ANNALI DI GEOFISICA, VOL. XXXIX, N. 1, January 1996

The iterative signal enhancing method for determining magnetotelluric impedance

Carla Braitenberg and Maria Zadro
Istituto di Geodesia e Geofisica, Università di Trieste, Italy

Abstract
The methodologies applied by our working group in Trieste for determining magnetotelluric impedance are presented, with special attention to the method used to evaluate the mean impedance over the entire observation interval using a robust procedure. The performance of our method is tested on an example data series, in order to compare it with other published methodologies. We make reference to a recorded time series of the EMSLAB traverse, which has been used to test eight different methods, the results of which have been collected in a published paper.

Key words *magnetotelluric impedance – EMSLAB Lincoln line*

1. Introduction

The evaluation of magnetotelluric sounding curves is divided into a sequence of successive analysis procedures, each one of which ought to be discussed separately and defined in its details, as the final result is strongly dependent on the single choices taken. The analysis starts with the preliminary evaluation of data quality over the entire recorded time period and the possible discarding of obviously disturbed data. The data are searched for spurious peaks and short interruptions, which must be corrected for. Furthermore the level of magnetic field activity must be assessed as a function of time, as it is well known that non-stationary phenomena like geomagnetic storms are generated by non-uniform sources, breaking the plane-wave hypothesis (*e.g.*, Egbert and Booker, 1986). Data decimation is part of the preliminaries, and is an essential tool for exploring the data in different frequency bands. The next step is the evaluation of the mean impedance over the data intervals selected, where crucial for the final result is the weighting criterion adopted. The last step is the decomposition of impedance, in order to obtain the sounding curves of apparent resistivities, phase and strike (Bahr, 1991; Groom and Bahr, 1992). In the following we focus our attention on the only problem of evaluation of mean impedance over the entire data set. The method of determining the magnetotelluric impedance developed in our working group has been explained previously in Braitenberg and Zadro (1990), showing the main features of the procedure only. Since then the methodology used has been refined, particularly the process of evaluating the mean impedance, that is as the weighted sum of the impedances obtained in subintervals of the observed sequence. The presentation of the method is not aimed to show that it is necessarily better than others (*e.g.*, Egbert and Booker, 1986), but that it is nonetheless acceptable in its application.

2. The method of calculating impedance

With $\boldsymbol{e}(f)$, $\boldsymbol{b}(f)$ the horizontal components of the telluric (T-) and the magnetic (B-) fields respectively, in frequency domain, the 2-D horizontal 2×2 impedance matrix $\boldsymbol{Z}(f)$ of the structure is defined by:

$$\boldsymbol{e}(f) = \boldsymbol{Z}(f)\,\boldsymbol{b}(f)$$

The computation of $\boldsymbol{Z}(f)$ is accomplished by an iterative signal to noise ratio enhancing method, in principle related to the technique developed by Kao and Rankin (1977). It also shares a common feature with the method developed by Egbert and Booker (1986), who introduce the «modified observation», where the observations are pulled toward their predicted values for large residuals. The impedance is supposed to vary slowly with frequency and is determined for M discrete frequencies f_k, $k = 1, M$, the impedance being constant over the frequency band defined by the two border frequencies f_{k1} and f_{k2}. In the n-th step of iteration the impedance at the frequency f_k is calculated from the relation:

$$|\,\boldsymbol{e}_{n-1}(f) - \boldsymbol{Z}_n(f_k)\boldsymbol{b}_{n-1}(f)\,|^2 = \min \quad f_{k1} \le f < f_{k2}$$

and the predicted T-field is then defined by:

$$\boldsymbol{e}_n(f) = q\boldsymbol{Z}_n(f_k)\boldsymbol{b}_{n-1}(f) + (1-q)\boldsymbol{e}_{n-1}(f) \quad f_{k1} \le f < f_{k2}$$

where q is the signal enhancing parameter.

The admittance $\boldsymbol{Q}_n(f_k)$ is then obtained from:

$$|\,\boldsymbol{b}_{n-1}(f) - \boldsymbol{Q}_n(f_k)\boldsymbol{e}_n(f)\,|^2 = \min \quad f_{k1} \le f < f_{k2}$$

and the predicted B-field is then defined by:

$$\boldsymbol{b}_n(f) = q\boldsymbol{Q}_n(f_k)\boldsymbol{e}_n(f) + (1-q)\boldsymbol{b}_{n-1}(f) \quad f_{k1} \le f < f_{k2}$$

The number L of iterations is chosen either fixed for all cases, or repeated until the difference between succeeding impedances is less than a fixed minimum value. Preference for one or the other method depends on the quality and initial signal to noise ratio of the data. The method is able to retrieve the correct impedance in the case of uncorrelated noise in the T- and B-fields, which causes a down weighting of the impedance components (Kao and Rankin, 1977). We operate the above procedure in time domain, applying it on band filtered data.

