Geophysical Journal International

Geophys. J. Int. (2015) **201,** 1765–1780
GJI Marine geosciences and applied geophysics

doi: 10.1093/gji/ggv107

The advantages of logarithmically scaled data for electromagnetic inversion

Brent Wheelock,[1] Steven Constable[2] and Kerry Key[2]

[1] *ExxonMobil Upstream Research Company, Houston, TX* 77389, *USA. E-mail: brentwheelock@gmail.com*
[2] *Scripps Institution of Oceanography, La Jolla, CA* 92093, *USA*

Accepted 2015 March 2. Received 2015 February 22; in original form 2014 October 06

SUMMARY

Non-linear inversion algorithms traverse a data misfit space over multiple iterations of trial models in search of either a global minimum or some target misfit contour. The success of the algorithm in reaching that objective depends upon the smoothness and predictability of the misfit space. For any given observation, there is no absolute form a datum must take, and therefore no absolute definition for the misfit space; in fact, there are many alternatives. However, not all misfit spaces are equal in terms of promoting the success of inversion. In this work, we appraise three common forms that complex data take in electromagnetic geophysical methods: real and imaginary components, a power of amplitude and phase, and logarithmic amplitude and phase. We find that the optimal form is logarithmic amplitude and phase. Single-parameter misfit curves of log-amplitude and phase data for both magnetotelluric and controlled-source electromagnetic methods are the smoothest of the three data forms and do not exhibit flattening at low model resistivities. Synthetic, multiparameter, 2-D inversions illustrate that log-amplitude and phase is the most robust data form, converging to the target misfit contour in the fewest steps regardless of starting model and the amount of noise added to the data; inversions using the other two data forms run slower or fail under various starting models and proportions of noise. It is observed that inversion with log-amplitude and phase data is nearly two times faster in converging to a solution than with other data types. We also assess the statistical consequences of transforming data in the ways discussed in this paper. With the exception of real and imaginary components, which are assumed to be Gaussian, all other data types do not produce an expected mean-squared misfit value of 1.00 at the true model (a common assumption) as the errors in the complex data become large. We recommend that real and imaginary data with errors larger than 10 per cent of the complex amplitude be withheld from a log-amplitude and phase inversion rather than retaining them with large error-bars.

Key words: Inverse theory; Magnetotellurics; Marine electromagnetic.

1 INTRODUCTION

Inversion is a central activity within the field of geophysics. It is the method by which measurements, often taken at or near the surface of the earth, are converted into a model for the distribution of material properties within the earth. This task of producing an earth model is often posed as an optimization problem, with a primary objective of reproducing, with some prescribed level of fidelity, a finite set of measurements. In the parlance of linear vector spaces, these measurements may be arranged into a data vector, $\mathbf{d} \in \mathbb{R}^N$, and the distribution of earth properties into a model vector, $\mathbf{m} \in \mathbb{R}^M$. Then, given a mapping which transforms a point in model space into a point in data space, the general goal of inversion is to find the point $\mathbf{m}$ that maps as closely as possible to $\mathbf{d}$; the mapping is provided by the specific physics involved in the measurement, and is appropriately called the forward calculation. The data vector may take multiple forms. All forms are derived from the same set of observations, but each may change the mapping, in turn changing the nature of the inverse problem. The only requirement in selecting the form in which a set of data will be inverted is that the forward calculation and the optimization method are modified accordingly.

Many geophysical processes are described in the form of complex Fourier coefficients of some transfer function, estimated from multiple time-series of observations. Such data can then be analysed in various forms, each packaged as pairs of complementary components. The basic three options we consider in this paper are: (1) real and imaginary components; (2) amplitude, raised to any power greater than zero, and phase; and (3) logarithmic amplitude

and phase, following from the definition of the logarithm of a complex number, where if $z = r\exp(i\theta)$, then $\ln z = \ln r + i\theta$, for $-\pi < \theta \leq \pi$. With magnetotelluric (MT) measurements, for example, these three forms translate to real and imaginary impedance, apparent resistivity and phase and logarithmic apparent resistivity and phase. All three forms are related by simple closed-form expressions, yet they differ, sometimes dramatically, in their effect on an inversion.

In this paper, we compare the effect that these different forms of data have on the course of an inversion, using synthetic MT and controlled-source electromagnetic (CSEM) data as examples, both treated in the Fourier-domain. The main MT data type we will discuss is impedance, a rank 2 tensor relating the naturally time-varying horizontal electric and horizontal magnetic field vectors measured at a point on the surface of the earth. In the CSEM example, each datum will be a single horizontal component of the electric field, measured at some offset from, and parallel to the moment of, a powered, time-varying, electric-dipole transmitter. Our aim is to determine the optimal form for these complex data by measure of the qualities that lead to fast and robust inversion. While the following illustrations make use of electromagnetic data, the statistical treatment in Section 4 is universal, and the advantages we discuss in other sections may extend to the inversion of other types of observations which exhibit a large dynamic range, either based in the frequency- or time-domain.

2 A SYNTHETIC MT STUDY

To highlight the differences between our three data forms, we wish to investigate conditions that impose high demands on non-linear inversion. We will use MT data which contain discontinuities and exhibit large dynamic range across frequencies. These are the sort of exotic symptoms that a coastline can produce in MT data. The 'MT coast effect', as it has become known, is a phenomenon observed particularly in marine MT data, in the vicinity of the coastline of a relatively resistive continent (e.g. Cox 1980; Constable *et al.* 2009; Key & Constable 2010; Worzewski *et al.* 2012). Within 2-D earth models, the following have been established as signatures of the MT coast effect: TE-mode phases that progress through all four quadrants, wrapping from -180° to 180° at some characteristic frequency; a cusp in the TE-mode apparent resistivity centered on the same characteristic frequency; and a broad-band depression of the TM-mode apparent resistivity relative to the expected value in the absence of a coast (Fig. 2). Historically, MT impedance phases that fell outside of the first quadrant had been considered unusual at the very least, if not cause to suspect 3-D structure or anisotropy in the subsurface (e.g. Egbert 1990; Weidelt & Kaikkonen 1994). Lately, 'out-of-quadrant' behaviour has been shown to be typical of thinly sedimented coastal marine data, and numerical modelling studies have proven that isotropic 2-D structure can yet permit it. Key & Constable (2010) calculate the Poynting vector of MT energy near the coastline in a 2-D model and illustrate that the TE-mode coast effect is a result of a localized reversal in the diffusion direction from downward to upward. Similar behaviour has also been observed in land-based MT data at the lateral edge of a thin surface conductor, and again 2-D numerical modelling confirmed it to be caused by a reversal in the standard MT diffusion direction (Selway *et al.* 2012). In this section, we will use synthetic MT measurements from a simple model purposefully designed to mimic the strong MT coast effect observed in a marine data set collected off the coast of central California (Wheelock 2013). We will then show that the wide dynamic range and extremely non-linear behaviour observed in TE-mode MT data call for special treatment in an inversion process, favouring one form of data above the others.

Fig. 1 depicts the simple 2-D coast model we use in this study of data forms. It has a gradual continental slope which levels out to a 4 km deep ocean, with seawater conductivity of 0.3 Ωm. The model has a uniform subseafloor resistivity which we will vary over multiple forward calculations. At each of the 20 seafloor sites we will calculate the MT quantities of impedance tensor, $\mathbf{Z}$, and vertical magnetic transfer function, or tipper, $\mathbf{V}$ [for more background on these MT measurements, consult Chave & Jones (2012)].

Figs 2 and 3 contain plots of two separate calculations from our simple coast model using two different seafloor resistivities, 90 and 115 Ωm [the calculations were made with the finite-element algorithm described in Wannamaker *et al.* (1987)]. These figures also show the 1-D values which correspond to models with the same seafloor resistivities but with no coast. Without dwelling on the physics of the coast effect, the point we make here is that coastlines create unique effects in the apparent resistivity, phase, and tipper data which vary in an extremely non-linear way with respect to the conductivity of the earth; this is particularly true with the TE-mode impedance and tipper. For example, Fig. 4 shows the variation of the TE-mode apparent resistivity at site 5, about 45 km away from

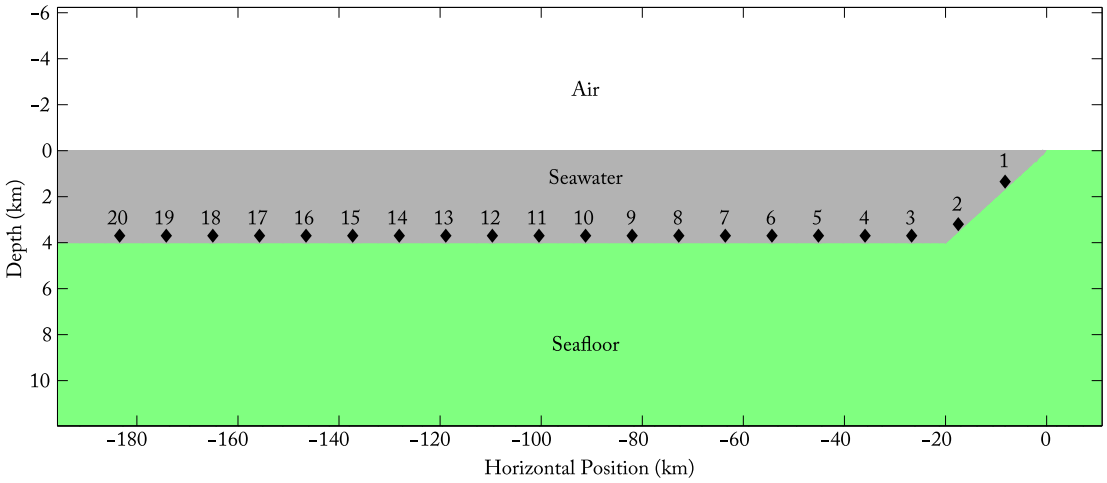

Figure 1. Model with a simple coast and uniform seafloor used in synthetic misfit study. The black diamonds indicate locations of hypothetical MT receivers labelled by site number.

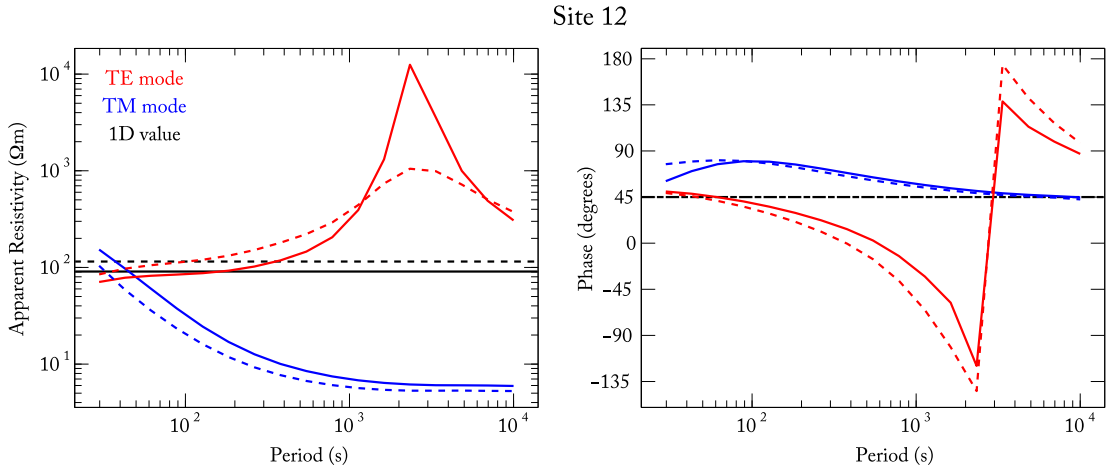

Figure 2. MT apparent resistivity and phase calculations for site 12 in the model pictured in Fig. 1. Two uniform seafloor resistivities were used: 90 Ωm (solid lines) and 115 Ωm (dashed lines). The black lines plot the respective values for models with equivalent seafloor resistivities but no coast (i.e. purely 1-D), in which case TE and TM modes are equivalent, and the phase is always 45°.

