Statistical Error of the Magnetotelluric Apparent Resistivity

Feb.99, Pamela Lezaeta. E-mail:plezaeta@whoi.edu

Summary

Estimation of the apparent resisistivity variance and expected value by assuming normal distributed tensor elements.

Expected value of ρ_a

The apparent resistivity (ρ_a) as function of period (T) and of an arbitrary tensor element Z is

$$\rho_a = 0.2 \cdot T(\operatorname{Re} Z^2 + \operatorname{Im} Z^2)$$

Its variance $Var(\rho_a)$ will depend on the real and imaginary impedance data errors $\Delta \operatorname{Re} Z$ and $\Delta \operatorname{Im} Z$, respectively, with known statistical distribution. By the majority of the data processing procedures, real and imaginary parts have equal errors, which implies that they are assumed uncorrelated since the variable is a complex number. This means that $\Delta \operatorname{Re} Z = \Delta \operatorname{Im} Z = \Delta |Z|$.

If the impedance tensor elements are random variables (r.v.'s), real and imaginary parts independently normal distributed ($N(a_{r,i}, \sigma)$ with mean values $u_r = \operatorname{Re} \hat{Z}$, $u_i = \operatorname{Im} \hat{Z}$ and standard deviation $\sigma = \Delta Z$), then the expected value of the square real and imaginary part of the tensor element, respectively is (see derivation in the appendix):

$$\begin{aligned} E(\operatorname{Re} Z^2) &= \operatorname{Re} \hat{Z}^2 + \sigma^2 \\ E(\operatorname{Im} Z^2) &= \operatorname{Im} \hat{Z}^2 + \sigma^2 \end{aligned} \qquad (2)$$

where $\operatorname{Re} \hat{Z}$ and $\operatorname{Im} \hat{Z}$ are the measured data assumed to be the respective expected values both with equally variance σ^2 (i.e. the square of the measured data error assumed to be the 68% confidence limit or its standard deviation). Thus we are able to determine the apparent resistivity expected value $E(\rho_a)$, which is subject to the ρ_a probability function distribution (p.f.d.). Following the property of independent r.v.'s summation we have that:

$$E(\rho_a) = E\left(0.2T\,|Z|^2\right) = 0.2T \times E\left(\operatorname{Re} Z^2 + \operatorname{Im} Z^2\right) = 0.2T \times [E\left(\operatorname{Re} Z^2\right) + E\left(\operatorname{Im} Z^2\right)]$$

and by using Eq.(2) the final expression for the expected value is obtained:

$$E(\rho_a) = 0.2T(2\Delta Z^2 + \operatorname{Re}\hat{Z}^2 + \operatorname{Im}\hat{Z}^2)$$
$$= 2\times 0.2T\Delta Z^2 + \hat{\rho}_a$$

This means that the measured (estimated) apparent resistivity $\hat{\rho}_a = 0.2T\times|\hat{Z}|^2$ is actually deviated from the expected value by the bias value $2\times 0.2T\Delta Z^2$. This fact relies on the p.d.f. characteristic of ρ_a.

Probability function distribution of ρ_a

Subject to the complex tensor element –assuming its real and imaginary parts as independent gaussian random variables (r.v.)– it follows that ρ_a has a non-central two degrees of freedom *chi* χ^2 p.f.d (see Appendix) with expected value

$$E(\rho_a) = 2\times 0.2T\cdot\Delta Z^2 + \hat{\rho}_a \quad . \qquad (3)$$

Its variance is obtained by using Eq.(2) and by knowing that the derivative of a normal distributed *N(u,σ)* characteristic function

$$\vartheta(t) = \exp(it\cdot u)\cdot\exp(-\frac{1}{2}\sigma^2 t^2)$$

of a r.v. X accomplishes the general relation for the expected value

$$E(X^n) = i^n\frac{\partial^n\vartheta(0)}{\partial t^n} \quad \Rightarrow \quad E(X^2) = i^2\frac{\partial\vartheta^2(0)}{\partial t^2} = u^2 + \sigma^2 \quad \text{and}$$

