Geophys. J. Int. (2013) doi: 10.1093/gji/ggt366
Geophysical Journal International Advance Access published October 25, 2013

On the statistics of magnetotelluric rotational invariants

Alan D. Chave

Deep Submergence Laboratory, Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, *USA. E-mail: achave@whoi.edu*

Accepted 2013 September 10. Received 2013 September 9; in original form 2013 May 28

SUMMARY

The statistical properties of the Swift skew, the phase-sensitive skew and the WAL invariants I_1-I_7 and Q are examined through analytic derivation of their probability density functions and/or simulation based on a Gaussian model for the magnetotelluric response tensor. The WAL invariants I_1-I_2 are shown to be distributed as a folded Gaussian, and are statistically well behaved in the sense that all of their moments are defined. The probability density functions for Swift skew, phase-sensitive skew and the WAL invariants I_3-I_4, I_7 and Q are derived analytically or by simulation, and are shown to have no moments of order 2 or more. Since their support is semi-infinite or infinite, they cannot be represented trigonometrically, and hence are inconsistent with a Mohr circle interpretation. By contrast, the WAL invariants I_5-I_6 are supported on $[-1,\ 1]$, and are inferred to have a beta distribution based on analysis and simulation. Estimation of rotational invariants from data is described using two approaches: as the ratio of magnetotelluric responses that are themselves averages, and as averages of section-by-section estimates of the invariant. Confidence intervals on the former utilize either Fieller's theorem, which is preferred because it is capable of yielding semi-infinite or infinite confidence intervals, or the less accurate delta method. Because section-by-section averages of most of the rotational invariants are drawn from distributions with infinite variance, the classical central limit theorem does not pertain. Instead, their averaging is accomplished using the median in place of the mean for location and an order statistic model to bound the confidence interval of the median. An example using real data demonstrates that the ratio of averages approach has serious systematic bias issues that render the result physically inconsistent, while the average of ratios result is a smooth, physically interpretable function of period, and is the preferred approach.

Key words: Probability distributions; Electromagnetic theory; Magnetotellurics.

1 INTRODUCTION

Rotational invariants are magnetotelluric parameters that remain identical for any orientation of the response tensor $\overleftrightarrow{\mathbf{Z}}$, and have long been investigated as an aid towards elucidating the resistivity structure beneath and around measurement sites. Recent work has elaborated the concept of rotational invariants and clarified the intertwining nature of electromagnetic distortion with dimensionality determination. Bahr (1988) defined a principal superimposition model for a distorted 2-D structure, and introduced the phase-sensitive skew as a rotationally invariant dimensionality indicator. Bahr (1991), as corrected by Prácser & Szarka (1999), built on this model and provided threshold values for four rotational invariants, including the phase-sensitive skew, to estimate the resistivity structure dimensionality. Weaver *et al.* (2000), extending earlier work by Szarka & Menvielle (1997), defined a set of seven independent rotational invariants $I_1 - I_7$ and a dependent rotational invariant Q (hereafter the WAL invariants) that, when combined with a single rotation parameter, comprise a complete representation of the eight elements in the complex response tensor $\overleftrightarrow{\mathbf{Z}}$. Martí *et al.* (2005) reconciled the Bahr and WAL approaches, and established threshold values for the invariants based on simulation. Jones (2012) provides a recent comprehensive review of distortion and rotational invariants.

Little corresponding attention has been devoted to the statistics of rotational invariant estimators, most of which are ratios of combinations of response tensor elements, as pertains to the Swift skew, Bahr's phase-sensitive skew and the WAL invariants $I_3 - I_7$ and Q. However, understanding their statistics is essential for quantitative interpretation, as an estimate for a statistic without a sense of its accuracy is nearly useless. For example, it is not sufficient to simply establish threshold values for the rotational invariants based on simulation to estimate the dimensionality of Earth; rather, it is essential to first establish their significance relative to zero before comparison to thresholds.

To illustrate the issues posed by entities that are the ratios of functions of elements of $\overleftrightarrow{\mathbf{Z}}$, as an exemplar, the statistical distribution of the Swift skew is derived in Section 2 from first principles based on a Gaussian model for the elements of $\overleftrightarrow{\mathbf{Z}}$. It is shown that all moments of the skew distribution of order 2 and higher do not exist, so that in particular the Swift skew variance is undefined, and the mean can be upwards or downwards biased without bound. Section 3 describes the distributions for phase-sensitive skew and the WAL invariants, deriving closed forms for $I_1 - I_4$ and establishing that phase-sensitive skew, $I_3 - I_4$, I_7 and Q have no variance and higher order moments. As a consequence, the classical central limit theorem does not pertain. In contrast, the WAL invariants $I_5 - I_6$ are statistically well behaved, and consistent with a beta distribution. Section 4 describes statistical inference for rotational invariants in two ways: using Fieller's method or the delta approximation when the invariant is the ratio of combinations of the magnetotelluric response tensor elements that are themselves averages, and using order statistics when estimates of the invariants are averaged. Section 5 illustrates some of the results using real data. Section 6 contains conclusions.

2 THE SWIFT SKEW DISTRIBUTION

All of the derivations in this paper were completed using Mathematica 9, and have been verified numerically and by simulation.

The earliest rotational invariant is the Swift skew (Swift 1967) given by the ratio of the magnitudes of the trace and antitrace of $\overleftrightarrow{\mathbf{Z}}$:

$$S = \frac{\left|Z_{xx} + Z_{yy}\right|}{\left|Z_{xy} - Z_{yx}\right|}. \tag{1}$$

The Swift skew is zero for a 1- or 2-D structure and non-zero for a 3-D structure, hence serves as a dimensionality indicator.

A statistical model for the elements of the magnetotelluric response tensor constitutes the basis for derivation of the Swift skew distribution. Let a row of $\overleftrightarrow{\mathbf{Z}}$ be estimated from a solution to the least-squares problem

$$\mathbf{e} = \overleftrightarrow{\mathbf{b}} \cdot \mathbf{z} + \boldsymbol{\varepsilon}, \tag{2}$$

where there are N observations (i.e. N Fourier transforms of N independent data sections at a given frequency), so that $\mathbf{e}$ is an electric field N-vector, $\overleftrightarrow{\mathbf{b}}$ is the $N \times 2$ magnetic field matrix, $\mathbf{z}$ is a 2-vector, $\boldsymbol{\varepsilon}$ is an N-vector of random errors and $\cdot$ denotes the inner product. It is well known (e.g. Chave & Thomson 2004; Chave 2012) that when the elements of the estimated residuals $\mathbf{r}$ for the random errors $\boldsymbol{\varepsilon}$ are complex N-variate Gaussian with zero mean and common variance σ^2, the elements of the solution $\hat{\mathbf{z}}$ to (2) are bivariate complex Gaussian with mean $\mathbf{z}$ and variance $\sigma^2(\mathbf{b}^{\mathrm{H}} \cdot \mathbf{b})^{-1}$, where the superscript H denotes the Hermitian conjugate. This condition typically holds at least approximately when (2) is solved using a bounded influence estimator. The additional simplifying assumption that the components of the magnetic field are uncorrelated will be made, so that $(\mathbf{b}^{\mathrm{H}} \cdot \mathbf{b})^{-1}$ is diagonal. This approximation will break down for some source types where the magnetic field is elliptically polarized.

Under these conditions, the bivariate Gaussian factors into the product of two univariate Gaussian distributions given by

$$f_{jk}\left(z^{\mathrm{r}}_{jk}, z^{\mathrm{i}}_{jk} | \mu^{\mathrm{r}}_{jk}, \mu^{\mathrm{i}}_{jk}, \sigma_{jk}\right) = \frac{1}{2\pi\sigma^2_{jk}} \mathrm{e}^{-\left(z^{\mathrm{r}}_{jk} - \mu^{\mathrm{r}}_{jk}\right)^2 / \left(2\sigma^2_{jk}\right)} \mathrm{e}^{-\left(z^{\mathrm{i}}_{jk} - \mu^{\mathrm{i}}_{jk}\right)^2 / \left(2\sigma^2_{jk}\right)}, \tag{3}$$

where z^{r}_{jk} and z^{i}_{jk} are the real and imaginary parts of the magnetotelluric response, the indices j and k may be either x or y, μ^{r}_{jk} and μ^{i}_{jk} are the real and imaginary parts of the population values of the magnetotelluric response and $\sigma^2_{jk} = \mathcal{E}[(z^{\mathrm{r}}_{jk} - \mu^{\mathrm{r}}_{jk})^2] = \mathcal{E}[(z^{\mathrm{i}}_{jk} - \mu^{\mathrm{i}}_{jk})^2]$ is the corresponding population variance.

As is shown in elementary statistics texts (e.g. De Groot & Schervish 2002, section 3.9), the distribution for the sum of two independent random variables drawn from any pair of probability density functions f_{i} is given by

$$g(y) = \int_{-\infty}^{\infty} f_1(y - z) f_2(z) \mathrm{d}z. \tag{4}$$

Eq. (4) is just the convolution of the two probability density functions, and hence the distribution for the difference of two independent random variables is their correlation.

Consequently, the distribution for the trace of $\overleftrightarrow{\mathbf{Z}}$ is

$$g_n\left(y_{\mathrm{r}}, y_{\mathrm{i}} | \mu^{\mathrm{r}}_{jk}, \mu^{\mathrm{i}}_{jk}, \sigma_{jk}\right) = \frac{1}{2\,\pi\,\left(\sigma^2_{xx} + \sigma^2_{yy}\right)} \exp\left[\frac{-\left(y_{\mathrm{r}} - \mu^{\mathrm{r}}_{xx} - \mu^{\mathrm{r}}_{yy}\right)^2}{2\left(\sigma^2_{xx} + \sigma^2_{yy}\right)}\right] \exp\left[\frac{-\left(y_{\mathrm{i}} - \mu^{\mathrm{i}}_{xx} - \mu^{\mathrm{i}}_{yy}\right)^2}{2\left(\sigma^2_{xx} + \sigma^2_{yy}\right)}\right]. \tag{5}$$

The distribution for the numerator of the skew is obtained by changing variables from the real and imaginary parts to the magnitude and phase, and then integrating over the support of the latter to obtain the marginal distribution of the magnitude. The joint distribution of $\eta = \sqrt{y^2_{\mathrm{r}} + y^2_{\mathrm{i}}}$ and $\phi = \tan^{-1}(y_{\mathrm{i}}/y_{\mathrm{r}})$ is

$$h\left(\eta,\ \phi | \mu_1, \sigma^2_1\right) = \frac{\eta}{2\,\pi\,\sigma^2_1} \exp\left[\frac{-\left(|\mu_1|^2 + \eta^2\right)}{2\sigma^2_1}\right] \exp\left[\frac{|\mu_1|\,\eta\cos\left(\phi - \xi\right)}{\sigma^2_1}\right], \tag{6}$$

where $\mu_1 = \mu_{xx} + \mu_{yy}$, $\sigma_1^2 = \sigma_{xx}^2 + \sigma_{yy}^2$ and $\xi = \tan^{-1}(\frac{\mu_{xx}^{\mathrm{i}} + \mu_{yy}^{\mathrm{i}}}{\mu_{xx}^{\mathrm{r}} + \mu_{yy}^{\mathrm{r}}})$. Substituting the series expansion of the second exponential term in (6):

$$\mathrm{e}^{\alpha \cos\theta} = I_0(\alpha) + 2\sum_{k=1}^{\infty} I_k(\alpha)\cos(k\theta), \tag{7}$$

where $I_{\mathrm{k}}(x)$ is a modified Bessel function of the first kind of order k, and performing the integration in ϕ over $[-\pi, \pi)$ gives

$$h_n\left(\eta|\mu_1, \sigma_1^2\right) = \frac{\eta}{\sigma_1^2}\exp\left[\frac{-\left(\eta^2 + |\mu_1|^2\right)}{2\sigma_1^2}\right] I_0\left(\frac{|\mu_1|\,\eta}{\sigma_1^2}\right). \tag{8}$$

Eq. (8) is the Rice distribution (Rice 1945) with parameters (μ_1, σ_1^2), and reduces to the Rayleigh distribution when $\mu_1 = 0$. The distribution $h_d(\eta|\mu_2, \sigma_2^2)$ for the denominator can be derived in an analogous manner, and is also the Rice distribution with parameters $\mu_2 = \mu_{xy} - \mu_{yx}$ and $\sigma_2^2 = \sigma_{xy}^2 + \sigma_{yx}^2$.