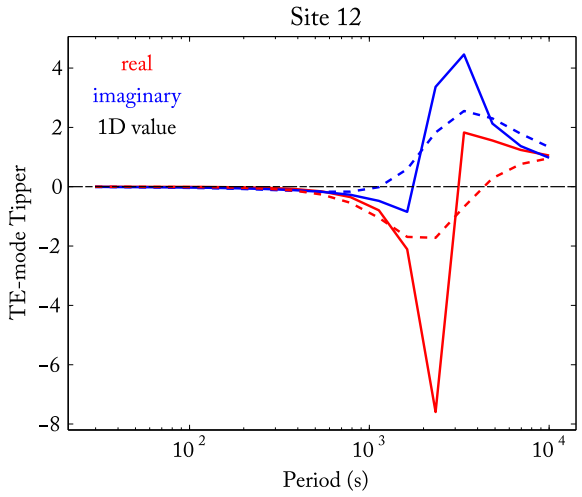

Figure 3. MT tipper calculations for site 12 in the model pictured in Fig. 1. Two uniform seafloor resistivities were used: 90 Ωm (solid lines) and 115 Ωm (dashed lines). The black lines plot the respective values for models with equivalent seafloor resistivities but no coast (i.e. purely 1-D), in which case all tippers are exactly zero.

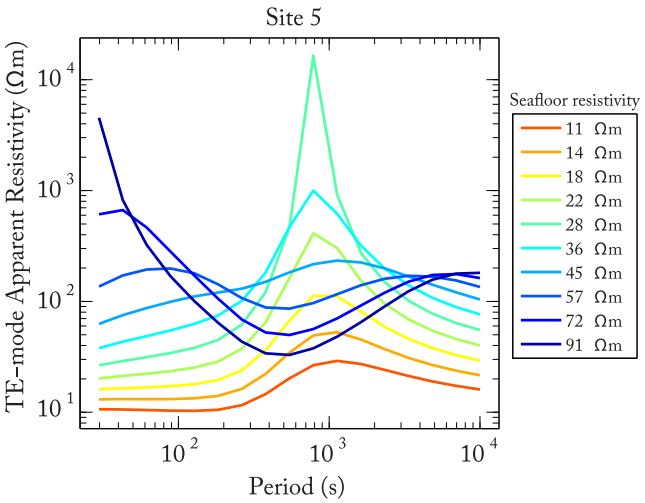

Figure 4. Variation in TE-mode apparent resistivity, at site 5 in Fig. 1, as a function of seafloor resistivity.

the coastline, as a function of seafloor resistivity. The progression in Fig. 4 highlights that the functional derivative with respect to resistivity is not monotonic for all frequencies. Such variable behaviour creates difficulty for most non-linear inversion algorithms which rely on a linear approximation to predict how a change in any given model parameter will affect a fit to the data. In contrast, the functional derivative with respect to the subsurface parameter for the analogous model without a coast (e.g. the black lines in Figs 2 and 3) can be shown to be monotonic (and, in the case of apparent resistivity, frequency independent), and therefore more predictable. Accordingly, we might guess that the a coast-free version of this single-parameter system presents far less of a challenge for an inversion algorithm.

In Section 2.2, we employ the Occam inversion algorithm (Constable *et al.* 1987), which, along with other non-linear deterministic algorithms (e.g. NLCG, ref. Appendix A), finds an optimal model by traversing the misfit space in a consistently downhill fashion (i.e. always in a direction of decreasing misfit). Due to this manner of progression, there are a few properties of misfit spaces that improve the performance of deterministic inversion schemes. We list these properties in the following paragraphs.

First, for the sake of argument, let us make the common assumption that the 'true' model maps into the global minimum in misfit space, regardless what form the data take. Given that, a deterministic inversion scheme has the best chance of recovering the 'true' model when a chosen misfit space, determined by the form of the data and the misfit functional, has the fewest number of minima, preferably only one. This is because deterministic inversion algorithms have no explicit mechanism for escaping local minima, as would, for example, a stochastic inversion method which may accept models in successive iterations that amount to an increase in misfit. Deterministic inversion methods greedily march in the direction of decreasing misfit until such a direction can no longer be found (i.e. the conditions of a local minimum); under such conditions, the current position in model-space is accepted as the final solution. However, it should be noted that the Occam algorithm is not necessarily impeded by a local minimum in data misfit because, within each iteration, Occam is not beholden to a single objective-function space. This is the space comprised of a weighted sum of the misfit-norm and the model-norm, where the relative weighting

is provided by a Lagrange (or 'trade-off') parameter. By varying the Lagrange parameter within each iteration, the Occam algorithm repeatedly changes which objective function it probes, providing an opportunity to lever itself out of what might be a local minimum in the data misfit space on its own. While it is not guaranteed that all model norms will provide the proper lever, the lost possibility of dodging local minima provides a strong argument against holding the Lagrange parameter fixed. Nonetheless, with fewer local minima in the data misfit space, *any* algorithm's prospects for finding the true model are improved.

Secondly, the uncertainties of the final model parameters resulting from inversion are smaller when the misfit surface surrounds the global minimum with steep walls; this is the literal origin of the phrase 'tightly constrained'. The steeper the walls in the misfit surface, the greater the increase in misfit as the model parameters stray from the final value. This leaves a narrower range of model space which may produce practically the same fit to the data, bestowing greater certainty to the final model.

Lastly, the misfit value of the true model must meet a statistically significant value, given the probability distributions of the data, so that one is aware when they have achieved a probable fit to the data set. This, of course, hinges upon the adequacy of the preceding estimation of data errors. In practice, the residuals from an inversion's final data fit rarely conform to the co-distribution of random errors assumed to exist in the data before inverting them. When this happens, either an improper statistical model was assumed or an improper physical model was used to describe the system which created the data. Often, it is the latter deficiency which affects geophysical inversion; finite computer simulations can rarely encompass the full detectable complexity of the earth. Acknowledging this, the goal in this paper is to isolate the former of these two effects through synthetic inversion, in which the physical model is guaranteed to be complete. We seek to understand how our three forms of data influence statistical expectations, independent of the physics which underly a real measurement, and the degree of accuracy with which they are modelled.

2.1 1-D misfit space

We now examine plots of the misfit spaces derived from our simple coast model with data packaged in a variety of possible subsets and forms. The intent is to appraise the misfit spaces of these different data packages in terms of the aforementioned properties that facilitate success for a deterministic algorithm. The following misfit spaces are 1-D, that is they vary with only one model parameter ($M = 1$). Low-dimensionality, while not representative of regularized inversion, where $M \gg N$ with N representing the number of data, makes the misfit spaces amenable to plotting and analysis. We expect some of the differences observed between data types to carry over to higher dimension problems, as we will show with some highly parametrized synthetic inversions in Section 2.2.

The misfit spaces are computed in the following way. We choose one seafloor resistivity to be the true model. MT forward calculations are made for the true model and independent Gaussian noise is added to the real and imaginary components of the impedance tensors and the tipper vectors to create synthetic data. The standard deviation of the pseudo-random, Gaussian errors are scaled to a percentage of amplitudes $|\mathbf{Z}|$ and $|\mathbf{V}|$, for $\mathbf{Z}$ and $\mathbf{V}$, respectively; this ensures that the real and imaginary parts of each synthetic datum are jointly distributed with a probability distribution that is circularly symmetric in the complex plane, an assumed characteristic of measured data. Note there are cases of geophysical measurements where this assumption does not hold (e.g. Myer *et al.* 2012). Next, MT forward calculations are made over many models that span a range of seafloor resistivities. These serve as the model predictions made by the forward functionals, $f_i[\mathbf{m}]$ for $i = 1, \ldots, N$. For each of these predictions, a normalized root-mean-square (rms) misfit is computed relative to our synthetic data. The rms misfit is defined as:

$$X_{\mathrm{rms}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i - f_i[\mathbf{m}])^2/\sigma_i^2}, \tag{1}$$

where d_i are the data, and σ_i estimates of their standard deviation. Making use of the formula for the expectation value $\mathrm{E}[||\mathbf{X}||]$, derived for the standardized random vector $\mathbf{X}$ in Parker (1994), we may bound the expected misfit with

$$\mathrm{E}[X_{\mathrm{rms}}] \leq 1 - \frac{1}{4N} + \frac{1}{32N^2} + O(N^{-3}). \tag{2}$$

For each of the data forms, we compare the misfit spaces of their individual data components. The total number of MT data used in this example produces an $N = 2040$. For each conglomeration of data in subsequent plots, N is large enough that we expect the rms misfit to be very close to 1 at the true model. Yet, (2) assumes the elements of $\mathbf{X}$ are Gaussian-distributed. While we assume this is true for complex impedance and tipper data, in Section 4 we show that with our transformed data this assumption progressively deteriorates as the relative magnitude of noise increases. Nonetheless, for the levels of noise added to the data in this section, the rms misfit at the true model should still be very close to 1.00.

Figs 5–7 depict the results of the misfit space calculations. The initial observation, common to all plots of the TE mode (left column in Figs 5 and 6), is that the solid red lines are much more jagged with multiple local minima than the other curves; the solid red lines of the TM mode have local minima as well, but far fewer, making their curves much smoother. The red lines relate to impedance data in the form of apparent resistivity, which we have shown form steep cusps in the TE mode spanning many orders of magnitude (Figs 2 and 4). By the criteria just discussed, apparent resistivity data are the worst option for MT inversion near a coast, and we infer the cause to be the large dynamic range, and unique non-linear nature of the TE mode. The solid orange lines, representing absolute tipper magnitude, are also quite rough, perhaps unsurprizing in light of the similarly dramatic variation of responses seen in Fig. 3. The solid dark blue and solid light blue lines, which represent the misfit spaces for the real and imaginary components of the impedance tensor and the tipper, respectively, also have multiple local minima, but these are less extreme than those of apparent resistivity and tipper amplitude. For this simple coastal model, we conclude that the real and imaginary components of the impedance tensor and tipper vector are the second worst option for inversion of coast-affected data. Note that the phase misfit spaces, for both tipper and impedance, have no such rash of local minima as in the case with the amplitude related data.