$$E(X^4) = 6u^2\sigma^2 + u^4 + 3\sigma^4$$

and that the variance of a r.v. $Y = X^2$ is per definition:

$$Var(Y) = E(Y^2) - \left(E(Y)\right)^2 \ .$$

Thus following the property of variance summation one obtains:

$$Var(\rho_a) = Var(cY) = c^2(Var\{\operatorname{Re} Z^2\} + Var\{\operatorname{Im} Z^2\})$$

$$= (0.2T)^2\times[E\{(\operatorname{Re} Z^2)^2\} + E\{(\operatorname{Im} Z^2)^2\} - (E\{\operatorname{Re} Z^2\})^2 - (E\{\operatorname{Re} Z^2\})^2]$$

$$\boxed{Var(\rho_a) = 4(0.2T)^2 \times (\sigma^4 + \sigma^2 \,|\, \hat{Z} \,|^2)} \qquad (4)$$

Now we can say that the real value ρ_a lays between the 86.5% confidence limit (property of the 2 degrees of freedom *chi*-χ^2 p.d.f.):

$$\left[E(\rho_a) - \sqrt{Var(\rho_a)}\right] < \rho_a < \left[E(\rho_a) + \sqrt{Var(\rho_a)}\right]$$

where the square root of the variance is the deviation of $E(\rho_a)$. By replacing $E(\rho_a)$ with the regular expression of Eq.(3) and $Var(\rho_a)$ with that of Eq.(4), we obtain thus through simple algebra the lower and upper deviation limits $\Delta\rho_-$ and $\Delta\rho_+$ (i.e., the error bars) of the measured data $\hat{\rho}_a = 0.2T \,|\, \hat{Z} \,|^2$, respectively:

$$E(\rho_a) - \sqrt{Var(\rho_a)} \Rightarrow \boxed{\Delta\rho_- = 2 \times 0.2T \cdot \sigma \cdot (\sigma - \sqrt{\sigma^2 + |\, \hat{Z} \,|^2})} \text{ where } \quad \Delta\rho_- < 0$$

$$E(\rho_a) + \sqrt{Var(\rho_a)} \Rightarrow \boxed{\Delta\rho_+ = 2 \times 0.2T \cdot \sigma \cdot (\sigma + \sqrt{\sigma^2 + |\, \hat{Z} \,|^2})} \text{ where } \quad \Delta\rho_+ > 0$$

which means that the real value ρ_a lays between the following confidence limit with respect to the measured data $\hat{\rho}_a$:

$$[\, \hat{\rho}_a - |\Delta\rho_-|\,] < \rho_a < [\, \hat{\rho}_a + \Delta\rho_+]$$

whose confidence limit is still the 86.5% because the region has been preserved. Thus this expression brings an asymmetrical error to the apparent resistivity, which relies on the asymmetry and non-centrality of its distribution function.

Fig. 1 shows the asymmetrical error of $\hat{\rho}_a$ for a normalised impedance as function of its relative error. It is seen that for impedances with over 50% percentage errors is the asymmetry more remarkable, abruptly increasing for larger errors. The asymmetry is of course directly proportional to the measured data deviation $2 \times 0.2T \cdot \Delta Z^2$ relative to the expected value .

Conclusion

We conclude that the estimation of a non-central *Chi*^{2} statistical distribution of ρ_a –for impedances with percentage relative errors not greater than approx 50% – results in negligible bias of the expected value giving very similar results to the symmetrical linear propagated error ([1]).