The probability density function of the ratio of two independent random variables is given by (e.g. De Groot & Schervish 2002, section 3.9)

$$g(y) = \int_{-\infty}^{\infty} |z| f_n(zy) f_d(z)\mathrm{d}z, \tag{9}$$

where f_{n} and f_{d} are the distributions of the numerator and denominator. Equivalently, the ratio of two random variables can be viewed as the product of the first and the inverse distribution for the second, in which case their probability density becomes

$$g\left(y\right) = \int_{-\infty}^{\infty} f_n\left(z\right) \bar{f}_d\left(y/z\right) \frac{\mathrm{d}z}{|z|}, \tag{10}$$

and $\bar{f}_{\mathrm{d}}$ is called an inverted distribution. The properties of inverted distributions and their implications for rotational invariants will be addressed in the next section.

Substituting the two Rice distributions into (9) yields the skew distribution

$$s\left(\zeta|\mu_1, \mu_2, \sigma_1^2, \sigma_2^2\right) = \frac{\mathrm{e}^{-|\mu_1|^2/\left(2\sigma_1^2\right)}\mathrm{e}^{-|\mu_2|^2/\left(2\sigma_2^2\right)}}{\sigma_1^2\sigma_2^2}\zeta\int_0^{\infty} z^3\mathrm{e}^{-\frac{z^2\zeta^2}{2\sigma_1^2}}\mathrm{e}^{-\frac{z^2}{2\sigma_2^2}} I_0\left(\frac{|\mu_1|\,z\,\zeta}{\sigma_1^2}\right) I_0\left(\frac{|\mu_1|\,z}{\sigma_2^2}\right)\mathrm{d}z. \tag{11}$$

Define the inverse coefficients of variation $\kappa_1 = |\mu_1|/\sigma_1$ and $\kappa_2 = |\mu_2|/\sigma_2$. Eq. (11) can be simplified by non-dimensionalizing the variables as $\zeta' = |\mu_2|\zeta/|\mu_1|$ and $z' = z/|\mu_2|$, where the former is just the Swift skew divided by its population value $|\mu_1/\mu_2|$. After dropping the prime notation on z, the transformed version of (11) becomes

$$s'(\zeta'|\kappa_1, \kappa_2) = \kappa_1^2\kappa_2^2\mathrm{e}^{-\kappa_1^2/2}\mathrm{e}^{-\kappa_2^2/2}\zeta'\int_0^{\infty} z^3\mathrm{e}^{-\kappa_1^2 z^2\zeta'^2/2}\mathrm{e}^{-\kappa_2^2 z^2/2} I_0(\kappa_1^2 z\zeta') I_0(\kappa_2^2 z)\mathrm{d}z. \tag{12}$$

Neither (11) nor (12) can be expressed in closed form, but both are easily integrated numerically.

The n-th non-central moment of (12) is

$$\mathcal{M}[\zeta'] = \kappa_1^2\kappa_2^2\mathrm{e}^{-\kappa_1^2/2}\mathrm{e}^{-\kappa_2^2/2}\int_0^{\infty} z^3\left[\int_0^{\infty}(\zeta')^{n+1}\,\mathrm{e}^{-\kappa_1^2 z^2(\zeta')^2/2}\, I_0(\kappa_1^2 z\zeta')\mathrm{d}\zeta'\right]\mathrm{e}^{-\kappa_2^2 z^2/2} I_0(\kappa_2^2 z)\mathrm{d}z. \tag{13}$$

The term in square brackets in (13) may be integrated directly:

$$\int_0^{\infty}\zeta^{n+1}\mathrm{e}^{-\kappa_1^2 z^2\zeta^2/2} I_0(\kappa_1^2 z\zeta)\mathrm{d}\zeta = \left(\sqrt{2}\right)^n\Gamma\left(\frac{n+2}{2}\right) L_{-(n+2)/2}(\kappa_1^2/2)(\kappa_1^2 z^2)^{-(n+2)/2}, \tag{14}$$

where $\Gamma(x)$ is the gamma function and $L_{\mathrm{v}}(x)$ is a Laguerre polynomial. When $n = 1$, (14) reduces to

$$\int_0^{\infty}\zeta^2\mathrm{e}^{-\kappa_1^2 z^2\zeta^2/2} I_0(\kappa_1^2 z\zeta)\mathrm{d}\zeta = \sqrt{\frac{\pi}{2}}\frac{\mathrm{e}^{\kappa_1^2/4}}{2\kappa_1^3}\left[\kappa_1^2 I_1\left(\kappa_1^2/4\right) + \left(\kappa_1^2 + 2\right) I_0\left(\kappa_1^2/4\right)\right] z^{-3}. \tag{15}$$

Substituting (15) into (13) and performing the integration over z gives the non-dimensional expected value

$$\mathcal{E}\left[\zeta'\right] = \frac{\pi}{4}\frac{\kappa_2}{\kappa_1}\mathrm{e}^{-\kappa_1^2/4}\mathrm{e}^{-\kappa_2^2/4} I_0\left(\kappa_2^2/4\right)\left[\left(\kappa_1^2 + 2\right) I_0\left(\kappa_1^2/4\right) + \kappa_1^2 I_1\left(\kappa_1^2/4\right)\right] \tag{16}$$

that depends in a complicated manner on the inverse coefficients of variation of the numerator and denominator of (1). However, in the high precision limit ($\kappa_{\mathrm{i}} \gg 1$), the leading term in the asymptotic expansion for the modified Bessel functions

$$I_v(z) \sim \frac{\mathrm{e}^z}{\sqrt{2\pi z}} \tag{17}$$

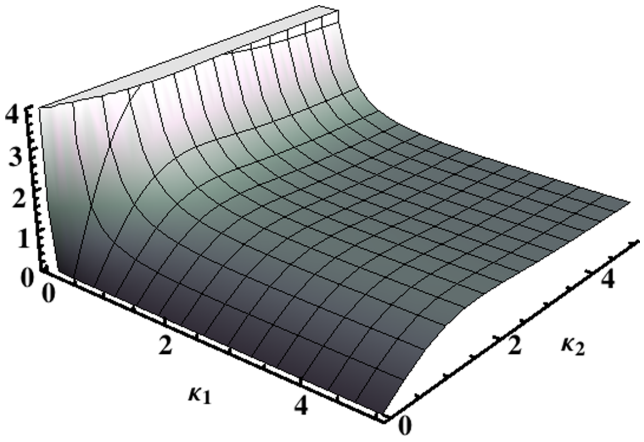

Figure 1. Surface plot of the non-dimensional expected value of the Swift skew distribution given by (16) as a function of the inverse coefficient of variation (mean divided by standard deviation) of the numerator κ_1 and denominator κ_2. The non-dimensional expected value is 1 when the dimensional expected value attains its population value μ_1/μ_2.

may be substituted, revealing that $\mathcal{E}[\zeta'] \to 1$, and hence the expected value of ζ is the population value $|\mu_1/\mu_2|$ for precise estimates of the numerator and denominator of (1). This breaks down for small κ_i. Fig. 1 shows the expected value of the normalized skew (16) as a function of κ_1 and κ_2 over the range [0, 5]. As the parameters simultaneously become large compared to one, the expected value becomes close to but slightly larger than unity; for example, ζ' is 1.0429, 1.0102 and 1.0001 as κ_1 and κ_2 are both 5, 10 and 100, respectively. However, as κ_1 becomes small, the expected value grows without bound regardless of the value of κ_2, and as κ_2 becomes small, the expected value goes to zero regardless of the value of κ_1.

Consequently, when (11) is taken to be the sampling distribution for Swift skew and the population parameters μ_1, μ_2, σ_1 and σ_2 are replaced by method of moments or maximum likelihood estimates, the sample equivalent of (1) can be biased upwards or downwards without bound unless its numerator and denominator are both sufficiently precise. Further, the sample Swift skew is always slightly upward biased even when the numerator and denominator are simultaneously precise. These bias properties may at least in part explain the unreliability as a dimensionality indicator that has frequently been ascribed to Swift skew, especially because upward bias will be apparent when $\hat{z}_{xx}$ and $\hat{z}_{yy}$ are small, as is the case for undistorted 1- or 2-D structures. As a result, discrimination of 1- and 2-D from 3-D structures using the Swift skew may be impaired even in the absence of galvanic distortion.

Since (14) is proportional to $1/z^{n+2}$, the moments of the Swift skew distribution of order two or more do not exist because

$$\lim_{x\to\infty} \int_0^x z^{1-n} \mathrm{e}^{-\kappa_2^2 z^2/2} I_0(\kappa_2^2 z) \mathrm{d}z = \infty \tag{18}$$

when $n \geq 2$. Consequently, the variance of the Swift skew distribution is infinite.

Fig. 2 shows the Swift skew distribution (12) for three pairs of the parameters κ_1 and κ_2. The distribution is skewed to the right with a mode that shifts as the parameters vary in accordance with Fig. 1. It can be demonstrated empirically (and will be shown exactly in the next section) that the upper tail of the Swift skew distribution is algebraic and proportional to $1/\zeta^3$ rather than exponential by presenting the probability density function on a log–log plot.

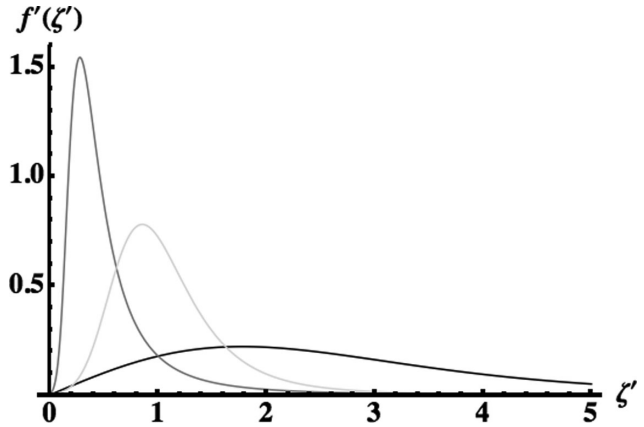

Figure 2. The probability density function (12) as a function of the non-dimensional skew for three pairs of the parameters κ_1 and κ_2 as follows: black line (0.5, 3.0), grey line (3.0, 0.5) and light grey line (3.0, 3.0). The corresponding non-dimensional expected values are 2.88, 0.62 and 1.14, respectively. Note that the distribution is skewed to the right, and that the location of the mode varies as the parameters change.

3 THE STATISTICAL DISTRIBUTIONS OF OTHER ROTATIONAL INVARIANTS

Heuristically, the unusual behaviour of the Swift skew distribution (11) occurs because (1) is a ratio and hence possesses a singularity when the denominator vanishes. Statistically, if the denominator is poorly characterized and hence highly variable, it may at times approach zero such that (1) is undefined. This phenomenon is sufficiently serious that the corresponding probability distribution is affected. Two simple examples serve as illustrations. It can easily be shown that if the denominator in a ratio is Gaussian, then none of the integer moments exist regardless of the distribution of the numerator. If the numerator is also Gaussian, and both the numerator and denominator possess zero mean, then the distribution for the ratio is Cauchy, appearing like a normal distribution but with $1/x^2$ rather than exponential tails. The Cauchy distribution has no integer moments, and in addition has the property that if the mean is estimated from a set of Cauchy random variables, it has the same distribution as each of them. This unusual behaviour is contrary to that of 'typical' random variables, where the mean becomes more precise as the number of data used to estimate it increases.

The Swift skew distribution can alternately be derived as the product of the numerator and the inverted distribution for the denominator using (10). Lehmann & Schaffer (1988) review the characteristics of inverted distributions, the most relevant of which is a proof that the inverted probability density function $g(y)$ of the random variable $Y = 1/X$ given the original distribution $f(x)$ has the property

$$\lim_{y\to\infty} \frac{g(y)}{1/(1+y^2)} = f(0^+). \tag{19}$$

Consequently, when the left-hand side of (19) tends to a finite constant, the inverted distribution has a right Cauchy tail, and is lighter or heavier than Cauchy when $f(0^+)$ is zero or infinity. An analogous relationship holds for the left tail. This establishes Cauchy tails (hence infinite mean and variance) for the inverted distributions of many common random variables, such as Gaussian, Student's t and exponential ones.