Once apparent resistivity and tipper amplitude are put into the logarithmic form, the jaggedness of their misfit spaces are reduced almost entirely, with one exception on the resistivity line. There remains a solitary positive cusp in the logarithmic tipper amplitude line, which occurs exactly at the resistivity of the seawater, coincident with a marked discontinuity in the tipper phase. As the seafloor resistivity approaches that of the seawater, the tipper amplitude goes to zero and its phase is undefined. Also, once the seafloor resistivity

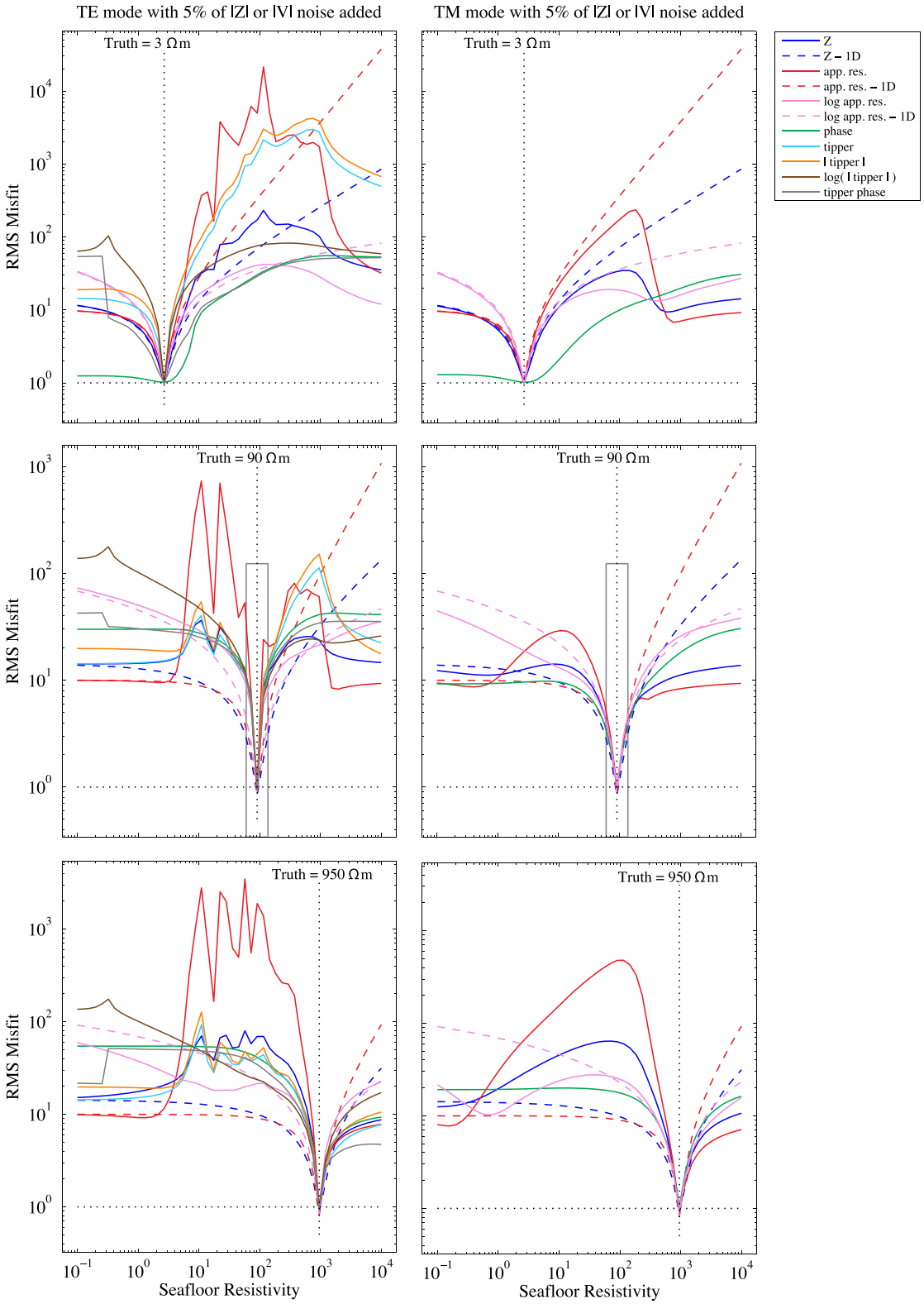

Figure 5. MT misfit spaces for different seafloor models and different forms of data (solid and long-dashed lines). The grey boxes in the middle row of plots outlines the range of models plotted in Fig. 6. Curves plotted in long-dashed lines are for the coast-free (1-D) case.

become less resistive than the seawater, the sign of the tipper flips, resulting in a sudden increase in its phase misfit. As we rarely expect the seafloor to be more conductive than the seawater above it, we do not see this cusp in misfit to be too much of a problem.

Another peculiarity is the shape of the impedance phase misfit curves (green lines) for the 3 Ωm true model (top row in Fig. 5), which is very different than that of the other more resistive true models. It does not have the sharp v-shaped minimum like the

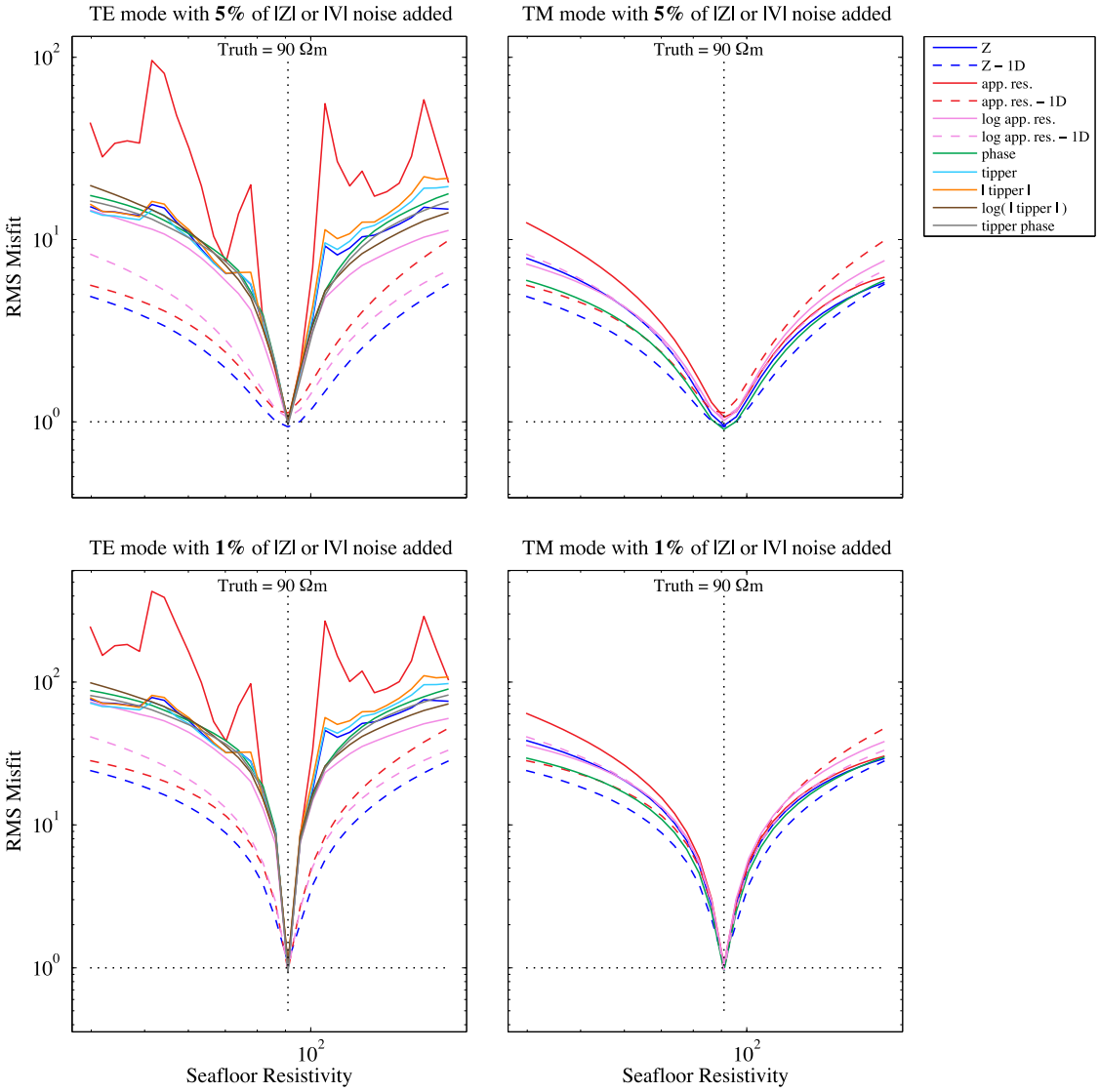

Figure 6. Top row: close-up of portions of the middle plots in Fig. 5 delineated by the grey box. Bottom row: same as the top row, except only 1 per cent noise was added to the synthetic data as opposed to 5 per cent. Lines and symbols follow the same conventions as in Fig. 5.

other data types. The phase misfit remains very small on the end of parameter space more conductive than 3 Ωm. This is because a 3 Ωm seafloor is not enough of a contrast with the seawater to produce a 'coast effect'. The coast is the only 2-D feature in our simple model. Without its 2-D distortion, the phase behaviour reverts back to a half-space response which is always 45°. Hence, there is very little model discrimination by phase data alone when the seafloor in both the true model and trial model are both filled with low resistivity.

Next, we note that all of the dashed curves in Figs 5–7, while derived using the same data forms as their solid-lined counterparts, are comprised of MT responses over a 1-D half-space. Without a coast effect, all of these misfit spaces are smooth on both sides of the absolute minimum and contain no local minima. As we suspected, the coast-free inverse problem is a far simpler one. Yet it should be noted that on the end of model space that is far more conductive than the true model, the misfit curves of impedance and apparent resistivity flatten, even without a coast, giving an inversion algorithm no clear direction; we will return to this phenomenon later. By nature of the smoothness of the misfit space, the uni-parameter coast-free inverse problem is guaranteed to converge on the true solution, regardless of the form that the data take, as long as the downhill search begins at a position in model space that is more resistive than the true model; this is not guaranteed once a coast is added to the model. However, we also note that the walls of the valleys in the misfit curves around the true models are steeper for the coast-affected data than they are for the coast-free data (Fig. 6). This reveals that the coast effect, in particular that of the TE mode, better constrains the seafloor resistivity. This is a consequence of the increase in wavenumber-richness of the electromagnetic fields that have been distorted by the shape of the land-sea interface (Wheelock 2013). Consider the fact that in our simple model no new information is gained by changing the location of the receiver when no coast (and no topography) is present; under the standard plane-wave source field, MT data from a 1-D earth consist of a single wavenumber, $k = 0$. Yet with a coast, energy from the same plane-wave source is dispersed into the higher wavenumbers, each providing an independent probe of the subsurface. This increase in independent constraints

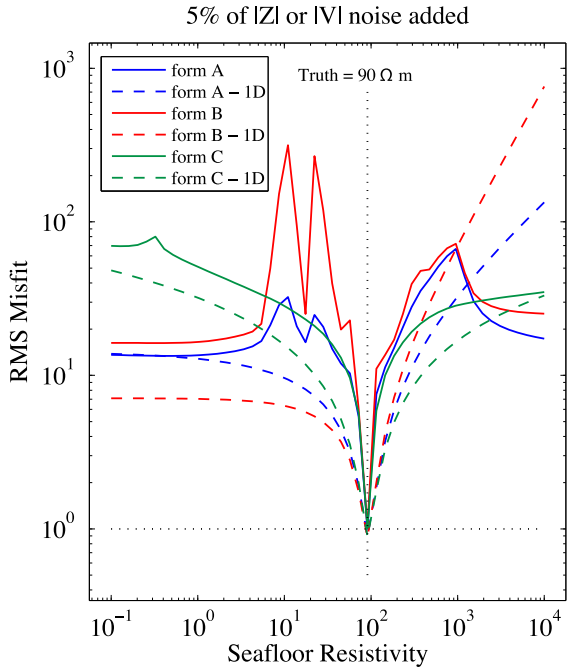

Figure 7. Misfit spaces for three common data ensembles: form A contains real and imaginary components of both the impedances and the tippers; form B contains apparent resistivities, impedance phases and the real and imaginary components of the tipper; and form C contains logarithmic apparent resistivities, impedance phases, logarithmic tipper amplitudes and tipper phases.

produced by the coast effect results in greater certainty about the subsurface.

The difference between the top and the bottom rows of plots in Fig. 6 is the magnitude of the Gaussian errors introduced into the synthetic data; the top row was made with a standard deviation of noise five times greater than that of the bottom row. The valleys in the misfit surfaces around the true model for data with smaller noise characteristics are much narrower and steeper than that with the larger noise characteristics. As linearized error propagation would predict, the lower the magnitude of data errors the better constrained, or the less uncertain, the final parameter value.

Rarely does one invert the TE mode without the TM mode, or apparent resistivity without impedance phase; in Fig. 5 we see flat zones or local minima in one data type that do not occur in its complimentary type, providing evidence for why inversion with an isolated data component might be a bad idea. So, in Fig. 7 we examine the joint misfit curves for three fundamental MT data ensembles that make use of all complimentary components: form A combines the real and imaginary components of both the impedance and the tipper; form B combines apparent resistivity and phase of the impedance with the real and imaginary components of the tipper; and form C combines the logarithmic apparent resistivity and phase of the impedance with the logarithmic amplitude and phase of the tipper. Note that the solid-line curves in Fig. 7 are comprised of misfits from both the TE and TM modes of impedance. The dashed-line curves in Fig. 7, however, derive from the purely 1-D model where the distinction of TE and TM mode dissolves and there is no tipper response (Figs 2 and 3).

