Comparison of the robust estimator with the least square estimator in magnetotelluric impedance estimation

Hang Chen

Abstract

See the abstract that has been submitted.

1.Introduction

Magnetotelluric (MT) is a geophysical exploration method to study the electric structure of the earth by using natural alternating electromagnetic fields (Chave and Jones, 2012). MT has been widely applied in tectonic studies (Jones and Hutton, 1979; Bedrosian and Feucht, 2014), geothermal exploration (Amatyakul, et al, 2015) and petroleum exploration (Hoversten et al, 1998). As an important step in MT data interpretation, time variation electromagnetic fields are typically Fourier transformed and linear regression algorithms are applied to obtain the estimation of MT response functions (Egbert and Booker 1986).

In the early time, the least square method (Sims et al.1971) was used to estimate the MT impedance tensor estimation, based on the assumption that the noise of measured MT data is Gaussian distributed with clean magnetic fields. Later on, the remote reference technique (Goubau et al. 1978; Gamble et al 1979) is proposed to handle the problems with noisy magnetic data. Least square method is easily failed due to the outliers in the electrical field, and robust estimators (Egbert and Booker 1986; Chave et al. 1987) are commonly applied, which can significantly reduce the effect of non-Gaussian noise in the electrical field data. Later on, the robust method became a standard method for MT data processing (Kelbert, 2020).

In this paper, I will use the Monte Carlo simulation in both synthetic and measured MT data to compare the robust and least square estimator. The synthetic MT data is generated by adding different distributed noise in electrical or magnetic field data and the measured MT data is collected in the Tibet plateau. The results from robust and least square estimator can give a lot of useful information to the practical applications of MT.

2.Methods

For one particular frequency band, we can stack the Fourier transformed electromagnetic fields into vector notations

$$\boldsymbol{E} = \begin{bmatrix} E_{x1} & E_{x2} & \cdots & E_{xN} \\ E_{y1} & E_{y2} & \cdots & E_{yN} \end{bmatrix} \boldsymbol{B} = \begin{bmatrix} B_{x1} & B_{x2} & \cdots & B_{xN} \\ B_{y1} & B_{y2} & \cdots & B_{yN} \end{bmatrix} \tag{1}$$

They are related via the impedance tensor

$$\boldsymbol{E} = \boldsymbol{Z}\boldsymbol{B}. \tag{2}$$

Equation (2) will be affected by the finite size of the sample and the presence of noise, and it becomes

$$\boldsymbol{E} = \boldsymbol{ZB} + \varepsilon. \tag{3}$$

For the least square method, the MT impedance tensor can be estimated by minimizing the equation (taking x-direction for example)

$$\sum_i \left|E_{xi} - (B_{xi}Z_{xx} + B_{yi}Z_{xy})\right|^2, \tag{4}$$

while the robust method (also called M-estimation) minimizes

$$\sum_i \rho\big(E_{xi} - (B_{xi}Z_{xx} + B_{yi}Z_{xy})\big). \tag{5}$$

where $\rho(r)$ is the Huber function and it is defined as

$$\rho(r) = \begin{cases} r^{\wedge}2/2 & r < r_0 \\ r_0|r| - r_0^2/2 & r \geq r_0 \end{cases}. \tag{6}$$

The r_0 usually equals to 1.5 （Huber, 2004). After minimizing the equation (4) and (5), the expression of least square impedance estimation is

$$\widehat{\boldsymbol{Z}} = [\boldsymbol{B}^*\boldsymbol{B}]^{-1}\boldsymbol{B}^*\boldsymbol{E} \tag{7}$$

where $\widehat{\boldsymbol{Z}}$ is the estimated impedance and * is the conjugate transpose and the expression of the robust estimator is

$$\widehat{\boldsymbol{Z}} = [\boldsymbol{B}^*\boldsymbol{WB}]^{-1}\boldsymbol{B}^*\boldsymbol{WE} \tag{8}$$

where $\boldsymbol{W}$ is a weight matrix, whose diagonal element is $\frac{\rho'(r)}{r}$.

The above is the basic theories about robust and least square estimator. To better examine their performance, we generate synthetic data with the estimated impedance from one site and compare solutions from the robust and stable MLE estimators. In these cases, the impedance is known. I use the measured horizontal magnetic fields as inputs to generate the electrical fields in the frequency domain, with random errors of different distribution added. The magnetic fields are treated differently (i.e., with and without noise). The error standard deviation used is 30 percent of the corresponding field amplitudes. I use three kinds of noise including Gaussian, Cauchy and Lévy noise. The probability density distribution (PDF) plot and Quantile-Quantile (QQ) plot of the three noise are shown in figure 1. It shows the skewness and/or kurtosis of the Cauchy and Lévy noise is different from the Gaussian noise. QQ plot shows that these two noises are long-tailed, which means they are very strong noise.