However, when $f(0^+) = 0$, as is the case for the Rice distribution (8), the infinite variance property does not always hold; suitable examples with finite variance include many inverted gamma or beta distributions. It is straightforward to show that the inverted distribution of $\xi = 1/\eta$ for the Rice distribution is

$$\bar{h}_{\mathrm{d}}\left(\xi|\mu_2,\sigma_2^2\right) = \frac{1}{\sigma_2^2\xi^3}\exp\left[-\left(1/\xi^2+\mu_2^2\right)/2\sigma_2^2\right] I_0\left(\frac{\mu_2}{\sigma_2^2\xi}\right). \tag{20}$$

The right tail of (20) is algebraic, falling off asymptotically as $1/\xi^3$, and it possesses a finite first moment (hence expected value) but an infinite second moment (hence variance). It is this property that leads to the corresponding behaviour of (11).

Since the denominator of the phase-sensitive skew of Bahr (1988) is identical to that for the Swift skew (1), it follows directly that its distribution also has algebraic tails. The numerator for the phase-sensitive skew contains the sum of product terms like (10) whose distributions are analytically intractable except in the geophysically uninteresting case of zero mean for all variables, and hence a closed form expression for the phase-sensitive skew distribution does not appear to be feasible. The phase-sensitive skew distribution can easily be simulated by drawing independent random Gaussian variates for the real and imaginary parts of the magnetotelluric responses, transforming them into the parameters S_1, S_2, D_1 and D_2 given by Simpson & Bahr (2005, section 5.2), and then computing the phase-sensitive skew using their eq. (5.32). As an example, the synthetic model from Weaver *et al.* (2000) at Site 3 for a period of 1000 s was used. The real and imaginary parts of the magnetotelluric response tensor were obtained using 10 000 random draws from Gaussian distributions with the means given by the Weaver *et al.* (2000) model and a common variance. The value of the variance is unimportant, although using a larger value samples the parameter space more thoroughly, and is preferred. Fig. 3 shows a kernel density estimate for the probability density function after trimming the abscissa. The kernel density estimator applies a Gaussian window at 100 points across a histogram to produce a smoothed estimate for the probability density function, and can include constraints on the support of the variable. The distribution for phase-sensitive skew displays

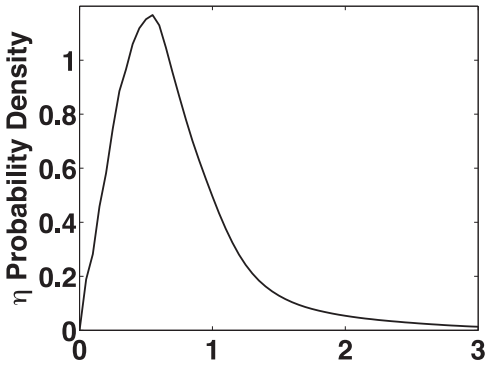

Figure 3. Kernel density estimate for the probability density function of the phase-sensitive skew computed using 10 000 independent random draws from the normal distribution for the real and imaginary parts using the parameters described in the text, followed by computation of the phase-sensitive skew statistic from them. The kernel density smoother was Gaussian with a bandwidth of 0.1004. The abscissa was truncated at 3 for plotting purposes, but the largest value in the simulation was about 76, reflecting the infinite variance nature of the phase-sensitive skew distribution.

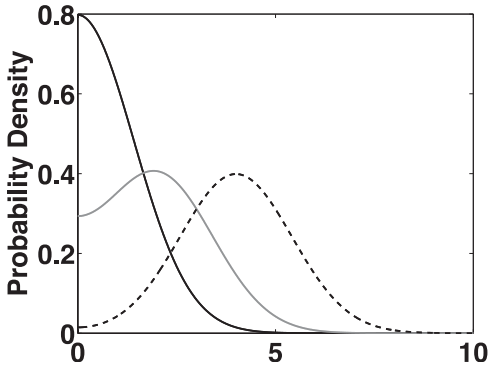

Figure 4. The folded Gaussian distribution (24) for the WAL invariants I_1 and I_2 with the population standard deviation σ_ξ set to unity and the population means μ_{ξ_1} and μ_{ξ_4} set to 0 (black line), 1 (grey line) and 2 (black dashed line).

a very long tail, as has been predicted. It can be shown that the right tail of Fig. 3 is asymptotically approximately $1/x^3$, and consequently the phase-sensitive skew distribution has a mean but no variance or higher order moments.

The distributions for the first two WAL invariants I_1 and I_2 can be derived using the starting model and approaches from Section 2. Using the notation of Weaver *et al.* (2000)

$$\begin{aligned}
\xi_1 + \mathrm{i}\,\eta_1 &= \left(Z_{xx} + Z_{yy}\right)/2, \\
\xi_2 + \mathrm{i}\,\eta_2 &= \left(Z_{xy} + Z_{yx}\right)/2, \\
\xi_3 + \mathrm{i}\,\eta_3 &= \left(Z_{xx} - Z_{yy}\right)/2, \\
\xi_4 + \mathrm{i}\,\eta_4 &= \left(Z_{xy} - Z_{yx}\right)/2,
\end{aligned} \tag{21}$$

it follows from (5) that

$$\begin{aligned}
f_{\xi_1}(y) &= \frac{1}{\sqrt{2\pi}\sigma_{\xi_1}} \mathrm{e}^{-\left(y-\mu_{\xi_1}\right)^2/\left(2\sigma_{\xi_1}^2\right)}, \\
f_{\xi_4}(y) &= \frac{1}{\sqrt{2\pi}\sigma_{\xi_4}} \mathrm{e}^{-\left(y-\mu_{\xi_4}\right)^2/\left(2\sigma_{\xi_4}^2\right)},
\end{aligned} \tag{22}$$

where $\mu_{\xi_1} = \mathcal{R}e[(\mu_{xx} + \mu_{yy})/2]$, $\mu_{\xi_4} = \mathcal{R}e[(\mu_{xy} - \mu_{yx})/2]$, $\sigma_{\xi_1}^2 = (\sigma_{xx}^2 + \sigma_{yy}^2)/8$ and $\sigma_{\xi_4}^2 = (\sigma_{xy}^2 + \sigma_{yx}^2)/8$. Using (4), the distribution for $\xi_1 + \xi_4$ is

$$f_1(x) = \frac{1}{\sqrt{2\,\pi}\sigma_\xi} \mathrm{e}^{-\left(x-\mu_{\xi_1}-\mu_{\xi_4}\right)^2/\left(2\sigma_\xi^2\right)}, \tag{23}$$

where $\sigma_\xi^2 = \sigma_{\xi_1}^2 + \sigma_{\xi_4}^2$. Transforming (23) to the distribution of $z = |x|$ follows from first principles: for $z \geq 0$, $\Pr[|x| < z] = \Pr[-z < x < z] = \Pr[x < z] - \Pr[x < -z] = F(z) - F(-z)$, where F is the cumulative distribution function. Taking the derivative gives the probability density. For the WAL invariant I_1, this results in a folded Gaussian distribution:

$$f_{I_1}(z) = f_1(z) + f_1(-z) = \frac{1}{\sqrt{2\,\pi}\sigma_\xi} \left(\mathrm{e}^{-\left(z-\mu_{\xi_1}-\mu_{\xi_4}\right)^2/\left(2\sigma_\xi^2\right)} + \mathrm{e}^{-\left(z+\mu_{\xi_1}+\mu_{\xi_4}\right)^2/\left(2\sigma_\xi^2\right)} \right), \tag{24}$$

where $z \in [0, \infty)$. The distribution for the WAL invariant I_2 has the same form with μ_{ξ_i} and σ_ξ replaced by μ_{η_i} and σ_η. Both of these distributions are well behaved in the sense that their mean and variance exist, along with higher order moments. Eq. (24) can be used for statistical inference about the first two WAL invariants after the parameters are replaced with method of moments or maximum likelihood estimates obtained from data. Fig. 4 shows (24) in standard form ($\mu_{\xi_1} = \mu_{\xi_4} = 0$, $\sigma_\xi = 1$) and with the mean non-zero.

The expected value of (24) is

$$\mathcal{E}[z] = \left(\sqrt{\frac{2}{\pi}} \mathrm{e}^{-\frac{\left(\mu_{\xi_1}+\mu_{\xi_4}\right)^2}{2\sigma_\xi^2}} + \frac{\left(\mu_{\xi_1} + \mu_{\xi_4}\right) \mathrm{erf}\left[\left(\mu_{\xi_1} + \mu_{\xi_4}\right)/\left(\sqrt{2}\sigma_\xi\right)\right]}{\sigma_\xi} \right) \sigma_\xi, \tag{25}$$

where erf(x) is the error function. For precise parameters [$(\mu_{\xi_1} + \mu_{\xi_4})/\sigma_\xi >> 1$], the expected value is the population value and hence is unbiased, but for imprecise data it will be upward biased without bound.

Prior to investigating the distributions for the remaining WAL invariants, it is instructive to use simulations to define their properties. The elements of (21) and then the WAL invariants $I_3 - I_7$ and Q were computed from random Gaussian draws using the Weaver *et al.* Model 3 parameters and a common variance, and a kernel density estimator was applied to obtain the empirical probability density function. Figs 5–7 show the results for I_3–I_4, I_5–I_6 and Q–I_7, respectively. The empirical probability densities for I_3 and I_4 (Fig. 5) are sharply peaked at

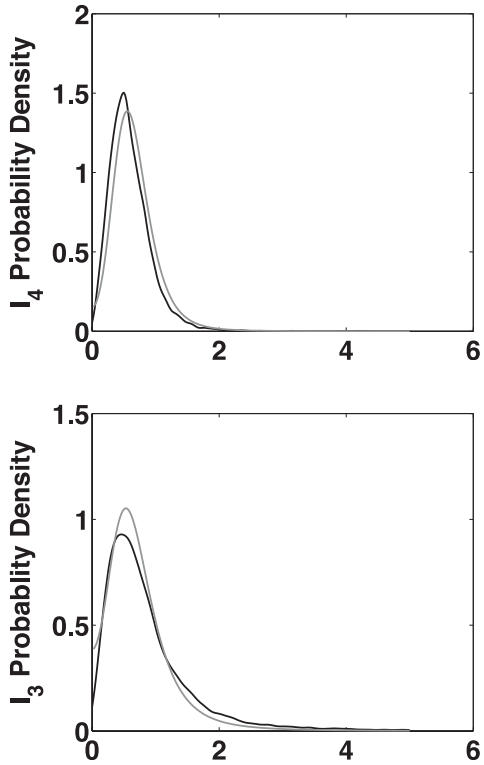

Figure 5. Kernel density estimates (dark lines) for the probability density functions of the WAL invariants I_3 (bottom) and I_4 (top) computed using 10 000 independent random draws from the normal distribution for the real and imaginary parts using the parameters described in the text, followed by computation of the WAL invariant statistics from them. The kernel density smoother was Gaussian with a bandwidth of 0.13. The abscissas were truncated at 5 for plotting purposes, but the largest values in the simulation were about 85 and 55, respectively, implying an infinite variance form for their distributions. The grey lines are (27) with the parameters estimated using the method of moments on the random draws.

abscissa values of about 0.5, and display long right tails. While the geometric interpretation of the WAL invariants in terms of Mohr circles requires that the support of I_3 and I_4 be [0, 1], this and many other simulations do not provide corroboration, and in fact suggest that the support is $[0, \infty)$. This assertion will be proved later when the distributions for I_3 and I_4 are derived. In contrast, the simulations for I_5 and I_6 (Fig. 6) yield probability densities whose support is $[-1, 1]$ and concentrated at the ends of the range. Because their support is finite, all of the moments of their distributions exist. This result is consistently observed for I_5 and I_6 over many simulations, and hence these WAL invariants can be represented trigonometrically. The simulation for I_7 (Fig. 7) is broadly peaked at the origin but with very long bilateral tails. Finally, the simulation for the dependent invariant Q is very long right tailed, and resembles those for I_3 and I_4. As a consequence, neither I_7 nor Q can be represented as trigonometric functions or using Mohr circles.

The WAL invariants I_1 and I_2 are also the denominators in the expressions for the third and fourth WAL invariants and for the dependent invariant Q. It is straightforward to show that the inverted distribution for (24) has Cauchy tails using (19), and hence the distributions for I_3, I_4 and Q will have algebraic tails.