We find that for both the 1-D data and the 2-D coast-affected data, only the form C, again that which involves logarithmic amplitudes, produces a misfit space compliant with a deterministic inversion. In Fig. 7, we see that the curves for data forms A and B, which do not involve logarithmic amplitudes, flatten as the model prediction moves towards a more conductive seafloor; this does not happen with the form C. To explain this flattening, we look first at the 1-D half-space curves where we can ignore impedance phase, which is always 45° regardless of seafloor resistivity, and the tippers which are always zero. We note that for the impedance and apparent resistivity data forms over a 1-D half-space, $f_i[\mathbf{m}] \to 0$ as $||\mathbf{m}|| \to 0$, where $\mathbf{m}$ represents the resistivity-depth profile of the earth. Now, let us analyse how this limit affects our different measures of misfit. For the non-log-scaled misfit we write,

$$\lim_{f_i \to 0} X^2 = \lim_{f_i \to 0} \sum_{i=1}^{N} \left(\frac{d_i - f_i[\mathbf{m}]}{\sigma_i} \right)^2 = \sum_{i=1}^{N} \left(\frac{d_i}{\sigma_i} \right)^2 = \text{const.} \qquad (3)$$

This explains the flattening in the misfit curves when the seafloor resistivity is much lower than that of the true model. The same flattening is seen in the models with a coast because the coast effect in impedance data vanishes as the seafloor resistivity approaches that of seawater; thus the half-space behaviour, $f_i[\mathbf{m}] \to 0$, again prevails in the 2-D setting as seafloor resistivity moves towards zero. Next, we write the definitions of the log-scaled data, model predictions and error estimates in terms of their non-log-scaled counterparts. With the log-scaled values denoted by the $\hat{\cdot}$, we have

$$\hat{d}_i = \log_{10} d_i, \qquad (4)$$

$$\hat{f}_i[\mathbf{m}] = \log_{10} f_i[\mathbf{m}], \qquad (5)$$

and

$$\hat{\sigma}_i = \frac{\sigma_i}{d_i \ln 10}, \qquad (6)$$

where (6) is found through linear propagation of error, which we discuss in Section 4.1. Then, the log-scaled misfit measure is

$$\begin{aligned} \hat{X}^2 &= \sum_{i=1}^{N} \left[\frac{\hat{d}_i - \hat{f}_i[\mathbf{m}]}{\hat{\sigma}_i} \right]^2 \\ &= \sum_{i=1}^{N} \left[\frac{(\log_{10} d_i - \log_{10} f_i[\mathbf{m}])(d_i \ln 10)}{\sigma_i} \right]^2 \\ &= \sum_{i=1}^{N} \left[\left(\frac{d_i}{\sigma_i} \right) \ln \left(\frac{d_i}{f_i[\mathbf{m}]} \right) \right]^2 . \end{aligned} \qquad (7)$$

Taking the limit again, we find

$$\lim_{f_i \to 0} \hat{X}^2 = \infty. \qquad (8)$$

With the log-scaled inverse problem, there is no flattening in the misfit curve on the conductive end of the parameter space. Conversely, we analyse the different misfit behaviours as the seafloor becomes more resistive. By noting that for the 1-D half-space, $f_i[\mathbf{m}] \to \infty$ as $||\mathbf{m}|| \to \infty$, we find

$$\lim_{f_i \to \infty} X^2 = \infty, \qquad (9)$$

and

$$\lim_{f_i \to \infty} \hat{X}^2 = \infty. \qquad (10)$$

For the current example we have shown that the gradient of the log-scaled misfit is non-vanishing at both extremes in parameter space, whereas the non-log-scaled misfit function is flat in the low end of parameter space. Flat misfit curves lead to divergence in a Jacobian or gradient-based inversion process as it is very difficult to

detect, within the limits of numerical precision, which direction in parameter space leads downhill in misfit; in other words, if the misfit gradient vector vanishes, no update is made to the model and no progress can be made towards a lower misfit. This behaviour at the limits of parameter space provides another argument for choosing the logarithm of apparent resistivity and the logarithm of tipper magnitude as the optimal forms in which to invert MT data. Shin & Min (2006) found similar benefits to using logarithmic misfit measures for seismic wavefield inversion.

Other non-standard misfit measures have been proposed to improve the robustness of inversion. The 'point-symmetric' misfit measure (Wilson *et al.* 2006), uses an *ad hoc* combination of each model prediction, f_i, with each datum, d_i, to normalize the corresponding residual in the following form,

$$X^2_{ps} = \sum_{i=1}^{N} \frac{(d_i - f_i[\mathbf{m}])^2}{1/2\left(d_i^2 + f_i^2[\mathbf{m}]\right)}. \tag{11}$$

It has been described as a way to weight data of dramatically different magnitude, say from different stations, time channels or frequencies, equally. Plotted as a function of $\log(f_i)$, this misfit formulation does achieve the objective of symmetry about the true solution, where $X^2_{ps} = 0$, but it has the undesirable quality of leveling towards $2N$ as the model predictions, f_i, stray far from the measurements, d_i; this misfit curve flattens at both ends of parameter space. Furthermore, the logarithmic misfit scaling is easier to justify as it is a model-independent way to equally weight diverse data, and does not compromise the statistical theory accompanying standardized residuals (Section 4).

Returning to (7), the only way the logarithmic form of data would lead to a flat misfit curve is when none of the $f_i[\mathbf{m}]$ change with a change in $\mathbf{m}$; we will encounter an example of this in Section 3. In this case, the data have no sensitivity to the model parameters, and the inverse problem begun in such a region of misfit space is inescapably ill-posed. Finally, we note that all forms of data make their closest approach to the 1.00 value of rms misfit at the true model. Thus, assuming one has a sufficient quantity of data with well-estimated data errors, the expected rms misfit of 1.00 holds unique significance in this single-parameter system, regardless of which data form is used. Again, we will show in Section 4.1 that this is not exactly true for the transformed data as the data errors become large, which suggests that it is better to remove relatively high-noise, transformed data from inversion, rather than continue to use them in a heavily down-weighted fashion.

2.2 2-D inversion

We close this MT section with a more practical illustration of how data forms affect the inverse problem. We perform synthetic, regularized inversions on each different data form generated from a highly parametrized model, as opposed to the single-parameter model addressed in the previous section. The 'true' model, replicated across the upper row of Fig. 8, is parametrized with 799 blocks below the same simple seafloor used previously; the inversion software is described in DeGroot Hedlin & Constable (1990). The background resistivity is 100 Ωm, with an anomalously conductive block of 5 Ωm. The estimated data errors, used for weighting in each inversion, were set using linear propagation formulae (Section 4) scaled to the variance of the Gaussian noise that was added to the components of raw impedance in the creation of the synthetic data. The lower three rows of plots in Fig. 8 give a synopsis of the final models resulting from inversions using different data forms, different quantities of synthetic noise, and different uniform

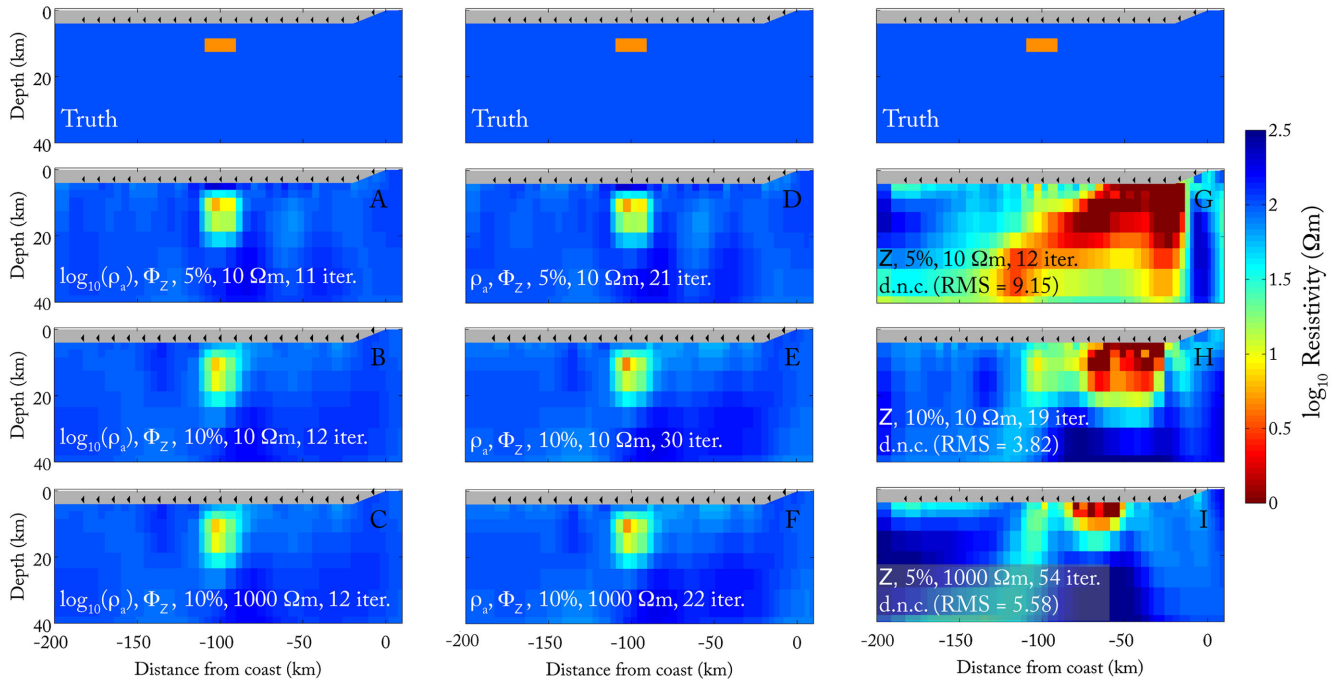

Figure 8. The true model and final models produced by synthetic inversion experiments with different data forms. The labels in each plot describe the setup for that particular inversion. The data forms used are logarithmic apparent resistivity ($\log_{10}(\rho_a)$), linear apparent resistivity (ρ_a), impedance phase ($\Phi_\mathbf{Z}$) and the real and imaginary components of the raw impedance ($\mathbf{Z}$). Synthetic Gaussian noise was added to the data for each inversion; their standard deviations are labelled in terms of percentage amplitude of the raw impedances, either 5 per cent or 10 per cent. Also, the labels report the uniform seafloor resistivity of the starting model, and the number of iterations required for each inversion to converge to the smoothest model with an rms misfit = 1.00, except for the cases where the inversion did not converge (d.n.c.), plateauing at a higher misfit.

starting models. Reported in each of these plots is the number of iterations required for the inversion to converge upon the smoothest model to fit the data to an rms misfit of 1.00; the designation 'd.n.c.' (did not converge) is used when a local minimum caused the inversion to founder at an rms misfit greater than 1.00.

Much of the behaviour encountered in these highly parametrized synthetic inversions confirms what was inferred from the single-parameter misfit curves in Section 2.1, but there are slight differences that are presumably caused by the increase in degrees of freedom. The logarithm of apparent resistivity paired with phase is found to be the optimal form for inversion of data distorted by 2-D bathymetry. The inversions using log-scaled data (Plots A, B & C) consistently achieve the expected value of rms misfit, recovering, in a smoothed form, models consistent with the true model; this is achieved regardless of the degree of noise added and regardless of the starting model used. Notice that the log-scaled inversion starting from a 10 Ωm half-space (Plot B) requires the same number of iterations as one starting from a 100 Ωm half-space (Plot C), demonstrating a symmetry in misfit behaviour about the true model similar to that observed for Form C of Fig. 7. This implies that an inversion of log-scaled data is nearly independent of the starting half-space value. A significant distinction for the log-scaled data form is that its inversion converges with the fewest iterations, about a factor of two fewer than the next best form of data. Thus, the logarithmic misfit measure is not only robust with respect to starting model, it is faster, too.