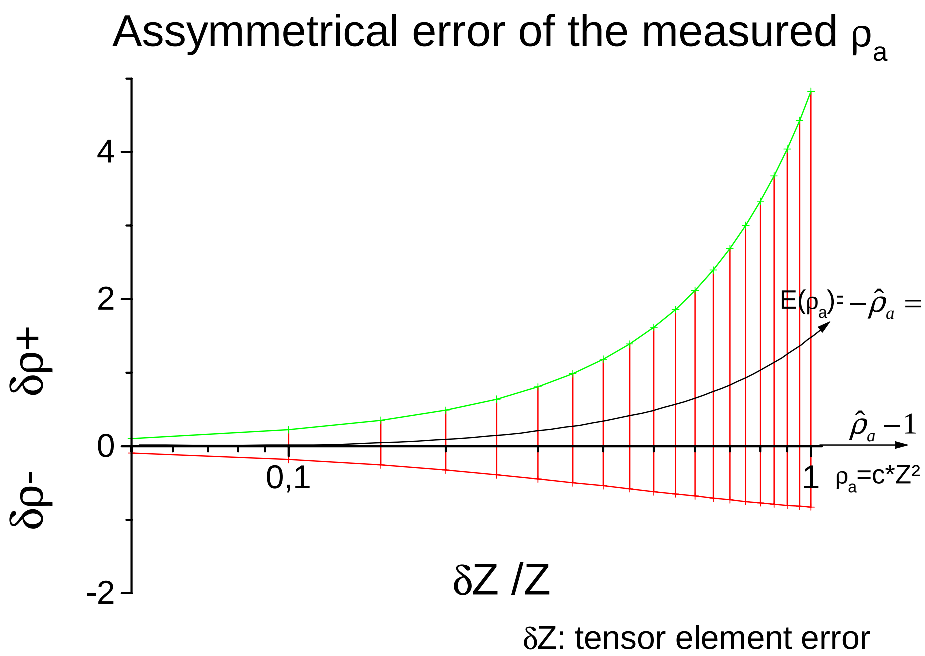

Fig.1: Graphical representation of the asymmetrical error ($\delta\rho^-$, $\delta\rho^+$) of the measured value $\hat{\rho}_a$ as function of the tensor element relative error $\frac{\Delta Z}{|\hat{Z}|}$ ($=\delta Z / Z$). Each value has been normalized by $0.2T\,|\hat{Z}|^2$, thus $\hat{\rho}_a = 1$, $E(\rho_a) \to E(\rho_a)/0.2T\,|\hat{Z}|^2$ and $\delta\rho_{+,-} = \Delta\rho_{+,-}/(0.2T\,|\hat{Z}|^2)$. In the graphic, $Z = |\hat{Z}|$.

Example on measured data.

The theory has been applied on data –processed with the Egbert code– of the campaign Bolivia'97, giving similar results to the usual linear propagated error calculation, provided of course that they are within the acceptable limits. For big errors it is observed that the expected apparent resistivity steeps up the curve, sometimes smoothing it at longer period bands (fig.2). This fact allows to re-considerate whether to use the measured data or the expected value when errors are over the acceptable limits (i.e., >50% percentage relative error). The expected phases are discussed in another document.

Appendix: The non-centrally X^2-parameter of the apparent resistivity.

For n independent r.v's U_i's each with standardized Normal d.f *N(0,1)*, the r.v. $U = \sum_{i=1}^{n}(U_i + \delta_i)^2$, where δ_i are constants, has a non-central chi-X^2 d.f of n degress of freedom and non-centrally parameter $\lambda = \sum_{i=1}^{n}(\delta_i^2)$. We denote this d.f. as $\chi^2(n,\lambda)$.

The expected value of the r.v. U of d.f. $\chi^2(n,\lambda)$ is:

$$E(U) = n + \lambda \quad , \qquad \text{(A1)}$$

and variance:

$$V(U) = 2(n + 2\lambda) \qquad \text{(A2)}$$

(e.g., Johnson & Kotz, 1970).

Then, for r.v's X_i's each with Normal d.f $N(a_i,\sigma)$, the r.v.

$$Y = \sum_{i=1}^{n} X_i^2 \qquad \text{(A3)}$$

has a non-central chi-X^2 d.f of n degress of freedom of the form:

$$\sigma^2\chi^2(n,\lambda) \qquad \text{(A4)}$$

and non-centrally parameter $\lambda = \sum_{i=1}^{n}(a_i^2/\sigma^2)$. (A5)

The expected value is: $E(Y) = \sigma^2(n + \sum_{i=1}^{n}\frac{a_i^2}{\sigma^2})$ (A6)

and the variance: $V(Y) = 2\sigma^4(n + 2\sum_{i=1}^{n}\frac{a_i^2}{\sigma^2})$. (A7)