The distributions for I_3 and I_4 can be expressed in closed form. Using (24), a corresponding expression for the numerator of I_3 and (9), the distribution for the WAL invariant I_3 is

$$f_{I_3}(x) = \frac{1}{2\,\pi\,\sigma_\xi^2} \int_0^\infty z \left(\mathrm{e}^{-(zx-\mu_b)^2/\left(2\sigma_\xi^2\right)} + \mathrm{e}^{-(zx+\mu_b)^2/\left(2\sigma_\xi^2\right)} \right) \left(\mathrm{e}^{-(z-\mu_a)^2/\left(2\sigma_\xi^2\right)} + \mathrm{e}^{-(z+\mu_a)^2/\left(2\sigma_\xi^2\right)} \right) \mathrm{d}z, \tag{26}$$

where $\mu_a = \mu_{\xi_1} + \mu_{\xi_4}$ and $\mu_b = \mu_{\xi_2} + \mu_{\xi_3}$. Performing the integration yields

$$\begin{aligned} f_{I_3}(x) = {} & \frac{1}{2\,\pi\,\sigma_\xi\,(1+x^2)^{3/2}} \exp\left(-\frac{\mu_a^2+\mu_b^2}{2\sigma_\xi^2}\right) \times \\ & \left[\sqrt{2\pi}\,(\mu_a - \mu_b x) \exp\left(\frac{(\mu_a-\mu_b x)^2}{2\sigma_\xi^2\,(1+x^2)}\right) \operatorname{erf}\left(\frac{\mu_a-\mu_b x}{\sqrt{2}\sigma_\xi\sqrt{1+x^2}}\right) + \right. \\ & \sqrt{2\pi}\,(\mu_a + \mu_b x) \exp\left(\frac{(\mu_a+\mu_b x)^2}{2\sigma_\xi^2\,(1+x^2)}\right) \operatorname{erf}\left(\frac{\mu_a+\mu_b x}{\sqrt{2}\sigma_\xi\sqrt{1+x^2}}\right) + \\ & \left. 4\sigma_\xi\sqrt{1+x^2} \right]. \end{aligned} \tag{27}$$

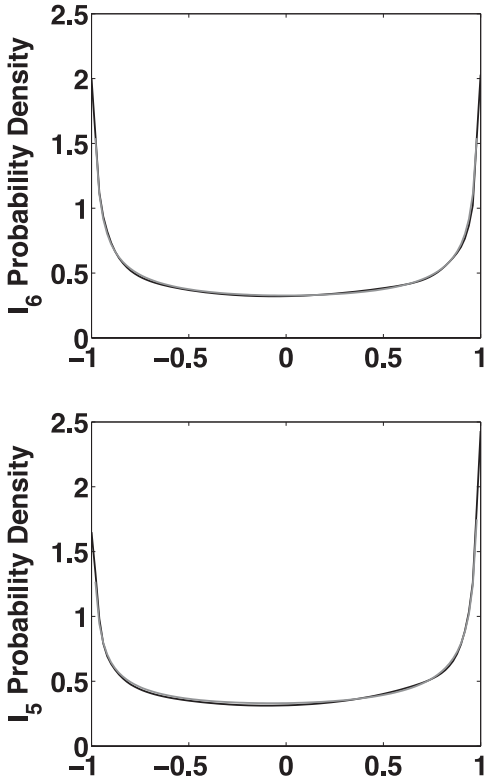

Figure 6. Kernel density estimates for the probability density functions of the WAL invariants I_5 (bottom) and I_6 (top) computed using 10 000 independent random draws from the normal distribution for the real and imaginary parts of the parameters described in the text, followed by computation of the WAL invariant statistics from them. The kernel density smoother was Gaussian with a bandwidth of 0.41. The support for the distributions is $[-1, 1]$, as is readily apparent in the figure. The grey line that is barely distinguishable from the empirical pdf is the beta probability density with parameters (0.52, 0.52) and (0.57, 0.50) for I_5 and I_6, respectively.

The distribution for I_4 is (27) with ξ_j replaced by the corresponding imaginary part η_j. Fig. 5 shows (27) for the Weaver *et al.* model using method of moments estimators for the parameters. It can easily be shown that (27) has an asymptotic $1/x^3$ right tail, and hence the distributions for I_3 and I_4 have a mean but no variance or higher order moments.

The probability distributions for the WAL invariants I_5 and I_6 cannot easily be expressed in closed form because the terms in both the numerator and denominator are products. However, the kernel density probability densities of Fig. 6 closely resemble beta distributions with parameters that lie between 0 and 1, and particularly the arcsine distribution where the parameters are both 1/2. Fig. 8 shows quantile–quantile plots for the simulations of I_5 and I_6 of Fig. 6 after standardizing them to lie on [0, 1] and using method of moments estimators for the beta function parameters. The results are nearly straight lines, supporting the hypothesis that I_5 and I_6 are approximate beta variables, and nearly arcsine variables. Fig. 6 shows the corresponding beta distributions that are nearly indistinguishable from the empirical distributions of I_5 and I_6.

The remaining WAL invariant I_7 is sufficiently complicated that analytic approaches for the inverted distribution are not feasible. However, infinite variance can be intimated through simulation using the results from Fig. 7. The empirical probability density for random values of I_7 consistently exhibits values that are very far from the centre of the distribution, as would be expected for such a distribution. Fig. 9 shows plots of the order statistics for the I_7 simulation against the quantiles for the Cauchy ($1/x^2$ tails) and Student's t with 2 degrees of freedom ($1/|x|^3$ tails). The result should be approximately linear if I_7 has tails like one of these distributions; note that it is the tail behaviour that is of interest, and there is no intent to suggest that I_7 is either Cauchy or t_2. The nearly straight line in the t_2 quantile–quantile plots, and the systematic short tailed form of the Cauchy quantile–quantile plots, suggests that I_7 has approximate $1/|x|^3$ tails, and hence that it has a mean but no variance or higher order moments. Finally, quantile–quantile plots for the dependent invariant Q (not shown) against a folded Student's t_2 distribution also suggest that its tail behaviour is asymptotically $1/|x|^3$.

Consequently, it has been either analytically or empirically demonstrated that the distributions for Swift skew (1), the phase-sensitive skew and the WAL invariants $I_3 - I_4$, I_7 and Q all have infinite variance. This has implications for inference or hypothesis testing when the invariants are statistics that are directly estimated from data and then averaged in some way because the classical central limit theorem does not pertain when the variance is undefined, and hence standard confidence interval estimation or hypothesis testing methods cannot be used.

However, when the rotational invariants are obtained as transformations of the magnetotelluric response tensor $\overleftrightarrow{\mathbf{Z}}$, with the latter an average entity estimated from data, it may be possible to use approximate methods to place error bounds on the invariants. This is the most widely used approach in practice [e.g. see the WALDIM code of Martí *et al.* (2009)]. The dichotomy between ratio of averages and averages

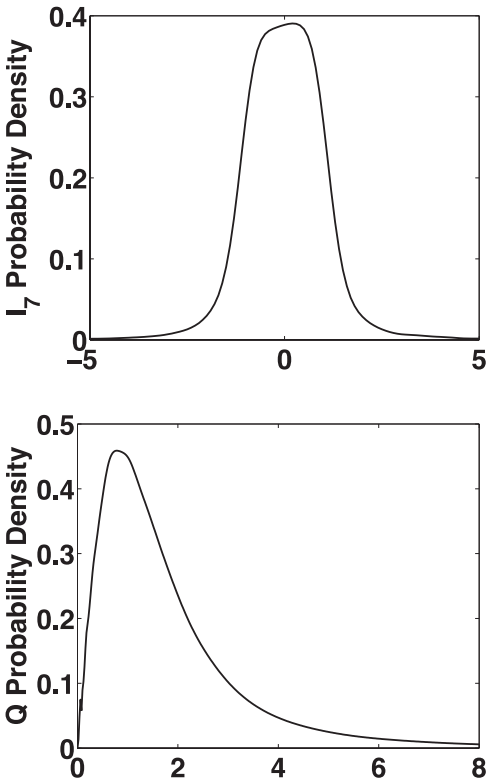

Figure 7. Kernel density estimates for the probability density functions of the WAL invariants Q (bottom panel) and I_7 (top panel) computed using 10 000 independent random draws from the normal distribution for the real and imaginary parts of the response tensor using the parameters described in the text, followed by computation of the WAL invariant statistics from them. The kernel density smoother was Gaussian with a bandwidth of 0.13. The abscissas were truncated at 8 and $[-5, 5]$ for plotting purposes, but the largest values in the simulation were about 85 and $[-45, 32]$, respectively, implying an infinite variance form for their distributions.

of ratios may seem to be inconsistent, but it is not providing that statistical rigor is maintained. If the numerators and denominators of the rotational invariants are first obtained from elements of $\overleftrightarrow{\mathbf{Z}}$ that are themselves averages, then each may be well behaved such that the classical central limit theorem pertains separately to the numerator and denominator of the rotational invariant, and hence either Fieller's theorem or the delta method may be used for statistical inference.

Given the complexity of the distributions for the rotational invariants, it may seem that the preferred approach is the ratio of averages one. However, there are several reasons why this may not make sense. First, approximations based on a Gaussian model are asymptotic results that will break down for an unquantifiably finite number of data, resulting in systematic bias and/or incorrect estimation of statistical uncertainty. Second, robust or bounded influence estimation for $\overleftrightarrow{\mathbf{Z}}$ is standard practice in magnetotellurics, but it is not clear that outlying data will have the same effect on the rotational invariants as they do on $\overleftrightarrow{\mathbf{Z}}$. Direct robust estimation of the rotational invariants is a better approach. Third, computation of statistics with bounded support, such as $I_5 - I_6$, as ratios of averages may not yield a result that is bounded, as will be demonstrated. Fourth, it will be shown that ratio of averages methods can lead to large systematic bias in practice when the numerator and/or denominator are not highly precise. Fifth, statistical inference based on the order statistics is simple to implement for the average of ratios approach, avoiding the complexity of the rotational invariant distributions. Finally, the use of simulations or resampling methods such as the bootstrap with rotational invariants must account for the infinite variance nature of their distributions, or else the result will be meaningless.

4 STATISTICAL INFERENCE FOR ROTATIONAL INVARIANTS

Ratios such as occur in most of the standard rotational invariants are a longstanding problem in statistics for which the estimation of a confidence interval is complicated. The available methods usually are applied to the ratio of the means of two statistics, although they apply equally to ratios constructed from regression parameters. Franz (2007) provides a comprehensive review of the topic.

It is easy to show that the ratio estimator is always biased. Let a ratio estimate $\hat{r}$ be obtained from the ratio of the sample means of the random variables Y and X. Assuming an infinite population, a first-order Taylor series expansion about the population parameters for a sample of size N yields

$$\mathcal{E}[\hat{r}] \approx r + \frac{1}{N\mu_X^2}\left(r\,\sigma_X^2 - cov[X, Y]\right). \tag{28}$$

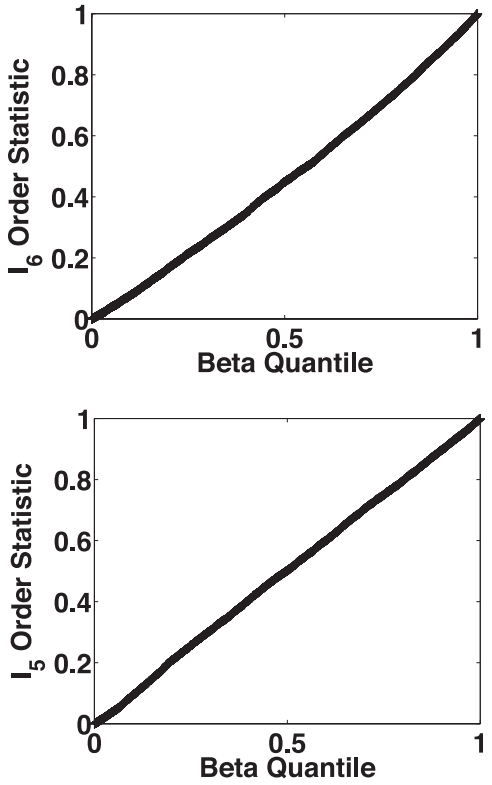

Figure 8. Quantile–quantile plots using order statistics of the simulations of I_5 and I_6 computed with the magnetotelluric model parameters from Fig. 6 after standardization to lie on [0, 1] against standard beta distribution quantiles. The beta distribution parameters were estimated using the method of moments, and are (0.55, 0.49) for I_6 and (0.56, 0.50) for I_5.