The next best form of data, contrary to our conclusions in Section 2.1, is linear apparent resistivity and phase, which also converged for all trials over noise levels and starting models (Plots D, E & F). Nonetheless, with this data form, larger components of random error hindered the inversion algorithm when the starting model was more conductive than the true background (Plot E). As the results of Section 2.1 would lead us to predict, when given a starting model more *resistive* than the true background, the inversion of the same high-noise data converged more rapidly (Plot F). It is observed that linearly scaled inversion is less robust with respect to the starting half-space resistivity, and again, the linear apparent-resistivity inversions required nearly twice as many iterations as their logarithmic counterparts. Lastly, the raw impedance (real and imaginary components) inversions failed at every attempt: all combinations of low noise, high noise, conductive starting model and resistive starting model. We hypothesize this is a consequence of the particular non-linearity the large dynamic range in the forward functionals caused by a strong coast effect, as inversions of synthetic raw impedance data from coast-free models converge routinely.

In Section 4.1, we show that use of data in the linear apparent resistivity form introduces a bias that is directly proportional to the variance of the Gaussian errors in the real and imaginary components of impedance. In contrast, there is no bias introduced with the logarithmic form of apparent resistivity, nor with the impedance phase, at least under a second-order approximation. Nonetheless, both the log-scaled and linearly scaled data have an expected value of rms misfit that deviates from the canonical 1.00 ever-increasingly as the relative errors in the impedance grow. One should be aware of this change in the expected rms misfit value for transformed data, particularly when synthetic data have relative errors greater than 10 per cent. We are reminded that, even with a high degree of control over the data errors, as with the synthetic studies performed here, a fanatical pursuit of an rms misfit of exactly 1 may be misguided depending upon the form of data inverted and the form in which the synthetic noise was added.

We conclude from these synthetic studies that an immediate improvement for 2-D inversion of marine MT data can be made by switching from using complex impedance elements, or linear apparent resistivity and phase, to logarithmic apparent resistivity and phase; improvement is also expected by switching to logarithmic amplitude and phase of the tipper data. In the absence of the coast effect, as with inland MT data, the smoothness of impedance curves implies that a 2-D inversion would not be dramatically affected by which form the data take. Nonetheless, there is still the flatness on the low-resistivity end of the misfit space that does not depend on the coast effect, and becomes a problem when the starting model is more conductive than the truth; this problem is only averted with log-scaled data.

2.3 Comments on 3-D inversion

As 3-D inversion becomes more widespread, where the use of all four elements of the impedance tensor (Z_{xx}, Z_{xy}, Z_{yx} and Z_{yy}) is required, the results presented here may be consequential. The diagonal elements of the impedance tensor do not inherit the intuitive interpretation that the form of apparent resistivity provides for the off-diagonal elements. Also, a logarithmic apparent resistivity and phase representation for the Z_{xx} and Z_{yy} impedance data, or equivalently, a logarithmic amplitude and phase representation of the V_{zx} or V_{zy} tipper data, seems unfit under certain resistivity model symmetries, where, in theory, these data are exactly zero. Under such circumstances, the logarithm would approach negative infinity, and the phase would be undefined. As a result, there has been a return to interpreting and inverting data in the form of real and imaginary impedance and tipper components in the growing number of 3-D MT studies (e.g. Newman & Alumbaugh 2000; Zhdanov *et al.* 2011; Tietze & Ritter 2013; Meqbel *et al.* 2014). Yet, with both actual data and synthetic data, there always exists some noise 'floor', which by definition never vanishes in its average amplitude, regardless of the vanishingly small amplitude of the ideal response; the diagonal impedance elements or tipper values will never truly be measured as zero. Correspondingly, the error-bars, or inverse data weights, must be fixed such that as the symmetries of the earth bring the diagonal impedance elements or tipper data toward extremely small values, they eventually lose all influence in the inversion. Although it is possible to accomplish this with ever-larger error-bars on log-scaled and phase data as they approach the noise floor, there are hidden consequences which may harm the performance of an inversion. We will return to this topic in Section 4, but for now discuss, for example, the consequences for phase. When the raw impedance or tipper data approach a floor in Gaussian noise, the fractional component of data errors become very large. This poses no problem for a least-squares (LS) inversion, which is often designed under the assumption of Gaussian-distributed data of a known, finite variance. However, once transformed into phase, these data are no longer Gaussian-distributed. Chave & Lezaeta (2007) show that as the error fraction of the raw impedance data become very large, the corresponding phase data tend towards being uniformly distributed. Thus, near the noise floor, phase does not conform to our inversion machinery, and moreover, with all values having nearly equal probability, phase loses its utility. For such reasons, we suggest removing both phase and log-scaled apparent resistivity data from an inversion, once their pre-transform relative error percentage exceeds a chosen value; we propose a 10 per cent threshold in Section 4.

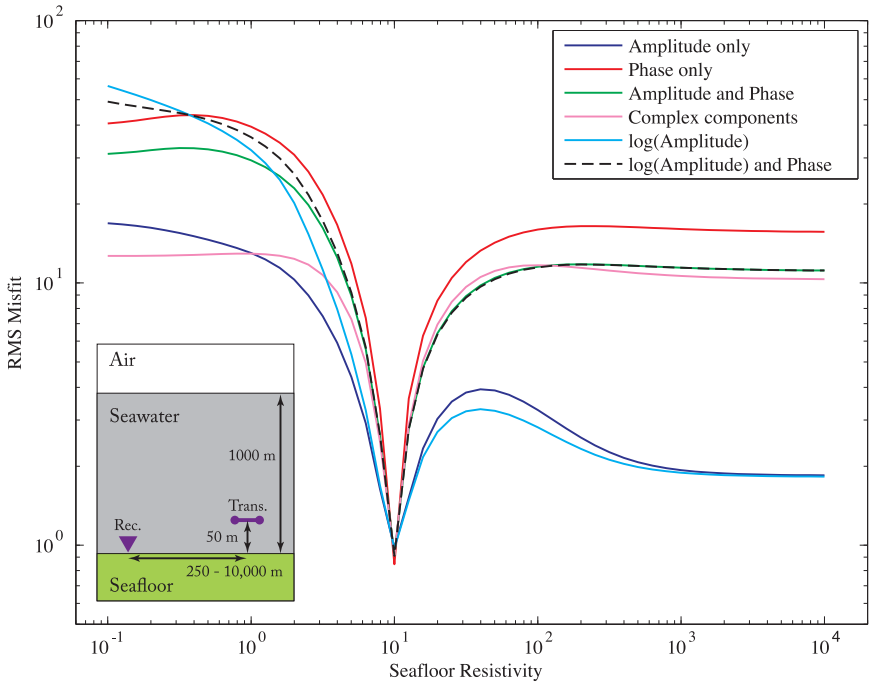

Figure 9. Marine CSEM misfit spaces for different forms of data and combinations thereof.

Once data with high relative-errors are culled from inversion, our preceding conclusions argue for the use of all four 3-D impedance elements in the form of logarithmic apparent resistivity and phase, as well as using logarithmic amplitude and phase with the tipper elements. We expect this to be especially true when the extreme effects of the coast are confronted. Yet log-scaled data may even offer an advantage in dealing with the wide dynamic range in amplitude between the diagonal and off-diagonal components of the impedance tensor, or the two components of the tipper vector. Finally, we remind the reader that from the perspective of the misfit functional, there is no difference between inverting MT data in the form of the logarithm of apparent resistivity and phase or the logarithm of impedance amplitude and phase; the exponent as well as the factors that do not change with the model, that is $1/(\omega\mu_0)$, all cancel out. It then seems natural that some of the benefits of log-scaled MT data should carry over to other complex data types which exhibit a large dynamic range. To bolster this premise, in the next section we turn to an illustration using CSEM data.

3 A SYNTHETIC CSEM STUDY

In this section, we again analyse a single-parameter misfit space, but this time pertaining to marine CSEM data. The forward calculations are made using the algorithm described in Key (2009). The misfit plots represent a CSEM data set consisting of the in-line electric fields from 40, evenly spaced, source–receiver offsets, ranging from 0.25 to 10 km, at a frequency of 0.25 Hz. Independent Gaussian noise with a standard deviation equal to 5 per cent of the electric field amplitude, plus a constant noise floor at 10^{-15} $\mathrm{VA}^{(-1)}$ $\mathrm{m}^{(-2)}$ was added to the real and imaginary components of the source-normalized fields. The model represents a simple deep-marine survey with the seafloor receiver residing in 1 km deep water and the transmitter towed 50 m above the flat seafloor.

To explore the single-parameter misfit space, the uniform resistivity below the seafloor is varied over many orders of magnitude. In Fig. 9, we see the results of this variation for an example where the 'true' seafloor resistivity was chosen to be 10 Ωm. Though not shown in this paper, the exercise which created Fig. 9 was repeated with different frequencies ranging between 0.1 and 10 Hz, even with multiple frequencies used simultaneously, and it was found that the relative character of the various misfit curves does not change. The same was true over trials with different 'true' seafloor resistivities, as well as different percentages of errors added to the data. Therefore, while we plot the misfit space for only one of these trials in Fig. 9, the analysis of the curves may broadly apply to CSEM data sets over simple structures.

The feature common to all forms of the CSEM data we use in this example is that their misfit curves flatten at the high resistivity end of parameter space. When the seafloor resistivity reaches 1000 Ωm, the skin depth of 0.25 Hz data is about 32 km, more than three times the greatest source-receiver offset in this synthetic CSEM data set. Once the skin depth becomes large compared with the maximum offset, attenuation below the seafloor is undetectable, and the fields are instead dominated by inductive dissipation in the seawater layer. Since the seawater resistivity does not vary from model to model, the misfit curve flattens, hence showing no sensitivity to resistivity changes below the seafloor.

Of all forms CSEM data can take, use of complex components appears to be the worst choice; their misfit curves drop slightly and then flatten on both ends of the parameter line as the model moves away from the true seafloor resistivity. The use of amplitude data (log- or linearly-scaled) on its own is safe on the low-end of the resistivity line but has a large local maximum separating the flat misfit on the high end of the resistivity line from the true minimum misfit model; this would surely confound any gradient-based inversion if the starting model is not sufficiently close to, or below, the true resistivity. The use of amplitude-only data at a single frequency was common practice in the early days of marine CSEM when phase was a very uncertain measurement (e.g. Johansen *et al.* 2005). With modern data it is always prudent to use phase, especially given that it provides the tightest constraints in misfit near the true

model (Fig. 9). Phase-only as well as joint amplitude-phase misfit curves do not have that dramatic local maximum on the high side of the resistively line, but they dip slightly and ultimately flatten on the low side of the resistivity line. The most favourable misfit behaviour is found with the joint log-amplitude and phase data ensemble, just as was the case for MT data. With log-amplitude and phase data, the flattening on the high end of the resistivity line is no worse than any of the other data forms, while the misfit at resistivities below the true model is always increasing away from the global minimum. The flat misfit curves for low-resistivities and non-log-scaled CSEM data can again be explained by the asymptotic reasoning of (3).

At present, most of the CSEM inversion codes that the authors are aware of invert the data as real and imaginary fields, especially those using the non-linear conjugate gradient method for large 3-D models. These codes would immediately benefit from reformulating their data misfit functional in terms of log-amplitude and phase. We describe the steps necessary to do that in Appendix A.

4 A STATISTICAL VIEW OF TRANSFORMATIONS OF COMPLEX DATA

One cannot properly solve an inverse problem without understanding the statistics of their fundamental data; the estimated uncertainties of a data set are as critical to an inverse problem as the data themselves. Likewise, upon transforming data into a new form, it is necessary to further analyse the effects that transformation has had on the statistics of the new data. In the following, we address this issue of the statistical repercussions of the specific data transformations discussed throughout this paper.