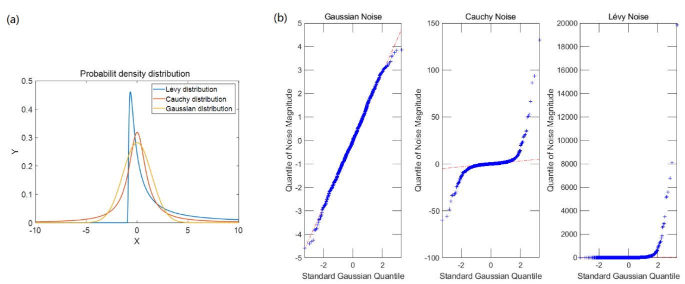

Figure 1. (a) The probability density distribution plot and (b) Quantile-Quantile (QQ) plot of

Gaussian, Cauchy and Lévy distributed noise.

3.Results

At first, I apply both the least square and the robust estimator to the synthetic data and compare their solution. As shown in figures 2 (a) and (b), both the methods can in general provide an accurate estimation of the apparent resistivity and phase (perfect match with the true model) when only Gaussian distributed noise in the electrical field. At several longest periods, slight bias from the true model is happened for solutions from both methods, probably due to the few data points at these periods. Relative to the apparent resistivity, the phase is more sensitive to the noise at the longest periods. However, when noise also exists in the magnetic field, as shown in figures (c) and (d), both methods have a slight down-biased estimation.

When stronger noise (Cauchy and Lévy distributed noise) exists in electrical field data like in figure 3 (a) and (b); figure 4 (a) and (b), robust estimator still performs well in most periods, while the least square estimator fails dramatically in these situations. It should be noted that with stronger noise in the electrical field, the robust estimator is still less effective to obtain unbiased estimation, especially in long periods. However, when the magnetic field is contaminated by these two extremely strong noises, both methods cannot give a precise estimation (e.g. figure 3 (c), (d), figure 4 (c) and (d)).

Finally, I apply the two estimators to a site from southeast Tibet Plateau to test their real applications in figure 5. On this site, the overall results differ slightly for the two estimators. However, at high frequency, the robust solution (for both apparent resistivity and phase) varies smoothly with frequencies, while, the least square solution is slightly more scattered.

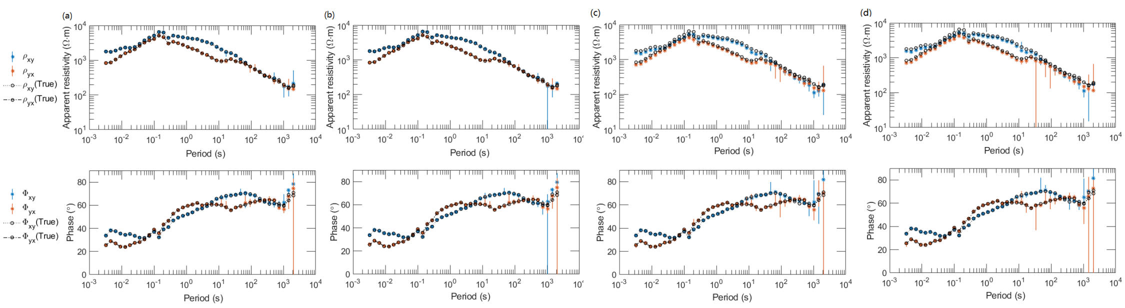

Figure 2. Apparent resistivity and phase comparison for (a) the least square method and (b) robust based on synthetic MT data with Gaussian distributed noise in the electric fields and comparison for (c) the least square method and (d) robust based on synthetic MT data with Gaussian distributed noise in the electric fields and magnetic fields.

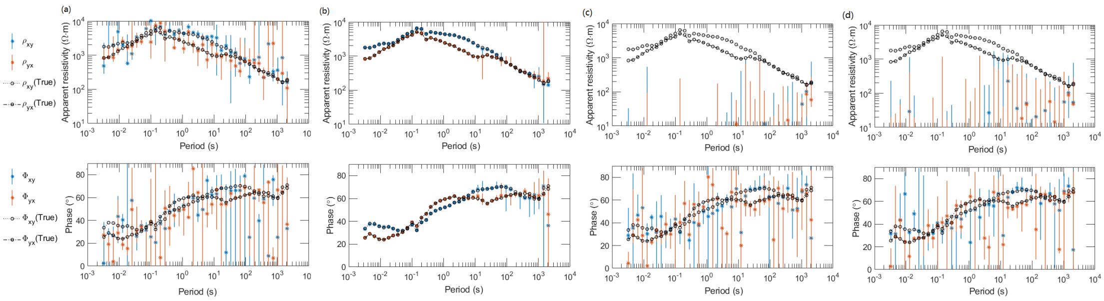

Figure 3. Apparent resistivity and phase comparison for (a) the least square method and (b) robust based on synthetic MT data with Cauchy distributed noise in the electric fields and comparison for (c) the least square method and (d) robust based on synthetic MT data with Gaussian distributed noise in the electric fields and Cauchy distributed noise in magnetic fields.