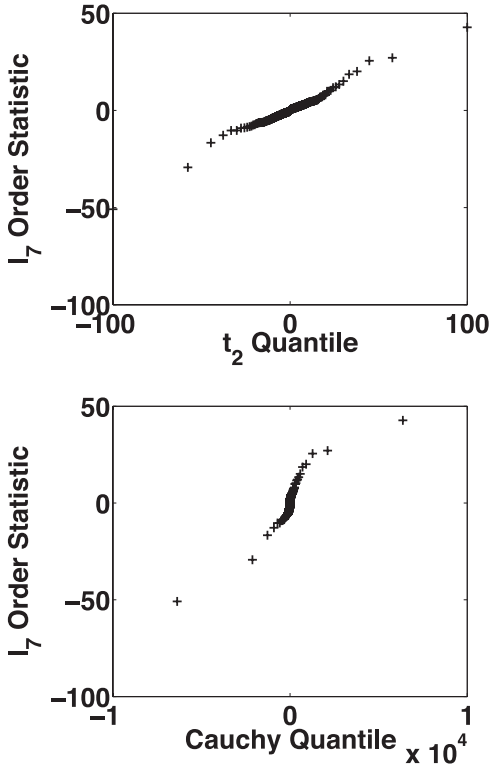

Figure 9. Plots of the order statistics of simulated values of the WAL invariant I_7 against quantiles of the Cauchy distribution (bottom) and the Student t distribution with 2 degrees of freedom (top panel). The simulated values for I_7 are computed from independent normal draws for the magnetotelluric response using the model of Fig. 7. Note that the quantile–quantile plot for t_2 is approximately linear, while the Cauchy quantile–quantile plot is systematically short tailed. This suggests approximate $1/|x|^3$ tails for the distribution of I_7.

Consequently, the sample ratio estimate bias increases if the coefficient of variation (ratio of the standard deviation to the mean) of the denominator becomes large and is reduced when the covariance of the two variables has the same sign as r. A naïve bias-corrected form follows by subtracting the second term of (28) from $\hat{r}$ after replacing the population terms with sample estimates. However, the bias is $O(1/N)$, and hence is small in most magnetotelluric applications where $N > 100$. An improved bias-corrected ratio estimator accurate to $O(1/N^2)$ may be obtained using the jackknife. Neither will be utilized in this study due to the sizes of typical samples. Further, neither (28) nor the jackknife ratio estimator account for other systematic sources of bias that may be much larger.

The most general approach to confidence interval estimation when the rotational invariants are constructed from the response tensor $\overleftrightarrow{\mathbf{Z}}$ is due to Fieller (1940, 1944, 1954). Zerbe (1978) provides a matrix formulation for regression that is especially pertinent. The construction of Fieller confidence limits for the ratio

$$r = \frac{\mathcal{L}_n \left[\mathbf{c}^{\mathrm{T}} \cdot \boldsymbol{\zeta}\right]}{\mathcal{L}_d \left[\mathbf{d}^{\mathrm{T}} \cdot \boldsymbol{\zeta}\right]}, \tag{29}$$

where $\mathbf{c}$ and $\mathbf{d}$ are four vectors of known constants, $\mathcal{L}_{\mathrm{n}}$ and $\mathcal{L}_{\mathrm{d}}$ are either the real or imaginary part operator and ζ is the column vector $(Z_{xx}\ Z_{xy}\ Z_{yx}\ Z_{yy})^{\mathrm{T}}$, follows by noting that since linear combinations of the elements of the real and imaginary parts of ζ are Gaussian, so is $\mathcal{L}_{\mathrm{n}}[\mathbf{c}^{\mathrm{T}} \cdot \zeta] - r\,\mathcal{L}_{\mathrm{d}}[\mathbf{d}^{\mathrm{T}} \cdot \zeta]$. Consequently, dividing the sample version of this term by an estimate of its standard deviation yields a quantity that is distributed as Student's t. That statistic is

$$\hat{T} = \frac{\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}} \cdot \hat{\zeta}\right] - r\,\mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}} \cdot \hat{\zeta}\right]}{\sqrt{\mathcal{V}ar\left(\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}} \cdot \hat{\zeta}\right]\right) - 2r\,\mathcal{C}ov\left(\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}} \cdot \hat{\zeta}\right],\ \mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}} \cdot \hat{\zeta}\right]\right) + r^2\,\mathcal{V}ar\left(\mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}} \cdot \hat{\zeta}\right]\right)}}. \tag{30}$$

$\hat{T}$ is a pivot because (30) has a distribution that does not depend on any of its constituent parameters. Let $t_{\alpha/n}$ be the $1 - \alpha/n$ quantile of Student's t distribution with Θ degrees of freedom, where Θ is the difference between the number of data and the number of parameters after bounded influence weights are applied and n will be specified later. The Fieller confidence limit follows from the definition

$$1 - \alpha/n = \Pr\left[-t_{\alpha/n} \le \hat{T} \le t_{\alpha/n}\right] = \Pr\left[\hat{\alpha}\, r^2 + \hat{\beta}\, r + \hat{\chi} = 0\right], \tag{31}$$

where

$$\begin{aligned}
\hat{\alpha} &= \left(\mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}} \cdot \hat{\zeta}\right]\right)^2 - t^2_{\alpha/n}\,\mathcal{V}ar\left(\mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}} \cdot \hat{\zeta}\right]\right), \\
\hat{\beta} &= 2\left[t^2_{\alpha/n}\,\mathcal{C}ov\left(\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}} \cdot \hat{\zeta}\right],\ \mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}} \cdot \hat{\zeta}\right]\right) - \left(\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}} \cdot \hat{\zeta}\right]\right)\left(\mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}} \cdot \hat{\zeta}\right]\right)\right], \\
\hat{\chi} &= \left(\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}} \cdot \hat{\zeta}\right]\right)^2 - t^2_{\alpha/n}\,\mathcal{V}ar\left(\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}} \cdot \hat{\zeta}\right]\right),
\end{aligned} \tag{32}$$

and corresponds to the set of r values for which $\hat{T}$ lies within the $1 - \alpha/n$ interquartile range of Student's t. This is given by

$$\frac{-\hat{\beta} - \sqrt{\hat{\beta}^2 - 4\,\hat{\alpha}\,\hat{\chi}}}{2\hat{\alpha}} \le r \le \frac{-\hat{\beta} + \sqrt{\hat{\beta}^2 - 4\,\hat{\alpha}\,\hat{\chi}}}{2\hat{\alpha}}, \tag{33}$$

provided that $\hat{\alpha} \ge 0$ and $\hat{\beta}^2 - 4\hat{\alpha}\hat{\chi} \ge 0$. Eq. (33) is asymmetric about the sample estimate for r.

Eq. (33) is the solution to (31) when the square of the denominator of (29) divided by its variance is significant, or $(\mathcal{L}_{\mathrm{d}}[\mathbf{d}^{\mathrm{T}} \cdot \hat{\zeta}])^2 / \mathcal{V}ar(\mathcal{L}_{\mathrm{d}}[\mathbf{d}^{\mathrm{T}} \cdot \hat{\zeta}]) > t^2_{\alpha/n}$. There are two other cases: when $\hat{\beta}^2 - 4\hat{\alpha}\hat{\chi} \ge 0$ and $\hat{\alpha} < 0$, the confidence interval is the complement of (33) given by

$$\left(-\infty,\ \frac{-\hat{\beta} + \sqrt{\hat{\beta}^2 - 4\,\hat{\alpha}\,\hat{\chi}}}{2\hat{\alpha}}\right] \le r \le \left[\frac{-\hat{\beta} - \sqrt{\hat{\beta}^2 - 4\,\hat{\alpha}\,\hat{\chi}}}{2\hat{\alpha}},\ \infty\right), \tag{34}$$

while if $\hat{\beta}^2 - 4\hat{\alpha}\hat{\chi} < 0$ and $\hat{\alpha} < 0$, the confidence interval is $(-\infty, \infty)$. In both of these cases, the confidence interval is infinite, and little or nothing can be said about r. The occurrence of infinite confidence intervals is a consequence of (29) being a ratio for which the denominator becomes arbitrarily close to zero, and Von Luxburg & Franz (2009) give an intuitive explanation. Eq. (34) pertains when the denominator of (29) is not significantly different from zero but the numerator is well defined. In that instance, as the numerator is divided by a number that is close to zero, the absolute value and sign of the result are uncontrolled, leading to a disjoint pair of infinite confidence intervals. If the numerator of (29) is also zero within the statistical uncertainty, then any result is possible, as zero divided by zero is undefined. In either case, the statistic of interest is not well specified, and little can be concluded from the data.

However, none of the magnetotelluric rotational invariants are the real or imaginary parts of linear combinations of the elements of $\mathbf{Z}$, so the Fieller method cannot be applied directly for small numbers of data, except for the WAL invariants I_3 and I_4 where the absolute value of (29) is the statistic of interest, and the Swift and phase-sensitive skews. In these cases, the operators $\mathcal{L}_{\mathrm{n}}$ and $\mathcal{L}_{\mathrm{d}}$ become the absolute value of the real or imaginary parts. Using (31), (29) is no longer distributed as Student's t, but rather approximates a folded version of the t distribution. The Fieller methodology may then be applied using the $1 - \alpha/n$ quantile of the folded Student t distribution for $t_{\alpha/n}$ in (31) and (32).

For the WAL invariants $I_5 - I_6$, the ratio of interest may be written as

$$s = \frac{\mathcal{L}_1[\zeta]}{\mathcal{L}_2[\zeta]}, \tag{35}$$

where $\mathcal{L}_1$ and $\mathcal{L}_2$ are operators that yield, respectively, product combinations of the real and imaginary parts of ζ and the absolute value of sums of the real and imaginary parts of ζ. For small numbers of data, the numerator of (35) is clearly non-Gaussian, and hence Fieller's method does not pertain. In the asymptotic limit, the numerator and denominator of (35) become Gaussian by the classical central limit theorem, and hence Fieller's method can be used, with the usual caveat that it is not possible to define the number of data required to reach the realm of validity. A similar approach can be used with the WAL invariant I_7.

A widely used alternative to Fieller's method is the delta method that is based on a Taylor series approximation to the variance of functions of random variables. This can be applied directly to a ratio y/x by expanding that quantity in a two-variable, first-order Taylor series and then taking the variance of the result, yielding

$$var\left(\frac{y}{x}\right) \approx \frac{\mathcal{E}(y)^2}{\mathcal{E}(x)^2}\left[\frac{var(y)}{\mathcal{E}(y)^2} - \frac{2cov(y,\,x)}{\mathcal{E}(x)\,\mathcal{E}(y)} + \frac{var(x)}{\mathcal{E}(x)^2}\right]. \tag{36}$$

Applying this to (29) gives

$$var\left(\frac{\mathcal{L}_n\left[\mathbf{c}^{\mathrm{T}}\cdot\hat{\boldsymbol{\xi}}\right]}{\mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}}\cdot\hat{\mathbf{z}}\right]}\right) \approx \frac{\left(\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}}\cdot\hat{\boldsymbol{\xi}}\right]\right)^2}{\left(\mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}}\cdot\hat{\boldsymbol{\xi}}\right]\right)^2}\left[\frac{var\left(\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}}\cdot\hat{\boldsymbol{\xi}}\right]\right)}{\left(\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}}\cdot\hat{\boldsymbol{\xi}}\right]\right)^2} - 2\frac{cov\left(\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}}\cdot\hat{\boldsymbol{\xi}}\right],\;\mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}}\cdot\hat{\boldsymbol{\xi}}\right]\right)}{\left(\mathcal{L}_{\mathrm{n}}\left[\mathbf{c}^{\mathrm{T}}\cdot\hat{\boldsymbol{\xi}}\right]\right)\left(\mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}}\cdot\hat{\boldsymbol{\xi}}\right]\right)} + \frac{var\left(\mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}}\cdot\hat{\boldsymbol{\xi}}\right]\right)}{\left(\mathcal{L}_{\mathrm{d}}\left[\mathbf{d}^{\mathrm{T}}\cdot\hat{\boldsymbol{\xi}}\right]\right)^2}\right]. \tag{37}$$

It is common practice (e.g. Simpson & Bahr 2005, eq. 5.18) to neglect the middle covariance term in (37), although this can lead to large, undetected errors, a point that was also emphasized by Booker (2013) in the context of the phase tensor.

Delta method confidence intervals using the Gaussian z-statistic may be computed with the standard error computed from (37). Such a confidence interval will always be symmetric about $\hat{r}$ and cannot be unbounded, in contrast to the Fieller confidence interval. The delta method will break down when the denominator of (29) or (35) statistically approaches zero, and will severely underestimate the confidence interval as that limit is approached. However, the delta method produces a reasonable approximation to the Fieller confidence interval when the denominator is sufficiently precise (Cox 1990; Franz 2007).