4.1 Error propagation

The use of each transformed data type within an inversion algorithm requires the propagation of the estimated errors from the original real and imaginary data components to the new transformed space (in this context, 'original' simply implies a form where the data probability distribution is known). Such propagated errors are critical for the weighting of residuals in a LS formulation, or, similarly, in defining the likelihood function for Bayesian methods. In this section, we provide the error propagation formulae for each data form. These formulae are approximations, therefore we investigate the rates at which they break down.

We also need to predict if any systematic bias is introduced to the data themselves by the various transformations out of the complex form. A LS solution attempts to produce a fit to the data whose residuals have a mean of zero. When the mean of our data on the whole is biased, the LS solution is misdirected away from the true solution. Furthermore, underdetermined LS inversions are only driven to the expected value for squared standardized residuals, which is usually higher than the minimum achievable squared misfit; the excess in degrees of freedom is such that the minimum misfit model involves spurious structure that is only needed to fit the errors in the data (Parker 1994). In Section 2.1, we mentioned using an rms misfit value of 1.00 as the target for inversion. In this section, we will drop the square-root and take the target to be the expected value for mean-squared standardized residuals. We will show that when the errors in the data become large, a target value of 1.00 no longer corresponds to the true model. Again, once we transform the data, we can no longer blindly aim for the same target misfit as before, else the inverse solution may not align with the true model.

We start by defining the basic random measurement as some complex number, either an impedance tensor element, a tipper vector element or simply an electromagnetic field component, each estimated at a given frequency. Our task is to calculate the mean and variance of functions of these random variables. The real and imaginary parts of these measurements are each contaminated by independent, identically distributed, Gaussian noise of zero mean and known variance, σ^2. For our complex datum, $z = x + iy$, we write:

$$Z = (x + X) + i(y + Y), \tag{12}$$

where a capital letter signifies a randomly perturbed quantity, and a lower case letter signifies the underlying true value. The probability density function of the bivariate noise is

$$\phi(X, Y) = \frac{\exp\{-(X^2 + Y^2)/2\sigma^2\}}{2\pi\sigma^2} \tag{13}$$

over the plane $\mathbb{R}^2$. We use the symbol δ to signify the standard error of a quantity, such that

$$\delta X = \sqrt{\mathrm{var(X)}} = \sigma. \tag{14}$$

Since X and Y are zero-mean, using the expectation operator, E, we have

$$\mathrm{E}[Z] = z; \tag{15}$$

given X and Y are independent, we quickly find

$$\delta Z^2 = \delta X^2 + \delta Y^2 = 2\sigma^2. \tag{16}$$

Now, we progress to more complicated functions of Z. For MT data, we need to calculate the mean and variance of the apparent resistivity, R, given by

$$R = \frac{1}{\mu\omega}\left[(x + X)^2 + (y + Y)^2\right], \tag{17}$$

the logarithm of apparent resistivity, $L = \log_{10}(R)$, and the impedance phase, P, given in units of radians by,

$$P = \arctan\left[(y + Y)/(x + X)\right]. \tag{18}$$

For CSEM data, the definition for phase remains the same as (18), and amplitude is given by

$$A = \sqrt{(x + X)^2 + (y + Y)^2}, \tag{19}$$

with the true amplitude $a \equiv |z|$. The logarithm of amplitude behaves identically to L in terms of relative variance, bias and mean-squared standardized misfit.

The standard method of calculating the variance of functions of a random variable is called linear propagation of errors, which utilizes the first-order terms in a Taylor series expansion of our functions of X and Y. Applied to R, L and P, we have the usual formulae for propagated errors in MT data:

$$\delta R^2 = \frac{4r}{\mu\omega}\sigma^2, \tag{20}$$

which is commonly used in the fractional standard error form

$$\frac{\delta R}{r} = \sqrt{\frac{4\sigma^2}{r\mu\omega}} = \frac{2\sigma}{|z|} = \sqrt{2}\frac{\delta Z}{|z|}, \tag{21}$$

and the remaining standard error forms are

$$\delta L = \frac{1}{\ln(10)}\left(\frac{2\sigma}{|z|}\right) = \frac{1}{\ln(10)}\left(\frac{\delta R}{r}\right), \tag{22}$$

and

$$\delta P = \frac{\sigma}{|z|} = \frac{1}{2}\left(\frac{\delta R}{r}\right). \tag{23}$$

For CSEM data or any complex measurement (displacement in the Fourier-domain, for example), the relative error in amplitude is

$$\frac{\delta A}{a} = \frac{\sigma}{|z|}, \tag{24}$$

and log-amplitude error, akin to (22), is given by,

$$\delta L = \frac{1}{\ln(10)}\left(\frac{\delta A}{a}\right). \tag{25}$$

One can also show, that to first-order $E[R] = r$, $E[P] = p$, $E[L] = l$ and cov(R, P) = cov(A, P) = cov(L, P) = 0; that is to say, to our first-order knowledge, each transformed pair of data are unbiased and uncorrelated.

The formulae (22) relative to (21), and (24) relative to (25), illustrate that data which are not necessarily equal, but whose errors scale with the magnitude of their corresponding datum, will have equivalent errors in the log-domain. This provides an easy check for accuracy in linear propagation of relative errors of a data set: if all data have the same percentage error, then plotted on a logarithmic scale the error-bars should look the same regardless of data amplitude. Furthermore, if subsequent LS inversion acts in the log-domain, the algorithm will assume the log-domain error-bars are symmetric, and therefore they should be plotted as so.

Notice that δL and δP are equal, ignoring the constants $1/\ln(10)$ and $1/2$, which cancel out in the inversion (Appendix A). It is often assumed that all the complex data have the same percentage amplitude of noise, in which case the error estimates for all the log-amplitude and phase data are the same. When the reciprocal of these error estimates are used as weights in inversion, all log-amplitude and phase data will naturally have equal importance regardless of their absolute magnitude; with CSEM data for example, the short- and long-offset data will automatically have the same weight. This observation is non-trivial because the same magnitude-independent weighting carries over to the forward calculation and functional derivatives as well. Thus, the transformation to log-amplitude and phase data has the same desirable effect of improving the condition of the linear system as do other pre-conditioners (e.g. Newman & Boggs 2004); this is exemplified by the factor of two increase in rate of convergence when switching to log-amplitude and phase inversion (e.g. Section 2.2). Yet, in contrast to other pre-conditioning strategies, log-amplitude and phase inversion requires no additional computational effort.

Now, we assess the inaccuracy of the first-order assumptions. To do this exactly, we would need to calculate complicated expectation integrals, take $E[P^2]$, for example. However, we are not interested in exact methods for bias estimation or error propagation, but instead an order of magnitude gauge on the degree of failure of our linear approximation. For this lesser task, we add the second-order terms to our Taylor series expansion of the arbitrary function, $F(x+X, y+Y)$, and are left with much simpler expectation integrals. Applying this methodology [for more detail see Wheelock (2013)], the second-order mean of F is found to be

$$E[F]_s = f + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\delta X^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}\delta Y^2, \tag{26}$$

and the second-order expansion for the variance of F is

$$\delta F_s^2 = \left(\frac{\partial f}{\partial x}\right)^2\sigma^2 + \left(\frac{\partial f}{\partial y}\right)^2\sigma^2 + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)^2\sigma^4 + \frac{1}{2}\left(\frac{\partial^2 f}{\partial y^2}\right)^2\sigma^4 + \left(\frac{\partial^2 f}{\partial x\,\partial y}\right)^2\sigma^4, \tag{27}$$

where the subscript s denotes 'second order'. The first term in (26) and the first two terms in (27) comprise the respective first-order versions. By applying (26), we see there is a small upward bias in the mean estimate of apparent resistivity, given by,

$$E[R]_s - r = \frac{2}{\mu\omega}\sigma^2. \tag{28}$$

Note that the second-order Taylor series expansion of R is exact since all higher order derivatives are zero. Normalizing (28) by the *first*-order δR, (21), to provide a measure of bias relative to our linearly propagated errors, we find

$$\frac{E[R]_s - r}{\delta R} = \frac{\sigma}{|z|}. \tag{29}$$

Thus, with 10 per cent error in R, the mean of R is a biased upward by 5 per cent of one linearly propagated standard deviation. This bias, though small for low-noise measurements, provides slightly more evidence against the use of the apparent resistivity data form. Similarly, the relative bias in amplitude, say for CSEM data, is

$$\frac{E[A]_s - a}{\delta A} = \frac{\sigma}{2a}, \tag{30}$$

where again a 10 per cent error in A, results in a 5 per cent bias in $E[A]_s/\delta A$. On the contrary, we find that the means of L and P can be considered bias-free by measure of the second-order expansion ($E[L]_s = l$, and $E[P]_s = p$). Considering this finding and especially those of Sections 2.1, 2.2 and 3, the log-amplitude (or log-apparent resistivity) and phase data forms are optimal for robust inversion, at least until the relative errors in the data exceed values amenable to prediction by a second-order approximation.

Using (27) to find the second-order relative forms of standard error in apparent resistivity, we have,

$$\frac{\delta R_s}{r} = \frac{2\sigma}{|z|}\sqrt{1 + \frac{\sigma^2}{|z|^2}}. \tag{31}$$

The first term outside of the square-root in (31) is the linear approximation, and the terms within the square-root represent the error in that approximation. The linearly approximated errors are underestimated by a very small amount. With 5 per cent noise added to the real and imaginary components of z ($\sigma = 0.05|z|$), the linear propagation of errors would predict 10 per cent noise in the apparent resistivity, whereas the true proportion of noise in the apparent resistivity is closer to 10.012 per cent. Similarly, the standard errors in the log of apparent resistivity and phase, under the second-order approximation, are

$$\delta L_s = \frac{2\sigma}{\ln(10)|z|}\sqrt{1 + \frac{\sigma^2}{|z|^2}}, \tag{32}$$

and,

$$\delta P_s = \frac{\sigma}{|z|}\sqrt{1 + \frac{\sigma^2}{|z|^2}}. \tag{33}$$

For the second-order variance in amplitude, we find empirically that (27) is not accurate enough given the square-root in A. Augmenting (26) up to fourth-order terms, we find a more accurate mean

amplitude

$$\mathrm{E}[A]_f = a + \frac{\sigma^2}{2a} + \frac{\sigma^4}{8a^3}, \tag{34}$$

and it is trivial to derive following exactly without appealing to any Taylor-series expansions

$$\mathrm{E}[A^2] = a^2 + 2\sigma^2. \tag{35}$$

Using (34) and (35) the second-order formula for the relative standard error in amplitude

$$\frac{\delta A_s}{a} = \frac{\sigma}{|z|}\sqrt{1 - \frac{\sigma^2}{2|z|^2}}. \tag{36}$$

In the case of amplitude, the linearly propagated errors are slightly overestimated; a 5 per cent linearly estimated error in reality is closer to 4.997 per cent. Computer generated random numbers confirmed the accuracy of our Taylor-series based formulae: (32) and (33) match simulations up to the percentage error, $\sigma/|z| \times 100 = 15$ per cent, beyond which they noticeably underpredict the failure of a linear-approximation; the amplitude formulae (30) and (36) perform well up to $\sigma/|z| \times 100 = 30$ per cent; and, again, formulae (29) and (31) are exact for all levels of percentage error. Finally, one can show that even by the second-order expansion, each pair of transformed data remain uncorrelated, $\mathrm{cov(R, P)} = \mathrm{cov(A, P)} = \mathrm{cov(L, P)} = 0$.