In an earlier section the expected value of the apparent resistivity was obtained:

$$E(\rho_a) = 2\times 0.2T\cdot\sigma^2 + \hat{\rho}_a \qquad \text{(A8)}$$

where $\rho_a = 0.2T(\operatorname{Re}Z^2 + \operatorname{Im}Z^2)$ is a function of the normal distributed r.v.'s *ReZ* and *ImZ.* We assign the r.v. *ReZ* as X_1 and *ImZ* as X_2, with normal distributions $N(a_1,\sigma)$ and $N(a_2,\sigma)$, respectively. We express the r.v. ρ_a as a function of the r.v. Y of eq.(A3):

$$\rho_a = 0.2TY$$

[1] The apparent resisitivty linear propagated error has the formula: $2\times 0.2T\cdot\Delta Z\cdot|\hat{Z}|$

for $Y = \sum_{i=1}^{2} X_i^2$, which after eq.(A4) has a d.f. $\sigma^2 \chi^2(2,\lambda)$.

Then the expected value of ρ_a after eq.(A6) is:

$$E(0.2TY) = 0.2T\sigma^2(2 + \sum_{i=1}^{2} \frac{a_i^2}{\sigma^2}) ,$$

result wich is equivalent to the expression of eq.(A8) obtained directly from the property of expected values summation for independent r.v.'s. The same is observed for the variance deduced in section 4.1.1. with that from eq.(A7):

$$V(0.2TY) = (0.2T)^2 2\sigma^4 (2 + 2\sum_{i=1}^{2} \frac{a_i^2}{\sigma^2}) .$$

Then ρ_a has a d.f.: $0.2T\sigma^2\chi^2(2,\lambda)$

with the non-centrally parameter: $\lambda = \sum_{i=1}^{2} (a_i^2/\sigma^2) = \frac{\mathrm{Re}\,\hat{Z}^2 + \mathrm{Im}\,\hat{Z}^2}{\sigma^2}$,

where we replaced the expected (and mean) values $a_1 = \mathrm{Re}\,\hat{Z}$ and $a_2 = \mathrm{Im}\,\hat{Z}$ as expressed in an earlier section for clarity.

Proof:

By expressing each normal distributed r.v. X_i as a function of its standardised Normal distribution U_i :

$$X_i = a_i + \sigma U_i$$

we have for the r.v. Y the relation:

$$Y = \sum_{i=1}^{n} X_i^2 = \sigma^2 \sum_{i=1}^{n} (U_i + a_i/\sigma)^2 .$$

Thereof it is easely deduced the equivalence between the r.v. Y with $U = \sum_{i=1}^{n} (U_i + \delta_i)^2$

following the d.f. $\chi^2(n, \sum_{i=1}^{n} \delta_i^2)$. Then

$$\delta_i = \frac{a_i}{\sigma}$$

implying a non-centrally parameter: $\lambda = \sum_{i=1}^{n} (a_i^2/\sigma^2)$

and a non-central chi-X² d.f. of the form: $\sigma^2 \chi^2 (n, \lambda)$.

The expected value by considering eq.(A1) is :

$$E(Y) = E(\sigma^2 U) = \sigma^2 E(U)$$

$$= \sigma^2 (n + \lambda) = \sigma^2 (n + \sum_{i=1}^{n} \frac{a_i^2}{\sigma^2})$$

and the variance (eq.(A2)):

$$V(Y) = V(\sigma^2 U) = \sigma^4 V(U)$$

$$= 2\sigma^4 (n + 2\lambda) = 2\sigma^4 (n + 2\sum_{i=1}^{n} \frac{a_i^2}{\sigma^2})$$

From these relations is deduced the density function of the apparent resistivity.