Hirschberg & Lye (2010) provide a lucid geometric representation and comparison of the Fieller and delta methods, and suggest that the delta method produces a good approximation to the Fieller interval when the denominator of (29) is precise and the signs of (29) and the covariance of the numerator and denominator in (37) are the same, but is less accurate when their signs differ, as was also shown by Herson (1975). They also indicate that, when their results differ, the Fieller interval provides better coverage, and hence is preferred.

Gleser & Hwang (1987) proved that any statistical method that is not able to produce infinite confidence intervals for a ratio will lead to arbitrarily large errors, or conversely, will have coverage probability that is arbitrarily small. For this reason, the Fieller approach is preferred to the delta method, and given that it does not require a substantial increase in computational burden, it is recommended that it be more widely used by the magnetotelluric community.

Implementation of either the Fieller or delta methods for the rotational invariants requires the full covariance matrix for the elements $(\xi_i,\,\eta_j)$ in (21). Standard magnetotelluric codes typically do not provide sufficient information to compute these quantities, although implementation is straightforward. Define the 4×4 complex covariance and pseudo-covariance matrices

$$\begin{aligned}\Sigma &= cov\left(Z^*_{jk},\;Z_{mn}\right),\\ \breve{\Sigma} &= cov\left(Z_{jk},\;Z_{mn}\right),\end{aligned} \tag{38}$$

where j, k, m, n can be either x or y in any order and * denotes the complex conjugate; for example, the first row of Σ is the four elements $[var(Z_{xx})\;cov(Z^*_{xx},\;Z_{xy})\;cov(Z^*_{xx},\;Z_{yx})\;cov(Z^*_{xx},\;Z_{yy})]$, where the last three elements are complex. Due to symmetry, only the upper triangles of Σ and $\breve{\Sigma}$ need be estimated. If the response functions comprise proper complex numbers, then the pseudo-covariance is identically zero, although this rarely holds for real data. Expanding the response elements into their real and imaginary parts, it is easy to show that

$$\begin{aligned}cov\left(Z^{\mathrm{r}}_{jk},\;Z^{\mathrm{r}}_{mn}\right) + i\;cov\left(Z^{\mathrm{r}}_{jk},\;Z^{\mathrm{i}}_{mn}\right) &= [\Sigma + \breve{\Sigma}]/2,\\ cov\left(Z^{\mathrm{i}}_{jk},\;Z_{mn}\right) - i\;cov\left(Z^{\mathrm{i}}_{jk},\;Z^{\mathrm{r}}_{mn}\right) &= [\Sigma - \breve{\Sigma}]/2.\end{aligned} \tag{39}$$

Consequently, the covariance between the real and/or imaginary elements of any pair of magnetotelluric response components may be obtained from the covariance and pseudo-covariance matrices. Their application to the WAL invariants using (21) and the definitions of $I_3 - I_7$ and Q requires further tedious but straightforward algebraic manipulation.

However, for the WAL invariants I_5–I_7, estimates for the variances of the numerator and denominator of (35) used in the denominator of (30) are also required, which in turn require estimates for the variance of products. The variance of the product of two random variables is given exactly by (Goodman 1960)

$$var\left(\xi_i\,\eta_j\right) = cov\left(\xi_i^2, \eta_j^2\right) + [var(\xi_i) + \mathcal{E}(\xi_i)^2][var\left(\eta_j\right) + \mathcal{E}\left(\eta_j\right)^2] - [cov\left(\xi_i, \eta_j\right) + \mathcal{E}(\xi_i)\,\mathcal{E}\left(\eta_j\right)]^2. \tag{40}$$

In practice, the covariance terms involving the squares of the parameters are difficult to compute parametrically, and hence the first term on the right-hand side of (40) is often neglected, leading to unknown errors. Further, the covariance of the numerator and denominator of (35) is also needed to complete the Fieller statistic $\hat{T}$, and is even more problematic. This makes Fieller's method difficult to apply with complicated statistics like the WAL invariants I_5-I_7 using parametric estimates for the variances and covariance, although this is equally true for its alternatives.

However, all of the elements of (40) and the covariance of the numerator and denominator of (35) can be obtained non-parametrically using the jackknife applied to response function estimates computed by deleting one datum at a time with replacement, as further described by Thomson and Chave (1991) and Chave (2012). Such an approach is advocated in this paper, and will be illustrated in the next section.

Neither the Fieller nor the delta method apply when estimates of the Swift skew, phase-sensitive skew or the WAL invariants $I_3 - I_4$, I_7 and Q are obtained, for example, over adjoining segments of a data set and then combined because of the infinite variance form of their distributions. Alternate location and scale estimators are required to replace the mean and variance. The order statistics $x_{(i)}$ obtained by sorting a data set into an ascending sequence exist for any distribution. The median is the middle-order statistic $x_{(\lfloor N/2 \rfloor+1)}$, where N is the number of data and the floor function $\lfloor z \rfloor$ is the largest integer not greater than z, and serves as a suitable replacement for the mean to estimate location. It can be shown (David 1981, section 2.5) that the confidence interval on a quantile $\xi_p = F^{-1}(p)$, where $F^{-1}(p)$ is the inverse cumulative distribution or quantile function at a probability level p, is independent of the actual data distribution. The probability that ξ_p lies between two order statistics $x_{(r)}$ and $x_{(s)}$ is bounded by

$$\Pr\left[x_{(r)} \le \xi_p \le x_{(s)}\right] = I_p\left(r,\ N-r+1\right) - I_p\left(s,\ N-s+1\right) = \sum_{i=r}^{s-1} \binom{N}{i} p^i \left(1-p\right)^{N-i} \ge 1 - \alpha/n, \tag{41}$$

where $I_p(i,\ j)$ is the incomplete beta function ratio (which is the cumulative distribution function for the beta distribution), $\ge$ on the probability indicates that the confidence interval is conservative [i.e. it holds at least $100(1-\alpha)$% of the time] and α is the total tail probability. For the median, (41) is easily solved by successive approximation for r and s after setting $p = 0.5$. Consequently, a statistically meaningful alternative to the Fieller or delta confidence interval for rotational invariants is available when they are estimated as averages of a ratio. Perhaps most importantly, use of the order statistics does not require parametric computation from complicated distributions like (27).

The bootstrap, and its linear approximation, the jackknife, are statistical methods that enable the construction of confidence intervals for complicated data without imposing statistical assumptions other than independence. However, the bootstrap (and by extension the jackknife) presents two difficulties when the statistic of interest is a ratio. First, all of the standard bootstrap estimators produce only finite length confidence intervals, and are not capable of producing infinite intervals. By the result of Gleser & Hwang (1987), this means that they will produce arbitrarily large errors with ratios under some circumstances. Second, Athreya (1987) and Knight (1989) showed that the bootstrap distribution for sums of random variables with infinite variance converges to a random rather than a fixed distribution, and hence confidence intervals constructed from the bootstrap distribution can be meaningless. Hwang (1995) proposed a bootstrap estimator that overcomes these problems by bootstrapping the pivot $\hat{T}$ in (30) rather than directly bootstrapping the ratio. Hwang first determines the $1-\alpha/n$ quantiles from the bootstrap distribution of $\hat{T}$ and then solves the quadratic equation in (31) to define the confidence interval. Because this bootstrap can yield disjoint or doubly infinite confidence intervals, it is not limited by the result of Gleser & Hwang (1987). Consequently, the Hwang (1995) bootstrap serves as a safe alternative to standard bootstrap confidence interval estimators. However, it typically requires large samples for good performance, and will not be further pursued in this paper.

5 EXAMPLES

Many of the principles enumerated in Sections 2–4 will be illustrated using an exemplar seafloor data set obtained by J.H. Filloux in 1983 offshore from the Bay of Plenty, New Zealand. This region is marked by severe SW–NE lineated topography aligned subparallel to the Kermadec Trench that is located several hundred kilometres to the southeast, and is almost surrounded by nearby land, suggesting that coastline- and bathymetry-induced three dimensionality are substantial. Site E (36°6'S, 177°5'E) provides the horizontal electric and three-component magnetic field variations recorded at 64/h for about 85 d. The horizontal magnetic field at Site B (36°52'S, 177°28'E) is used as a remote reference. The data were processed using the bounded influence estimator of Chave & Thomson (2004). Fig. 10 shows all four components of the magnetotelluric response expressed as apparent resistivity and phase. Three dimensionality is readily apparent, especially in the phase that spans three quadrants for all elements of the response tensor.

It is (unfortunately) common practice to present double-sided confidence intervals with the total tail probability α set to 0.05 for each component (real or imaginary part, or apparent resistivity or phase) for a single element of the magnetotelluric response tensor, or at best to apportion the tail probability between the components of each element. This results in the tail probability for all four tensor elements taken together becoming 0.4 or 0.2, respectively. A better practice is the Bonferroni approach that apportions the tail probability among all of the significant elements, with the simplest implementation being allocation of an equal fraction to each one. For a 2-D structure where the diagonal tensor elements are zero, the antidiagonal elements should each be assigned $\alpha/4$ of the tail probability; double-sided confidence intervals then follow by using the $1-\alpha/8$ quantile of Student's t distribution. When all of the response tensor elements are important, as in this study, each of its elements receives $\alpha/8$ of the tail probability, and the pertinent quantile is the $1-\alpha/16$ one. When the degrees of freedom become sufficiently large that the t distribution transitions to Gaussian, the relevant quantiles are 1.96 when only a single tensor element is

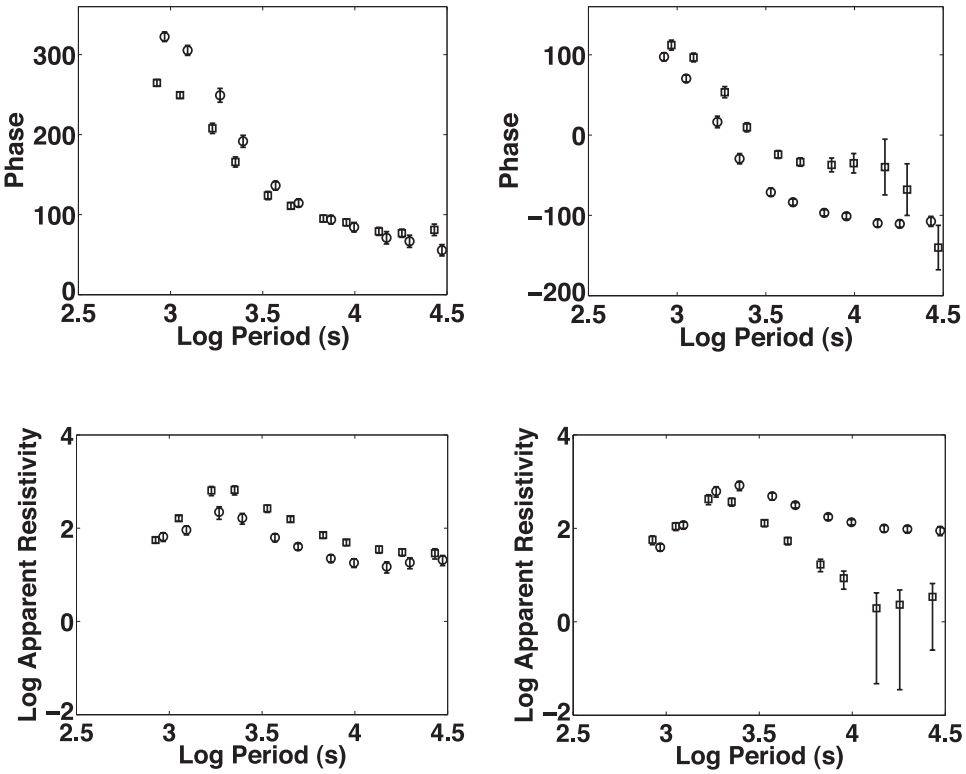

Figure 10. The apparent resistivity and phase for (left-hand panels) Z_{xx} and Z_{xy} and (right-hand panels) Z_{yx} and Z_{yy}. The data are in geomagnetic coordinates. Open squares denote the results for Z_{xx} and Z_{yy}, while open circles represent Z_{xy} and Z_{yx}; Z_{xy} and Z_{yx} have been offset to the right for clarity. The error bars show double-sided 95 per cent confidence limits obtained from the jackknife with the tail probability apportioned equally between apparent resistivity and phase in all four components of the response tensor, so that the total tail probability is 0.05. The apparent resistivity of the diagonal components is non-trivial in both orientations, while the phase spans three quadrants in all components, suggesting strong 3-D effects.

important, 2.49 when four elements are relevant and 2.73 when the entire tensor is of interest. For a folded distribution, the quantiles are 2.24, 2.73 and 2.96, respectively, and hence slightly larger than for an unconstrained distribution. Failure to utilize the correct tail probability will result in under specification of the confidence intervals and potentially to overfitting of models to data. In Fig. 8, the total tail probability is 0.05 across all eight of the elements shown. This comment pertains equally to the rotational invariants and to the response tensor elements themselves, and will be applied for the remainder of this study.