The second-order estimates of bias in the mean and deviations from our linearly propagated errors are minute and seemingly negligible when the percentage of noise in the complex data is small. Yet a mean-squared misfit measure is sensitive to both bias in the mean and bias in the linearly propagated standard errors. Again using Taylor-series expansions as above, we find that the expected values for the mean-squared misfit measures for the transformed data are perturbed from 1, the value expected for the fundamental complex data. The predictive formulae for the *sum*-squared misfit of N data points in the various forms are

$$\mathrm{E}\left[\sum_{i=1}^{N}\left(\frac{R_i - r_i}{\delta R_i}\right)^2\right]_s = \sum_{i=1}^{N}\left(1 + 2\frac{\sigma_i^2}{|z_i|^2}\right), \tag{37}$$

$$\mathrm{E}\left[\sum_{i=1}^{N}\left(\frac{A_i - a_i}{\delta A_i}\right)^2\right]_s = \sum_{i=1}^{N}\left(1 - \frac{\sigma_i^2}{4|z_i|^2}\right), \tag{38}$$

$$\mathrm{E}\left[\sum_{i=1}^{N}\left(\frac{L_i - l_i}{\delta L_i}\right)^2\right]_s = \sum_{i=1}^{N}\left(1 + \frac{\sigma_i^2}{|z_i|^2}\right), \tag{39}$$

and

$$\mathrm{E}\left[\sum_{i=1}^{N}\left(\frac{P_i - p_i}{\delta P_i}\right)^2\right]_s = \sum_{i=1}^{N}\left(1 + \frac{\sigma_i^2}{|z_i|^2}\right). \tag{40}$$

The effect on apparent resistivity data is the worst of all because both bias in the mean and the breakdown of linear error-propagation distort constructively. With 10 per cent error in all the complex components, the mean-squared misfit for the apparent resistivities of the true model is 1.02. The same noise propagated into log-amplitude and phase would produce a mean-squared misfit at the true model of 1.01; for amplitude, it is 0.9975. Knowing formulae (37)–(38), one can simply adjust their target misfit for inversion according to the percentage of noise they estimate to exist in their data. Yet, even when using the preferred log-amplitude and phase data, a safer strategy to avoid problems with bias in the mean and standard errors is to simply remove from inversion any transformed datum involving a fundamental data-error percentage, $\sigma/|z| \times 100$, greater than 10 per cent. Of course, inversion is not required to treat all data in the log-domain. So, with a flexible algorithm, one could retain only the high-error data, properly down-weighted, in the fundamental complex form, while treating the low-error data in the log-amplitude and phase form.

4.2 Gaussian conformance

A LS inversion method, like that used in the inversions of Section 2.2, has the following advantages when the random errors in the inverted data are uncorrelated, zero-mean and Gaussian-distributed of equal variance: it is a maximum likelihood, unbiased and minimum variance estimator; the latter two conditions are shown by the Gauss-Markov theorem (e.g. Priestly 1981). We have assumed that an LS inversion delivers these advantages when used with the fundamental complex data described in this paper. To confer the same advantages to log-scaled data, in Section 4.1, we showed that the Gauss-Markov conditions (uncorrelated, zero-mean and known variance) can be satisfied by errors in log-amplitude and phase data, at least, when the relative variance remains small. What is left is to assess the extent to which such transformed data appear Gaussian.

In the preceding sections, illustrations have dealt with the complex data of MT impedance or CSEM fields. It is likely that the errors in these data are uncorrelated and Gaussian-distributed as they are derived from a discrete Fourier transform (FT) of random fields regularly sampled in time. The Central Limit Theorem informs us that the FT, involving the weighted sum of many of these random samples, tends towards coefficients which are Gaussian-distributed as the number of samples involved increases, regardless of the original distribution of each individual sample in time. Furthermore, multiple FT coefficients from adjacent time-windows are often stacked together to increase the signal-tonoise ratio, a step which, again by the Central Limit Theorem, bolsters the assumption of Gaussian-distributed errors in our complex data. When we convert these stacked complex data to amplitude, squared-amplitude (as with apparent resistivity), log-amplitude, and phase, each a non-linear transformation, we would prefer the resulting probability distribution of data errors to remain Gaussian so that the desired performance of an LS inversion is preserved.

Stochastic inversion methods which make use of Markov chain Monte Carlo (MCMC) sampling (e.g. Ray *et al.* 2013) are perhaps more sensitive to the full probability distribution of the transformed data, rather than just their mean and standard deviation. For some of the transformed data we have discussed, the distributions are well known; amplitude data follow a Rice distribution, while apparent resistivity data follow a non-central chi-squared distribution. In short, the distributions of our converted data errors are not Gaussian, and an exploration into their precise forms is beyond the scope of this paper; for a more thorough treatment of this topic, see Chave & Lezaeta (2007). Instead, we are only interested in circumstances in which a realistic data set can appear to have Gaussian-distributed errors, so that we may continue to employ LS for deterministic methods or Gaussian likelihoods for MCMC machinery.

To quantify the Gaussian semblance of our data errors, we apply the Kolmogorov-Smirnov (KS) test (Massey 1951), using as the null hypothesis that the synthetic transformed data used in Section 2 are

Gaussian-distributed. The KS test produces a statistic, d_N, which is a measure of the difference between the hypothesized cumulative distribution function and the observed cumulative step-function of an N-sized sample. The key is that the probability distribution of d_N is known, and can be derived independently of the sample's cumulative distribution. Then, armed with the value of d_N calculated for a certain sample, the null hypothesis may be rejected with a quantifiable level of significance.

The synthetic data we will use are based on those of the solid lines in Fig. 2 but repeated for a survey of 50 MT receivers. Each receiver provides complex data at 17 frequencies in both the TE and TM modes resulting in the data quantity of $M = 1700$ per complementary data type. For example, we use 1700 real and 1700 imaginary data. Independent realizations of Gaussian-distributed random numbers are added to the real and imaginary data, with standard deviations scaled to a fraction of the amplitude of each complex pair, $\sigma/|z|$. Then, the noisy and clean data are both converted into apparent resistivity, amplitude, log-apparent resistivity (analogous to log-amplitude), and phase, and residuals for each are calculated. Note that it is crucial that the phase data are unwrapped before residuals are computed (Appendix A). Each of the residuals are then normalized by linearly propagated error estimates (Section 4.1), as they would be in an inversion scheme. The KS statistic, d_N, is then calculated for each type of normalized residual; the results are plotted in Fig. 10.

First, notice the random fluctuations of the d_N; we must not forget that the statistic itself is a random number. However, there is a clear underlying trend to the d_N. As we would expect, the trends for the d_N that pertain to the real and imaginary parts of impedance are flat, indicating that the null hypothesis, 'the residuals are Gaussian-distributed', is acceptable for all levels of noise. Similarly, the amplitude residuals appear to be Gaussian-distributed over the tested range of noise levels. The phase errors appear to be Gaussian-distributed until the percentage of noise exceeds 30 per cent.

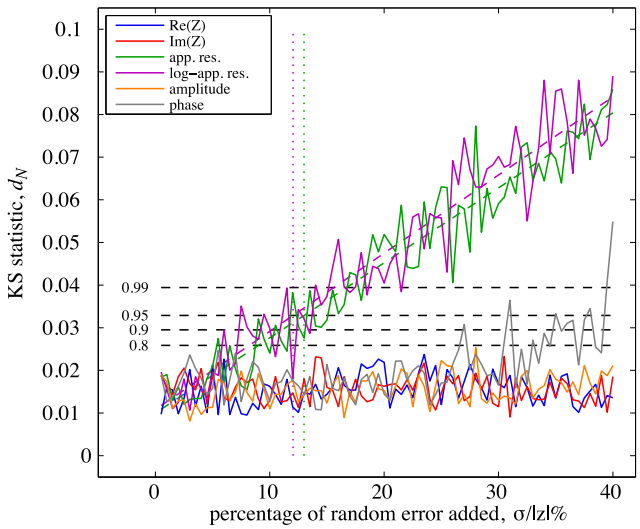

Figure 10. KS statistic, d_N, for various MT data types given variably proportioned Gaussian noise added to the real and imaginary parts of impedance. Confidence thresholds for the rejection of the hypothesis that the data are Gaussian-distributed are shown with the dashed horizontal black lines. Dashed green and magenta lines show the linear trend fit to the KS statistics for apparent resistivity and log-apparent resistivity, respectively. The dotted vertical lines mark the point on the percentage-noise line beyond which apparent resistivity and log-apparent resistivity residuals may be rejected as Gaussian-distributed with 95 per cent confidence.

However, the apparent resistivity and log-apparent resistivity residuals only seem Gaussian when the percentage errors are even smaller. We fit linear trends to the noisy d_N for apparent resistivity and log-apparent resistivity, and find those trends cross the 95 per cent confidence level for rejection of the null hypothesis at 12 per cent and 13 per cent added noise, respectively. Thus, when the estimated noise in the complex data exceeds about 12 per cent of its magnitude, it is unwise to assume errors in apparent resistivity or log-amplitude are Gaussian-distributed. So long as errors in the complex data remain small, stochastic inversions may be applied to log-amplitude and phase data using the usual Gaussian likelihood function. Though we have yet to see a stochastic inversion of log-amplitude and phase data, for small data errors we do not expect the log-amplitude and phase derived posterior distributions to differ from the those derived from real and imaginary data. However, the improvement in smoothness and removal of vanishing gradients in the misfit space, as demonstrated in Sections 2.1 and 3, are expected to reduce the 'burn-in' period, and perhaps even speed-up the convergence upon the distributions of solution models.

5 CONCLUSIONS

In this paper, we have analysed the effect of common data transformations for complex data, constructing examples from frequency-domain electromagnetic geophysical methods. We have used single-parameter misfit curves as well as multiparameter synthetic inversions to decide which data transformation is optimal for robust and fast inversion. Our conclusion is that the logarithmic-amplitude and phase transformation is the optimal form. It produces the smoothest misfit space which is rarely flat, two qualities that facilitate an inversion's reliable convergence to the true solution. In contrast, we show that all non-log-scaled data result in flat misfit spaces on the end of parameter space which produces observations of diminishing magnitude.

For the case of MT data, this analysis applies to impedance and tipper data, where both are best inverted in terms of log-amplitude and phase. We showed that coast-affected data reap the greatest benefit from a log-amplitude and phase transformation, due to their large dynamic range and non-linear nature with respect to frequency and seafloor resistivity. It was also shown that while the coast-effect can hamper inversion of non-log-scaled data, it nonetheless provides better constraints on the subsurface, regardless of the form the data take; the gradients in misfit are steeper near the true solution when a coast is involved. For CSEM, again the large dynamic range in the data produces flatness on the low-resistivity end of parameter space when data are not log-scaled. No form of scaling for CSEM data removes the flatness on the high-resistivity end of parameter space, but this is due to vanishing seafloor sensitivity. We also showed that with both MT and CSEM data, there are pitfalls when using, for example, only amplitude or only TM-mode data; the use of complementary data types, log-amplitude with phase or TE-mode with TM-mode, removes local minima from the misfit space.

Perhaps the most encouraging finding is that a log-amplitude and phase inversion converges almost two times faster than inversions of all other data types. Our high-dimensional synthetic inversions also showed that a log-amplitude and phase inversion is far more robust with respect to starting model and with respect to the degree of noise in the data.

Finally, we analysed the statistics of each transformed data type under the assumption that the fundamental real and imaginary data are Gaussian-distributed. We found that for small errors (less than

10 per cent of amplitude in the real and imaginary components) the assumptions necessary for using least-squares and a mean-squared standardized-misfit target of 1.00 in a deterministic, log-amplitude and phase inversion still hold. Similarly, a Gaussian likelihood may still be used within an MCMC stochastic inversion of log-amplitude and phase data, so long as the relative errors remain below about 12 per cent of the amplitude (a consequence of the probability density function of log-amplitude, and less so of phase). We provide formulae to predict changes in the means, standard errors and the mean-squared misfit for the transformed data, as the fundamental errors grow. Nevertheless, we recommend that data with fundamental errors of 10 per cent or greater be removed from log-amplitude and phase inversion; if data with large errors must be used, a more statistically sound choice is to invert them in the fundamental complex form.