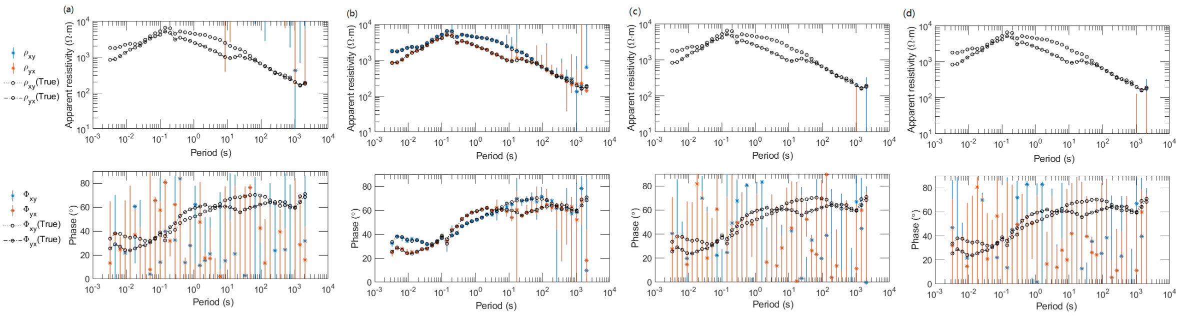

Figure 4. Apparent resistivity and phase comparison for (a) the least square method and (b) robust based on synthetic MT data with Lévy distributed noise in the electric fields and comparison for (c) the least square method and (d) robust based on synthetic MT data with Gaussian distributed noise in the electric fields and Lévy distributed noise in magnetic fields.

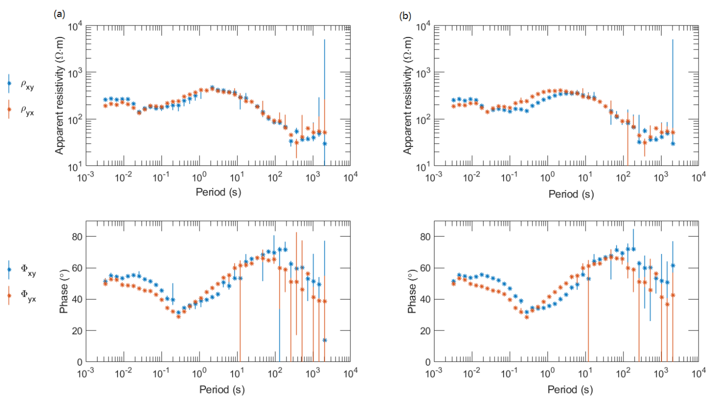

Figure 5. Apparent resistivity and phase estimates were obtained by applying (a)robust method, (b)stable MLE method to the data from the Tibet plateau.

4.Discussion and conclusion

In this paper, we make a complete comparison between the performances of the least square estimator and robust estimator. For this purpose, we generate synthetic MT data with desired errors of different distributions to model different real measurement errors. For the synthetic data with clean magnetic fields, the robust estimator will perform better than the least square estimator when extreme noise in the electrical field, and these two methods perform similarly with Gaussian noise in the electrical field. The existence of magnetic noise causes apparent resistivity bias for both methods. However, their phases are estimated accurately, indicating for data with Gaussian error contaminated input, the phase still can be recovered accurately. When input channels are contaminated by Cauchy or Lévy distributed errors, no meaningful interpretation can be obtained for the two methods. In measured MT data, these two estimators perform very similarly and the robust method is a little better than the least square method. This may be because the measured data are not affected by extreme noise.

Reference

Amatyakul, P., Rung-Arunwan, T. and Siripunvaraporn, W., 2015. A pilot magnetotelluric survey for geothermal exploration in Mae Chan region, northern Thailand. *Geothermics*, *55*, pp.31-38.

Bedrosian, P.A. and Feucht, D.W., 2014. Structure and tectonics of the northwestern United States from EarthScope USArray magnetotelluric data. *Earth and Planetary Science Letters*, *402*, pp.275-289.

Chave AD, Thomson DJ, Ander ME 1987 On the robust estimation of power spectra, coherences, and transfer functions. *J Geophys Res* 92:633–648

Egbert, G.D. and Booker, J.R., 1986. Robust estimation of geomagnetic transfer functions. *Geophysical Journal International*, 87(1), pp.173-194

Gamble, T.D., Goubau, W.M. and Clarke, J., 1979. Magnetotellurics with a remote magnetic reference. *Geophysics*, 44(1), pp.53-68.

Hoversten, G.M., Morrison, H.F. and Constable, S.C., 1998. Marine magnetotellurics for petroleum exploration, Part II: Numerical analysis of subsalt resolution. *Geophysics*, *63*(3), pp.826-840.

Huber, P.J., 2004. *Robust statistics* (Vol. 523). John Wiley & Sons

Jones, A.G. and Hutton, R., 1979. A multi-station magnetotelluric study in southern Scotland—II. Monte-Carlo inversion of the data and its geophysical and tectonic implications. *Geophysical Journal International*, *56*(2), pp.351-368.

Kelbert, A., 2020. EMTF XML: New data interchange format and conversion tools for electromagnetic transfer functions. *Geophysics*, *85*(1), pp.F1-F17.

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