In this work, the upper triangles of Σ and $\tilde{\Sigma}$ in (38) are estimated using the jackknife applied to all four response tensor components, yielding the four real covariance matrices on the left-hand side of (39) by addition and subtraction. This is computationally simpler than obtaining (38) directly, and in addition accounts for the second-order covariance term in (40) without its direct estimation.

Fig. 11 compares the WAL invariant I_4 for the New Zealand data with confidence intervals computed using the Fieller and delta methods for $\alpha = 0.05$. The denominator is quite precise for this WAL invariant, with the inverse squared coefficient of variation of its denominator ranging from several hundred to several thousand. Slight asymmetry of the Fieller interval is apparent, but overall the two methods produce nearly identical confidence intervals. However, there is a suggestion for upward bias when the confidence interval is relatively large, especially at the third shortest period of 1687 s. This bias is not explainable by (28), and constitutes an unknown source of systematic error. It is possible that it is caused by breakdown of the central limit theorem conditions that are implicit to both the Fieller and delta methods, although this would be difficult to prove. Further, neither the estimates nor their confidence intervals in Fig. 11 are confined to [0, 1]. This is required by a Mohr circle interpretation of these invariants, but (27) has already demonstrated that I_3 and I_4 have support of $[0, \infty)$.

Fig. 12 compares the WAL invariant I_6 with confidence intervals computed using the Fieller and delta methods. The denominator of I_6 is less precise than that for I_4, with its inverse squared coefficient of variation ranging from several tens to several hundred. The least precise denominator estimates occur at 1125 and 4500 s, and are evidenced by large error bars and obvious systematic bias away from zero, with the bias evidently proportional to the length of the confidence interval. The delta method consistently underestimates the lower confidence bound and overestimates the upper confidence bound compared to the Fieller method. Further, the systematic bias that occurs for data that by most standards would be regarded as precise is not fully accounted for by the confidence interval estimates. More importantly, the estimates for I_6 sometimes lie outside of its support of $[-1, 1]$, reflecting the fact that taking the ratio of averages of the numerator and denominator after separately passing to a Gaussian limit does not align with the inherent support of the statistic. Consequently, it is possible to obtain physically meaningless results with either the Fieller or delta approaches.

Fig. 13 compares the WAL invariant I_7 computed using the Fieller and delta methods with the tail probability decreased to 0.00001 (or the Gaussian quantile increased to 4.24). This choice forces the longest period estimate, whose denominator inverse squared coefficient of

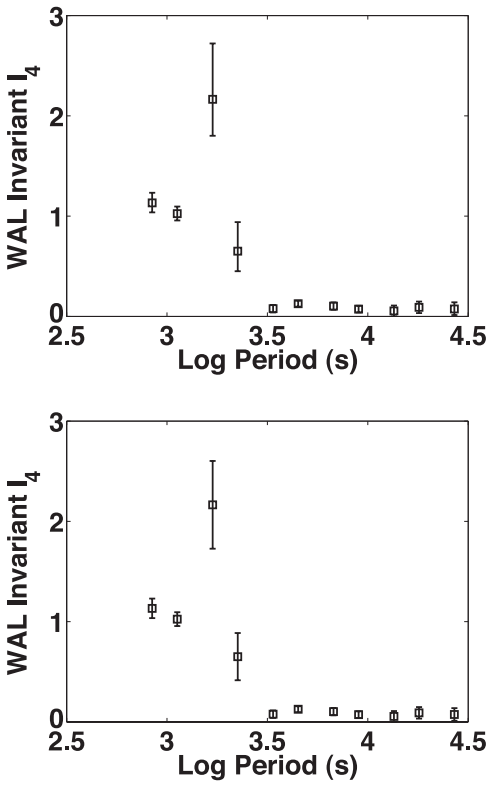

Figure 11. The WAL invariant I_4 for the New Zealand data with confidence intervals estimated using the $1-\alpha/16$ quantile of the folded Gaussian distribution with $\alpha = 0.05$ for the Fieller method (top panel) and the delta method (bottom panel). Note the slight asymmetry in the Fieller confidence intervals, but the overall consistency produced by the two methods due to the high precision of the denominator of the I_4 ratio.

variation is 17.8, to have a disjoint semi-infinite confidence interval given by (34), with the excluded values being $(-200.1, 0.53)$. The delta method gives a confidence interval that is finite but wildly in error by comparison at the same period. Note also the upward bias when the confidence interval is large, as was also seen in Figs 11 and 12.

As a simple implementation of the average of ratios estimation of the WAL invariants, raw estimates were obtained by computing the magnetotelluric response section-by-section over a data set. The raw response functions were then transformed into the section-by-section estimates of the WAL invariants I_3-I_7, and hence they have the statistical properties that were demonstrated in Section 3. The median value is used to estimate location, which has the advantage of providing a degree of robustness. Uncertainty on the median is obtained from the order statistics using (41).

Fig. 14 shows the median estimates for the WAL invariant I_4. The result is a smooth function of period with no evidence for the sharp upward bias at some periods that is apparent in Fig. 11, especially at 1687 s. There is a large discrepancy between Figs 11 and 14 at long periods, with the former showing nearly zero response at periods over $\sim$3000 s, while the median value in Fig. 11 lies between 0.6 and 0.7, although it must be recognized that the two figures display different statistics. However, the empirical probability density (not shown) using the New Zealand data displays all of the characteristics seen in Fig. 5, and is well represented by (27). Given the obvious bias in Fig. 11 and the consistency of the statistical model for I_4, it must be concluded that the ratio of means approach is unreliable in the present context.

Fig. 15 compares the median and mean for the WAL invariant I_6, with confidence intervals estimated using (41) and the jackknife, respectively. Comparison with Fig. 12 displays qualitative similarity at the longest periods, but Fig. 15 does not reproduce the sharp negative excursions seen in Fig. 12 at 1125 and 4500 s, and produces an invariant that is a smooth function of period. The confidence intervals on the mean are systematically smaller than those for the median, reflecting the well-known efficiency properties of these two location estimators; for Gaussian data, the variance of the median is $\pi/2$ times the variance of the mean. Fig. 16 compares the empirical probability densities for I_6 at periods of 4500 and 1687 s. At the longer period, the mean and median are very similar in Fig. 15, and the corresponding probability density is tilted upwards to the lower end of the support. At the shorter period, the mean is consistently closer to the origin compared to the median, and the probability density is much flatter but with a large tail at the lower end. The median will always tend towards the high left peak, while the mean will tend towards the centre as the distribution flattens, qualitatively explaining the difference between the mean and median at short periods in Fig. 15.

Fig. 17 shows the median of the WAL invariant I_7 for the New Zealand data, along with confidence intervals determined from the order statistics. The result is a smooth function of period whose median value (but not confidence interval) is confined to $\sim[0, 1]$. However, this is not due to bounds on the support of I_7, as shown by the empirical probability density function at a period of 1687 s in Fig. 18. As is also seen in Fig. 7, the support of I_7 is $(-\infty, \infty)$, and the probability density in Fig. 18 is bimodal. There is a suggestion of bimodality in Fig. 7 due to

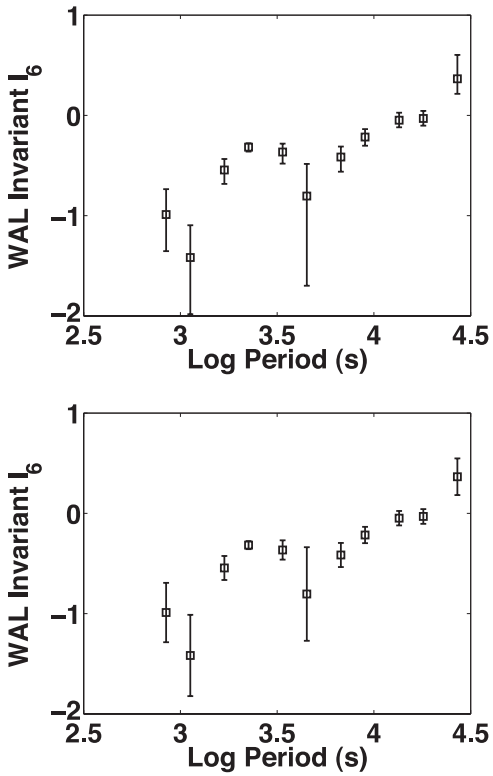

Figure 12. The WAL invariant I_6 for the New Zealand data with confidence intervals estimated using the $1-\alpha/16$ quantile of the Gaussian distribution with $\alpha = 0.05$ for the Fieller method (top panel) and the delta method (bottom panel). Note the strong asymmetry in the Fieller confidence intervals, and the obvious bias at 4500 and 1687 s that is not accounted for by the confidence bounds. Note further that both the estimates and their confidence limits lie outside $[-1, 1]$ at some periods.

the breadth of the peak, and it is possible that reducing the kernel density estimator bandwidth would resolve it. As for the WAL invariants I_4 and I_6, the median of the section-by-section estimates of I_7 does not exhibit the upward bias that is especially evident at 1687 s in Fig. 13.

6 CONCLUSIONS

This paper constitutes a study of the statistics of magnetotelluric rotational invariants, a subject that has not received focused attention in the past. Rotational invariants are the ratios of combinations of elements of the magnetotelluric response tensor, and two approaches to their statistics have been examined in detail. The first is the traditional method of first estimating the magnetotelluric response tensor from data, which inherently involves an averaging process, and then regarding the invariant as a ratio of these averages. The second approach utilizes raw estimates of the rotational invariant that are then averaged in some way.

Statistical inference for the ratio of averages approach requires the assumption that the numerator and denominator in the invariant separately reach the Gaussian limit through classical central limit theorem reasoning. Confidence intervals on the ratio can be estimated either through Fieller's theorem or using the delta method. The former produces exact confidence intervals under some circumstances, and is capable of producing disjoint semi-infinite or infinite confidence intervals when the numerator and denominator of the invariant are not precise. In contrast, the delta method will always produce finite confidence intervals, and agrees with the Fieller interval only for precise data. Given that the computational burden for Fieller's method is only slightly higher than for the delta method, and in view of its higher accuracy, it is recommended that it be more widely used by the magnetotelluric community.

The statistical distributions for the Swift skew, phase-sensitive skew and the WAL invariants have been studied using a Gaussian model for the magnetotelluric response for the case where the rotational invariants are themselves random variables. Analytic expressions for the probability density functions of the Swift skew and the WAL invariants I_1-I_4 have been derived. The WAL invariants I_1-I_2 are distributed as a folded Gaussian, and have well-defined statistical moments, although the mean is upward biased without bound. By contrast, the Swift skew, phase-sensitive skew and I_3-I_4 distributions have algebraically descending right tails, and have a well-defined mean but infinite variance with support of $[0, \infty)$. The statistical distributions of the WAL invariants I_5-I_7 and Q have been studied through simulation. It has been demonstrated that I_7 and Q have algebraically descending tails with support of $(-\infty, \infty)$ and $[0, \infty)$, respectively, and have finite means but infinite variances. By contrast, I_5-I_6 have support of $[-1, 1]$, hence all of their integer moments exist. The invariants I_5-I_6 have a beta distribution that is upward concave and concentrated at the extremes of its support. Because the variance is undefined for Swift skew, phase-sensitive skew and the WAL invariants I_3-I_4, I_7 and Q, confidence intervals on the mean cannot be estimated using the classical central

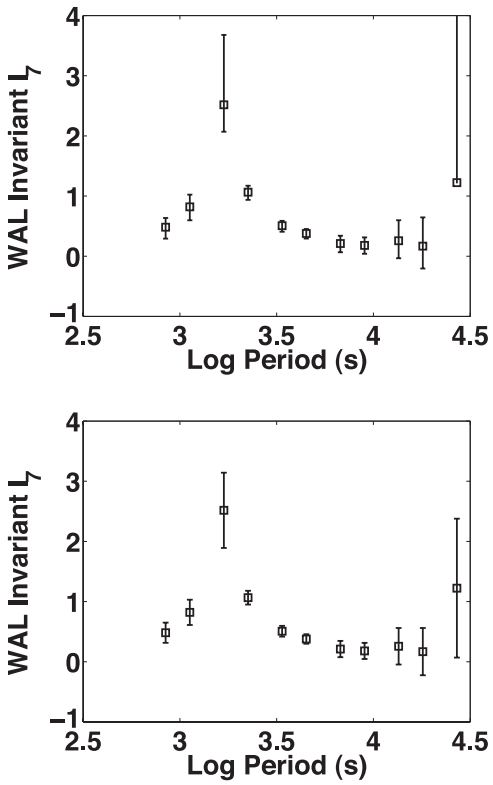

Figure 13. The WAL invariant I_7 for the New Zealand data with confidence intervals estimated using the Fieller (top panel) and delta (bottom panel) methods with the tail probability reduced to 0.00001. This choice forces the longest period estimate to have a disjoint semi-infinite Fieller confidence interval that is undetectable using the delta method. The second semi-infinite confidence interval for the longest period Fieller estimate is $(-\infty, -200.1]$, and hence not shown.