ACKNOWLEDGEMENTS

This work was supported in part by the Seafloor Electromagnetic Methods Consortium at Scripps Institution of Oceanography. We thank Robert L. Parker for discussions which helped guide the statistical aspect of our research. We also thank Andrei Swidinsky, Max Moorkamp and one other anonymous referee for their thorough reading of the manuscript, and for their suggestions which certainly improved this text.

REFERENCES

Chave, A. & Jones, A.G., 2012. *The Magnetotelluric Method: Theory and Practice,* Cambridge Univ. Press.

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

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

Constable, S., Key, K. & Lewis, L., 2009. Mapping offshore sedimentary structure using electromagnetic methods and terrain effects in marine magnetotelluric data, *Geophys. J. Int.,* **176**(2), 431–442.

Cox, C., 1980. Electromagnetic induction in the oceans and inferences on the constitution of the earth, *Geophys. Sur.,* **4,** 1–20.

DeGroot Hedlin, C. & Constable, S., 1990. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data, *Geophysics,* **55**(12), 1613–1624.

Egbert, G., 1990. Comments on 'Concerning dispersion relations for the magnetotelluric impedance tensor' by E. Yee and KV Paulson, *Geophys. J. Int.,* **102,** 1–8.

Johansen, S.E., Amundsen, H.E.F., Røsten, T., Ellingsrud, S., Eidesmo, T. & Bhuiyan, A.H., 2005. Subsurface hydrocarbons detected by electromagnetic sounding, *First Break,* **23**(1081), 31–36.

Key, K., 2009. 1D inversion of multicomponent, multifrequency marine CSEM data: methodology and synthetic studies for resolving thin resistive layers, *Geophysics,* **74**(2), F9–F20.

Key, K. & Constable, S., 2010. Coast effect distortion of marine magnetotelluric data: insights from a pilot study offshore northeastern Japan, *Phys. Earth planet. Inter.,* **184**(3–4), 194–207.

Massey, F.J. Jr. 1951. The Kolmogorov-Smirnov test for goodness of fit, *J. Am. Stat. Assoc.,* **46**(253), 68–78.

Meqbel, N.M., Egbert, G.D., Wannamaker, P.E., Kelbert, A. & Schultz, A., 2014. Deep electrical resistivity structure of the northwestern U.S. derived from 3-D inversion of USArray magnetotelluric data, *Earth planet. Sci. Lett.,* **402,** 290–304.

Myer, D., Constable, S., Key, K., Glinsky, M.E. & Liu, G., 2012. Marine CSEM of the Scarborough gas field, Part 1: Experimental design and data uncertainty, *Geophysics,* **77**(4), E281–E299.

Newman, G.A. & Alumbaugh, D.L., 2000. Three-dimensional magnetotelluric inversion using non-linear conjugate gradients, *Geophys. J. Int.,* **140,** 410–424.

Newman, G.A. & Boggs, P.T., 2004. Solution accelerators for large-scale three-dimensional electromagnetic inverse problems, *Inverse Prob.,* **20**(6), S151–S170.

Parker, R.L., 1994. *Geophysical Inverse Theory,* Princeton Univ. Press.

Priestly, M.B., 1981. *Spectral Analysis and Time Series,* Academic Press.

Ray, A., Alumbaugh, D.L., Hoversten, G.M. & Key, K., 2013. Robust and accelerated Bayesian inversion of marine controlled-source electromagnetic data using parallel tempering, *Geophysics,* **78**(6), E271–E280.

Selway, K., Thiel, S. & Key, K., 2012. A simple 2-D explanation for negative phases in TE magnetotelluric data, *Geophys. J. Int.,* **188,** 945–958.

Shin, C. & Min, D.-J., 2006. Waveform inversion using a logarithmic wavefield, *Geophysics,* **71**(3), R31–R42.

Tietze, K. & Ritter, O., 2013. Three-dimensional magnetotelluric inversion in practice—the electrical conductivity structure of the San Andreas Fault in Central California, *Geophys. J. Int.,* **195,** 130–147.

Wannamaker, P.E., Stodt, J.A. & Rijo, L., 1987. A stable finite element solution for two-dimensional magnetotelluric modelling, *Geophys. J. R. astr. Soc.,* **88**(1), 277–296.

Weidelt, P., 1982. Response characteristics of coincident loop transient electromagnetic systems, *Geophysics,* **47**(9), 1325–1330.

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

Wheelock, B.D., 2013. Electromagnetic imaging of the crust and upper mantle across the continental margin in Central California, *PhD thesis,* University of California, San Diego.

Wilson, G.A., Raiche, A.P. & Sugeng, F., 2006. 2.5D inversion of airborne electromagnetic data, *Explor. Geophys.,* **37**(4), 363–371.

Worzewski, T., Jegen, M. & Swidinsky, A., 2012. Approximations for the 2-D coast effect on marine magnetotelluric data, *Geophys. J. R. astr. Soc.,* **189**(1), 357–368.

Zhdanov, M.S., Smith, R.B., Gribenko, A., Cuma, M. & Green, M., 2011. Three-dimensional inversion of large-scale EarthScope magnetotelluric data based on the integral equation method: geoelectrical imaging of the Yellowstone conductive mantle plume, *Geophys. Res. Lett.,* **38**(8), L08307, doi:10.1029/2011GL046953.

APPENDIX A: ADAPTATION OF GAUSS-NEWTON AND NLCG

For non-linear inversion, two commonly used algorithms are the Gauss-Newton and Non-linear Conjugate Gradient (NLCG) methods. In this appendix, we describe how to adapt these two algorithms to invert complex data in the log-amplitude and phase domain. Note that we will describe the adaptation of algorithms that were previously designed for frequency-domain, complex data, such as for MT impedance, or CSEM electromagnetic fields. However, the principles are easily transferred to non-negative, real data, consistent with certain time-domain electromagnetic methods (e.g. Weidelt 1982). So long as the data are always positive, the log-scaling we describe in the following text may be used with real data by simply ignoring the phase component.

The adaptation of the Gauss-Newton method is quite simple. One need only redefine the normalized misfit measure, or the residuals to be squared and summed, and re-scale the Jacobian matrix. In essence, the same two steps underlie the transformation of NLCG, but they are slightly obscured by the gradient calculation when using adjoint-reciprocity. We first turn to Gauss-Newton where the Jacobian matrix is calculated and stored in full. The scaling of the Jacobian follows from the chain rule:

$$\hat{J}_{ij} = \frac{\partial \hat{f}_i}{\partial m_j} = \frac{\partial \hat{f}_i}{\partial f_k}\frac{\partial f_k}{\partial m_j} = \frac{\partial \hat{f}_i}{\partial f_k} J_{kj}, \tag{A1}$$

where we have reused the 'hat' notation from equations (3)–(10). The functional $\partial \hat{f}_i / \partial f_k$ represents the data-space transformation whose behaviour for MT data is defined by the first-order derivatives of equation (17), after taking logarithm, and equation (18) with respect to x and y. The analogue formulae for log-amplitude data, say for CSEM electric fields, as opposed to logarithmic apparent resistivity, are similarly derived. In fact, with the proper normalization, it makes no difference whether MT data are treated as the log-amplitude and phase of a complex impedance element itself, or the more traditional logarithmic apparent resistivity and phase. Therefore in the remaining development we drop the apparent resistivity formulation and work in the more general terms of log-amplitude and phase, a characterization common to all complex data types.

The normalized residuals may be re-scaled as was done for log-amplitude type data with the terms inside the brackets of (7). But, for the sake of generality, here we will not assume equal errors in the real and imaginary parts, X and Y, respectively, of the complex data, Z. The two forms of normalized residuals are then,

$$\frac{\Delta L}{\delta L} = \frac{|Z|^2}{\sqrt{X^2 \delta X^2 + Y^2 \delta Y^2}} \ln\left(\frac{|Z|}{|z|}\right), \tag{A2}$$

for the log-amplitude data, L, and

$$\frac{\Delta P}{\delta P} = \frac{|Z|^2}{\sqrt{X^2 \delta Y^2 + Y^2 \delta X^2}} (P - p), \tag{A3}$$

for the phase data, P, which will have units of radians throughout this appendix. We have reused notation conventions similar to Section 4.1. While upper-case letters still denote random variables, that is our measured data, the same letter in lower case denotes the corresponding non-random value of the forward functional's data prediction. It is necessary to check for phase-wrapping when dealing with phase residuals. This may be done by checking alternative residuals $\Delta P + 2\pi$ and $\Delta P - 2\pi$, as well as the original ΔP, and selecting the smallest one among them. The Jacobian, however, is not affected by phase-wrapping.

Given that primer, we now turn to the NLCG method. We will follow the description and notation found in Newman & Boggs (2004), and refer to equations therein with the prefix N. In addition to re-scaling of the residuals used in the objective function, which we discussed in the preceding paragraph, the transformation of the NLCG inverse algorithm to the log-amplitude and phase data space involves re-scaling the data component of the objective-function's gradient, $\nabla \phi_d$. In essence, this data gradient is a collapsed Jacobian, where each row of the Jacobian is scaled by the corresponding data residual before being summed column-wise; the resulting vector has a dimension equal to the number of the unknowns in the model. Newman & Boggs (2004) describe the calculation of $\nabla \phi_d$ which uses adjoint reciprocity. This method has the advantage of never having to explicitly compute or store the full Jacobian matrix which can be immense for realistic 3-D problems. The data gradient is computed in part by solving the same linear system that represents the forward functional, (N19), but where the source vector, (N18), is modified to contain the normalized data residuals inserted at their respective point of observation in the model. The solution of this linear system produces complex field values covering the entire model domain. Thus, the data gradient is often described as a reverse-propagation of the data residuals onto the model space. In order to transform an NLCG algorithm from inverting for the complex data to inverting for the log-amplitude and phase of those data, we need only to modify the source vector once more, while all else in the linear system remains the same.

Substituting $\hat{J}$ for J in (N16) and inserting (A1), we find

$$\hat{J}_{ji} = G_{jp} K^{-1}_{pl} Q_{lk} \frac{\partial \hat{f}_k}{\partial f_i}. \tag{A4}$$

We combine the final two multiplicands in (A4), the interpolator matrix, Q, and the non-linear data-transformation functional, rewritten as $T[f] = \partial \hat{f}_i / \partial f_k$, with the source vector to get,

$$\hat{v} = Q^T T[f]^T \hat{D}^T \{\hat{D}(\hat{d} - \hat{f})\}^* = Q^T \hat{s}. \tag{A5}$$

Here we have also introduced a new diagonal weighting matrix, $\hat{D}$, which uses our transformed error estimates δL and δP. To simplify the recoding of an NLCG algorithm which was designed for complex data, we formulate a new complex data vector, $\hat{d}$, with log-amplitude in the real part and phase in the imaginary part like so,

$$\hat{d} = \ln(|Z|) + i \arg(Z) = L + iP. \tag{A6}$$

The corresponding model prediction is

$$\hat{f} = \ln(|z|) + i \arg(z) = l + ip. \tag{A7}$$

Recall that $f = z = x + iy$. We perform complex differentiation using the Wirtinger derivative to find the functional form of $T[f]$, yielding

$$\frac{\partial \hat{f}}{\partial f} = \frac{1}{2}\left(\frac{\partial l}{\partial x} + \frac{\partial p}{\partial y}\right) + \frac{i}{2}\left(\frac{\partial p}{\partial x} - \frac{\partial l}{\partial y}\right). \tag{A8}$$

The non-linear functional (A8) simplifies rather satisfyingly to,

$$\frac{\partial \hat{f}}{\partial f} = \frac{e^{-ip}}{|z|}. \tag{A9}$$

Finally, combining (A2), (A3), (A5) and (A9), we find that the pre-interpolation, transformed source vector, $\hat{s}$ from the right-hand side of (A5), is given by,

$$\hat{s} = \frac{e^{-ip}}{|z|}\left(\frac{\Delta L}{\delta L^2} - i\frac{\Delta P}{\delta P^2}\right). \tag{A10}$$