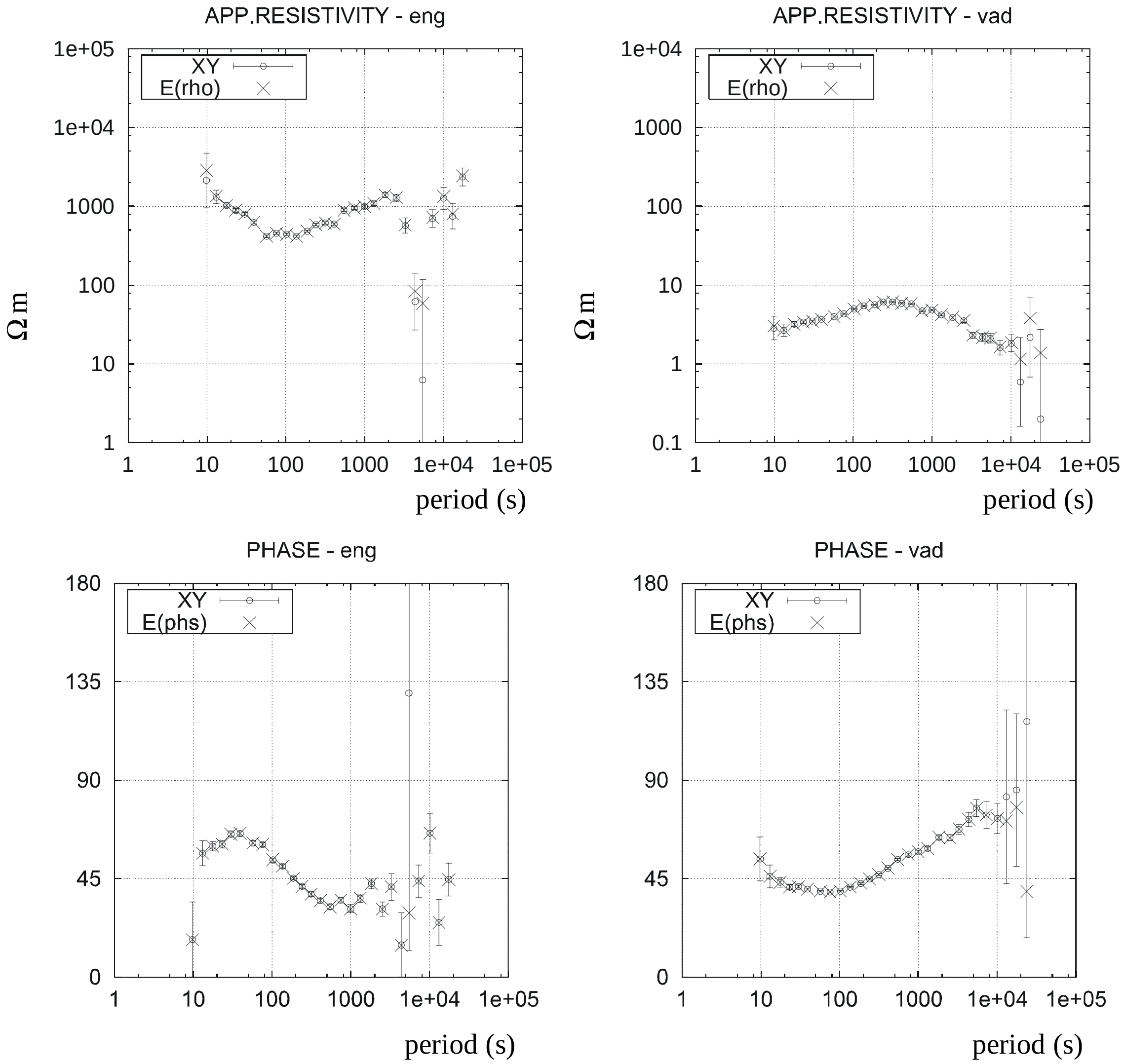

Fig.2 : *Above*: Example of apparent resistivity data (dots), expected values E(rho) and the asymmetrical app. resistivity error bars after the chi-X² statistic assumption (see text) for two site located near the Pacific Ocean and in the Bolivian Altiplano, respectively. Data were processed with the robust method of Egbert including also a remote reference station. *Bottom*: The corresponding data phases (dots) with their linear propagated errors and expected values E(phs) after polar transformation.