006. Mantle dynamics beneath the East Pacific Rise at 17°S: insights from the mantle electromagnetic and tomography (MELT) experiment, *J. geophys. Res.,* **111,** B02101, doi:10.1029/2004JB003598.

Bostick, F.X., 1977. A simple almost exact method of MT analysis, Workshop on Electrical Methods in Geothermal Exploration, *U.S. Geol. Surv.,* Contract No. 14080001-8-359.

Chave, A.D., 2012. Estimation of the magnetotelluric response function, in *The Magnetotelluric Method: Theory and Practice,* pp. 165–218, eds Chave, A.D. & Jones, A.G., Cambridge Univ. Press.

Chave, A.D. & Thomson, D.J., 2004. Bounded influence estimation of magnetotelluric response functions, *Geophys. J. Int.,* **157**(3), 988–1006.

Chave, A.D. & Lezaeta, P., 2007. The statistical distribution of magnetotelluric apparent resistivity and phase, *Geophys. J. Int.,* **171**(1), 127–132.

David, H.A. & Nagaraja, H.N., 2003. *Order Statistics,* 3rd edn, John Wiley and Sons.

Egbert, G.D., 1997. Robust multiple-station magnetotelluric data processing, *Geophys. J. Int.,* **130**(2), 475–496.

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

Groom, R.W. & Bailey, R.C., 1989. Decomposition of magnetotelluric impedance tensors in the presence of local three-dimensional galvanic distortion, *J. geophys. Res.,* **94**(B2), 1913–1925.

Groom, R.W. & Bailey, R.C., 1991. Analytic investigations of the effects of near-surface three-dimensional galvanic scatterers on MT tensor decompositions, *Geophysics,* **56,** 496–518.

Jones, A.G., 1983. The problem of current channelling: a critical review, *Surv. Geophys.,* **6,** 79–122.

Jones, A.G. & Jödicke, H., 1984. Magnetotelluric transfer function estimation improvement by a coherence-based rejection technique, in *Proceedings of the 54th Annual International Meeting,* Soc. Expl. Geophys., Atlanta, GA, 2–6 December, Expanded Abstract, Vol. 3, pp. 51–55.

Jones, A.G., Chave, A.D., Egbert, G., Auld, D. & Bahr, K., 1989. A comparison of techniques for magnetotelluric response function estimation, *J. geophys. Res.,* **94**(B10), 14 201–14 213.

Jones, A.G. *et al.*, 2009. Area selection for diamonds using magnetotellurics: examples from southern Africa, *Lithos,* **112S,** 83–92.

Kitagawa, G., 1979. On the use of AIC for the detection of outliers, *Technometrics,* **21,** 193–199.

Kitagawa, G., 1981. The corrigenda of "On the use of AIC for the detection of outliers", *Technometrics,* **23,** 320–321.

Kitagawa, G. & Akaike, H., 1982. A quasi Bayesian approach to outlier detection, *Ann. Inst. Stat. Math.,* **34,** 389–398.

Ledo, J., 2005. 2-D versus 3-D magnetotelluric data interpretation, *Surv. Geophys.,* **26,** 511–543. (Corrected version in Ledo, J., 2006. 2-D versus 3-D magnetotelluric data interpretation, *Surv. Geophys.*, **27,** 111–148.)

Ledo, J., Queralt, P., Marti, A. & Jones, A.G., 2002. Two-dimensional interpretation of three-dimensional magnetotelluric data: an example of limitations and resolution, *Geophys. J. Int.,* **150**(1), 127–139.

Matsuno, T. *et al.*, 2010. Upper mantle electrical resistivity structure beneath the central Mariana subduction system, *Geochem. Geophys. Geosyst.,* **11,** Q09003, doi:10.1029/2010GC003101.

McNeice, G.W. & Jones, A.G., 2001. Multisite, multifrequency tensor decomposition of magnetotelluric data, *Geophysics,* **66,** 158–173.

Michael, J.R., 1983. The stabilized probability plot, *Biometrika,* **70,** 11–17.

Muller, M.R. *et al.*, 2009. Lithospheric structure, evolution and diamond prospectivity of the Rehoboth Terrane and western Kaapvaal Craton, southern Africa: constraints from broadband magnetotellurics, *Lithos,* **112S,** 93–105.

Niblett, E.R. & Sayn-Wittgenstein, C., 1960. Variation of electrical conductivity with depth by the magneto-telluric method, *Geophysics,* **25,** 998–1008.

Park, S.K. & Mackie, R.L., 2000. Resistive (dry?) lower crust in an active orogen, Nanga Parbat, northern Pakistan, *Tectonophysics,* **316,** 359–380.

Parker, R.L., 2010. Can a 2-D MT frequency response always be interpreted as a 1-D response? *Geophys. J. Int.,* **181**(1), 269–274.

Parker, R.L. & Booker, J.R., 1996. Optimal one-dimensional inversion and bounding of magnetotelluric apparent resistivity and phase measurements, *Phys. Earth planet. Inter.,* **98,** 269–282.

Parker, R.L. & Whaler, K.A., 1981. Numerical methods for establishing solutions to the inverse problem of electromagnetic induction, *J. geophys. Res.,* **86**(B10), 9574–9584.

Porsani, M.J., Niwas, S. & Ferreira, N.R., 2001. Robust inversion of vertical electrical sounding data using a multiple reweighted least-squares method, *Geophys. Prospect.,* **49,** 255–264.

Pynnönen, S., 1992. Detection of outliers in regression analysis by information criteria, in *Proceedings of the University of Vaasa,* Vaasa, Finland, Discussion Papers 146, 8 p.

Rodi, W. & Mackie, R.L., 2001. Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion, *Geophysics,* **66,** 174–187.

Santos, A.B., Sampaio, E.E.S. & Porsani, M.J., 2005. A robust two-step inversion of complex magnetotelluric apparent resistivity data, *Studia Geophysica et Geodaetica,* **49,** 109–125.

Stephens, M.A., 1974. EDF statistics for goodness of fit and some comparisons, *J. Am. Stat. Assoc.,* **69,** 730–737.

Stirling, D.W., 1982. Enhancements to aid interpretation of probability plots, *J. R. Stat. Soc. Ser. C (Statisticians),* **31,** 211–220.

Tango, T., 1986. Estimation of normal ranges of clinical laboratory data, *Stat. Med.,* **5,** 335–346.

Varentsov, Iv.M., 2002. A general approach to the magnetotelluric data inversion in a piecewise-continuous medium, *Izvestiya, Phys. Solid Earth,* **38,** 11–33.

Wannamaker, P.E., Hohmann, G.W. & Ward, S.H., 1984. Magnetotelluric responses of three-dimensional bodies in layered earths, *Geophysics,* **49,** 1517–1533.

Weidelt, P. & Chave, A.D., 2012. The magnetotelluric response function, in *The Magnetotelluric Method: Theory and Practice,* pp. 122–164, eds Chave, A.D. & Jones, A.G., Cambridge Univ. Press.

Weidelt, P. & Kaikkonen, P., 1994. Local 1-D interpretation of magnetotelluric B-polarization impedances, *Geophys. J. Int.,* **117**(3), 733–748.

Wessel, P. & Smith, W.H.F., 1998. New, improved version of the Generic Mapping Tools released, *EOS Trans. AGU,* **79,** 579.