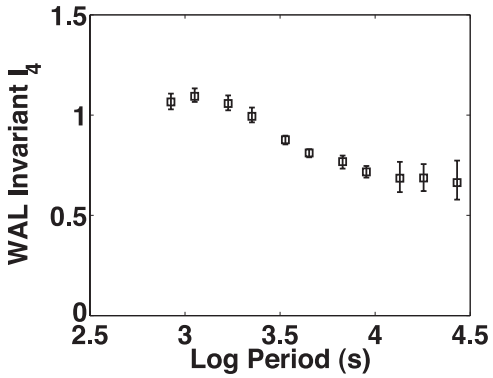

Figure 14. The median value of the WAL invariant I_4 computed from section-by-section estimates of the magnetotelluric response for the New Zealand data set. The error bars are $1 - \alpha/16$ values from (41) with $\alpha = 0.05$.

limit theorem. Instead, use of the median rather than the mean is advocated, and exact confidence intervals of the median can be bounded from the order statistics.

These results are illustrated using a seafloor data set collected in a coastal environment with very strong topography that is markedly 3-D, exhibiting a phase in all four components that spans three quadrants. Estimation of the rotational invariants using the ratio of averages method shows consistency of the Fieller and delta method confidence intervals for precise data, but marked differences when they are not. More importantly, the results display substantial upward (i.e. away from the origin) systematic bias even for data that would conventionally be regarded as precise. The bias is not of statistical [i.e. $O(1/N)$] origin, and may be due to undetectable breakdown of central limit theorem conditions. Since the bias is both systematic and large, it renders useless ratio of averages estimates for the rotational invariants in this instance. The ratio of averages approach is also shown to give estimates that are not confined to the finite support of the WAL invariants $I_5 - I_6$, yielding inconsistent and unphysical results.

Empirical probability density functions estimated from the seafloor data are consistent with those from theory and simulations. Average of ratios estimates were computed by section-by-section computation of the magnetotelluric response and its transformation into rotational

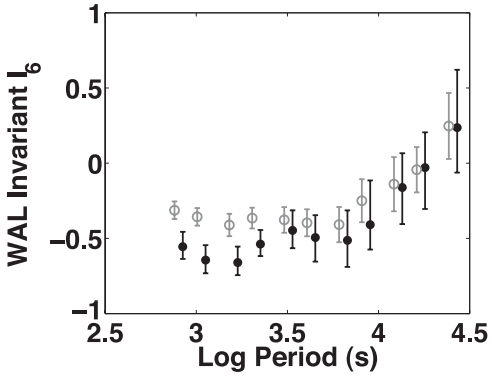

Figure 15. The WAL invariant I_6 for the New Zealand data set computed from section-by-section estimates of the magnetotelluric response. The solid circles are the median values with black error bars computed from the order statistics using (41), while the open circles are the mean values with grey jackknife confidence intervals. The error bars are $1 - \alpha/16$ values with $\alpha = 0.05$.

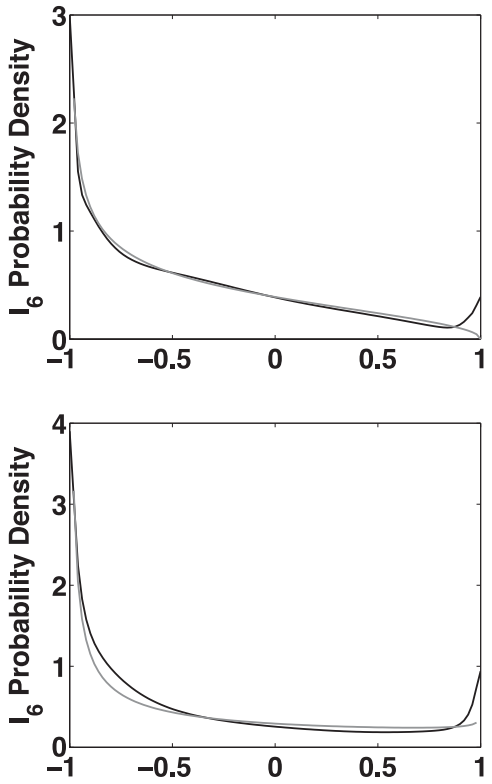

Figure 16. The empirical probability density function (black lines) for the WAL invariant I_6 with the New Zealand data at (top panel) 4500 s and (bottom panel) 1687 s. The kernel density estimator was Gaussian with a bandwidth of 0.4. The grey lines are the best-fit beta distribution using method of moments values for the parameters of (0.64, 1.49) at 4500 s and (0.37, 0.88) at 1687 s, respectively.

invariants. Use of the median with confidence intervals computed from the order statistics is shown to yield reliable, physically interpretable estimates for the rotational invariants. It is recommended that this approach be more widely adopted.

Because the support of all of the WAL invariants save $I_5 - I_6$ is either semi-infinite or infinite, they cannot be expressed as trigonometric functions, and hence cannot be represented using Mohr circles. This in no way invalidates the use of the WAL invariants for quantitative interpretation of magnetotelluric data and their dimensionality, but does invalidate their geometric representation.

ACKNOWLEDGEMENTS

This paper is the consequence of valuable discussions with Alan Jones, Gary Egbert and John Booker regarding rotational invariants, and was supported by NSF grant EAR1015185.

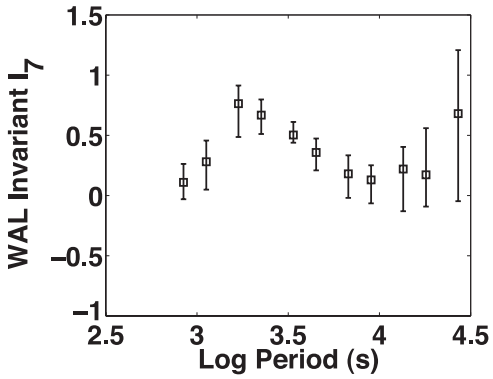

Figure 17. The median of the WAL invariant I_7 for the New Zealand data computed from section-by-section estimates of the magnetotelluric response together with confidence intervals estimated from the order statistics using (41). The error bars are $1 - \alpha/16$ values with $\alpha = 0.05$.

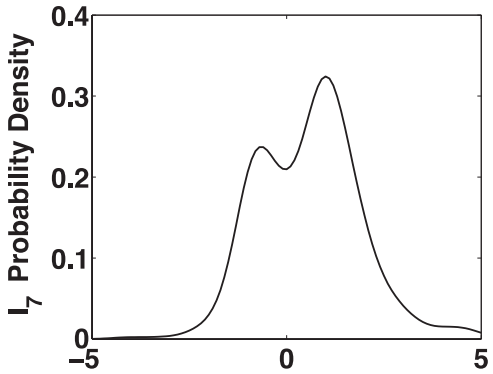

Figure 18. The empirical probability density function for the WAL invariant I_7 at a period of 1687 s computed using a kernel density estimator with a Gaussian smoother and a bandwidth of 0.15.

REFERENCES

Athreya, K.B., 1987. Bootstrap of the mean in the infinite variance case, *Ann. Stat.,* **15,** 724–731.

Bahr, K., 1988. Interpretation of the magnetotelluric impedance tensor—regional induction and local telluric distortion, *J. Geophys.,* **62,** 119–127.

Bahr, K., 1991. Geological noise in magnetotelluric data: a classification of distortion types, *Phys. Earth planet. Inter.,* **66,** 24–38.

Booker, J.R., 2013. The magnetotelluric phase tensor: a critical review, *Surv. Geophys.*, in press.

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,** 988–1006.

Cox, C., 1990. Fieller's theorem, the likelihood and the delta method, *Biometrics,* **46,** 709–718.

David, H.A., 1981. *Order Statistics,* 2nd edn, John Wiley and Sons, 360 pp.

De Groot, M.H. & Schervish, M.J., 2002. *Probability and Statistics,* 3rd edn, Addison-Wesley, 816 pp.

Fieller, E.C., 1940. The biological standardization of insulin, *J. R. Stat. Soc.,* **7**(Suppl.), 1–64.

Fieller, E.C., 1944. A fundamental formula in the statistics of biological assays and some applications, *Q. J. Pharm. Pharmacol.,* **17,** 117–123.

Fieller, E.C., 1954. Some problems in interval estimation, *J. R. Stat. Soc.,* **16,** 175–185.

Franz, V.H., 2007. Ratios: a short guide to confidence limits and proper use, arXiv:0710.2024v1, unpublished manuscript.

Gleser, L.J. & Hwang, J.T., 1987. The nonexistence of $100(1 - \alpha)\%$ confidence sets of finite expected diameter in errors-in-variables and related models, *Ann. Stat.,* **15,** 1351–1362.

Goodman, L.A., 1960. On the exact variance of products, *J. Am. Stat. Assoc.,* **55,** 708–713.

Herson, J., 1975. Fieller's theorem vs the delta methods for significance intervals for ratios, *J. Stat. Comp. Simul.,* **3,** 265–274.

Hirschberg, J. & Lye, J., 2010. A geometric comparison of the delta and Fieller confidence intervals, *Am. Stat.,* **64,** 234–241.

Hwang, J.T., 1995. Fieller's problems and resampling techniques, *Stat. Sinica,* **5,** 161–171.

Jones, A.G., 2012. Distortion of magnetotelluric data: its identification and removal, in *The Magnetotelluric Method: Theory and Practice,* pp. 219–302, eds Chave, A.D. & Jones, A.G., Cambridge Univ. Press.

Knight, K., 1989. On the bootstrap of the sample mean in the infinite variance case, *Ann. Stat.,* **17,** 1168–1175.

Lehmann, E.L. & Schaffer, J.P., 1988. Inverted distributions, *Am. Stat.,* **42,** 191–194.

Martí, A., Queralt, P. & Ledo, J., 2009. WALDIM: a code for the dimensionality analysis using the rotational invariants of the magnetotelluric tensor, *Comput. Geosci.,* **35,** 2295–2303.

Martí, A., Queralt, P., Jones, A.G. & Ledo, J., 2005. Improving Bahr's invariant parameters using the WAL approach, *Geophys. J. Int.,* **163,** 38–41.

Prácser, E. & Szarka, L., 1999. A correction to Bahr's "phase deviation" method for tensor decomposition, *Earth Planets Space,* **51,** 1019–1022.

Rice, S.O., 1945. Mathematical analysis of random noise, *Bell. Syst. Tech. J.,* **24,** 26–156.

Simpson, F. & Bahr, K., 2005. *Practical Magnetotellurics,* Cambridge Univ. Press, 254 pp.

Swift, C.M. Jr., 1967. A magnetotelluric investigation of an electrical conductivity anomaly in the southwestern United States, *PhD thesis,* Massachusetts Institute of Technology, 222 pp.

Szarka, L. & Menvielle, M., 1997. Analysis of rotational invariants of the magnetotelluric response tensor, *Geophys. J. Int.,* **129,** 133–142.

Thomson, D.J. & Chave, A.D., 1991. Jackknife error estimates for spectra, coherences, and transfer functions, in *Advances in Spectral Analysis and Array Processing,* Vol. **1,** pp. 58–113, ed Haykin, S., Prentice-Hall.

Von Luxburg, U. & Franz, V.H., 2009. A geometric approach to confidence sets for ratios: Fieller's theorem, generalizations and the bootstrap, *Stat. Sinica,* **29,** 1095–1117.

Weaver, J.T., Agarwal, A.K. & Lilley, F.E.M., 2000. Characterization of the magnetotelluric tensor in terms of its invariants, *Geophys. J. Int.,* **141,** 321–336.

Zerbe, G.O., 1978. On Fieller's theorem and the general linear model, *Am. Stat.,* **32,** 102–105.