3. Evaluation of the average impedance

Due to the nonstationarity of the data, interruptions and the possible variation of the source fields, it is convenient to divide the complete data sequence into N valid subintervals. The estimate of the impedance representing the entire data set is then obtained by averaging the impedances calculated in each subinterval. We use a robust averaging scheme, defined by Claerbout and Muir (1973) as a method «relatively insensitive to a moderate amount of nonstationarity and outliers», and taken up for MT studies by various authors (*e.g.*, Egbert and Booker, 1986; Chave and Thomson, 1989). We have adopted a 2-step data adaptive scheme which combines down weighting of samples with large residuals from the average impedance, enhancing those values obtained on data intervals with high signal to noise ratio (SNR) in the T-field. In the first step the weighted impedance means are obtained, the weights being equal to the SNR of the T-field for the i-th subinterval at frequency f_k, defined by:

$$\mathrm{SNR}^i(f_k) = \int_{fk1}^{fk2} \frac{|\boldsymbol{Z}^i(f_k)\boldsymbol{b}^i(f)|^2}{|\boldsymbol{e}^i(f) - \boldsymbol{Z}^i(f_k)\boldsymbol{b}^i(f)|^2}\,df$$

The weighted mean impedance is calculated as defined in the following:

$$\overline{\boldsymbol{Z}}(f_k) = \sum_{i=1}^{N} w^i(f_k)\boldsymbol{Z}^i(f_k),$$

where the weights are defined as follows; those sequences for which the SNR is less than unity are eliminated:

$$w^i(f_k) = \begin{cases} \mathrm{SNR}^i(f_k) & \text{for} \quad \mathrm{SNR}^i(f_k) \ge 1 \\ 0 & \text{for} \quad \mathrm{SNR}^i(f_k) \langle 1 \end{cases}$$

In the second step the averaging is repeated, down weighting those cases for which the deviation of impedance from the mean impedance is greater than the median med(dZ) of the distribution, following the Huber weight function (Chave and Thomson, 1989):

$$w^i(f_k) = $$

$$= \begin{cases} \mathrm{SNR}^i(f_k) & \text{for } dZ^i(f_k) \le \mathrm{med}\,(dZ(f_k)) \\ \mathrm{SNR}^i(f_k)\, \mathrm{med}\,(dZ(f_k)) / dZ^i(f_k) & \text{for } dZ^i(f_k) > \mathrm{med}\,(dZ(f_k)) \end{cases}$$

The error estimate of impedance is found calculating the jackknife variance (Chave and Thomson, 1989); application of the jackknife is preferable with respect to a distribution based variance estimate, due to the difficulties in assigning a valid distribution of the impedances obtained for all subintervals.

4. Resolved frequencies

Considering the slow frequency dependence of impedance, we have chosen to evaluate impedance for $M = 6$ frequencies per decade. The frequencies f_k, $k = 1$, M are set equidistantly on a logarithmic scale. This implies a broadening of the bandwidth which is proportional to the frequency f_k. Successive frequencies are in a proportion equal to $f_{k+1} = 10^{1/M} f_k$, and the bandwidth b_k is defined by $b_k = 2df_k$, with $d = (10^{1/M}-1)/(10^{1/M}+1)$.

The length of data sequences is fixed to 720 samples, varying the sampling in order to analyse the different frequency ranges of the magnetic field as the harmonics of the Solar quiet component, the variations and the pulsations. Table I lists the sampling (dt), the lengths (T) of the analyzed time periods, and the inverse of the lower and upper frequencies defining each of the three frequency ranges used ($P1$ and $P2$, respectively).

5. Sounding curves for the EMSLAB01 site

The records of one site of the EMSLAB – Juan de Fuca project were used to compare the relative efficacies of different MT-analysis codes given the same data. The paper of Jones *et al.* (1989) where the results of this comparison have been published, gives an excellent opportunity to compare our MT analysis code with the other ones considered. The same test was also taken up in the recent paper of Spagnolini (1994). The data of the site 1 of the EMSLAB Lincoln line long-period land MT transect were kindly made available by Alan Jones for this purpose. The site was 10.6 km from the ocean and located on the Coast Range sediments (Wannamaker *et al.*, 1989). The entire data set consists of 10 weeks of recording from July 18 to September 23, 1985, with a sampling interval 20 s. The test of the different methods is applied on a 5 day long time segment, characterized by high B-field activity. In the paper of Jones *et al.* (1989) 8 different methods ranging from conventional spectral single station analysis to remote reference robust processing were tested. We compare our method with the best performance single station scheme cited in Jones *et al.* (1989) and with the recently published method of Spagnolini (1994). Table II lists a brief description of the methods. Figure 1 reports the apparent resistivities with their associated standard errors. For comparison, the resistivity curves given by Jones *et al.* (1989) for the robust remote reference method of Chave and Thomson (1989) are superposed, where the calculation was done using the entire data set of 10 weeks of data.

We evaluated the response curves in the range of 120 s to 6700 s, using the sampling

Table I. Sampling dt, window length T and upper $P1$ and lower $P2$ limit of the periods analyzed.

dt	T	$P1$	$P2$
15 min	6 days	4 h	2 h
2 min	24 h	111 min	10 min
20 s	4 h	20 min	2 min

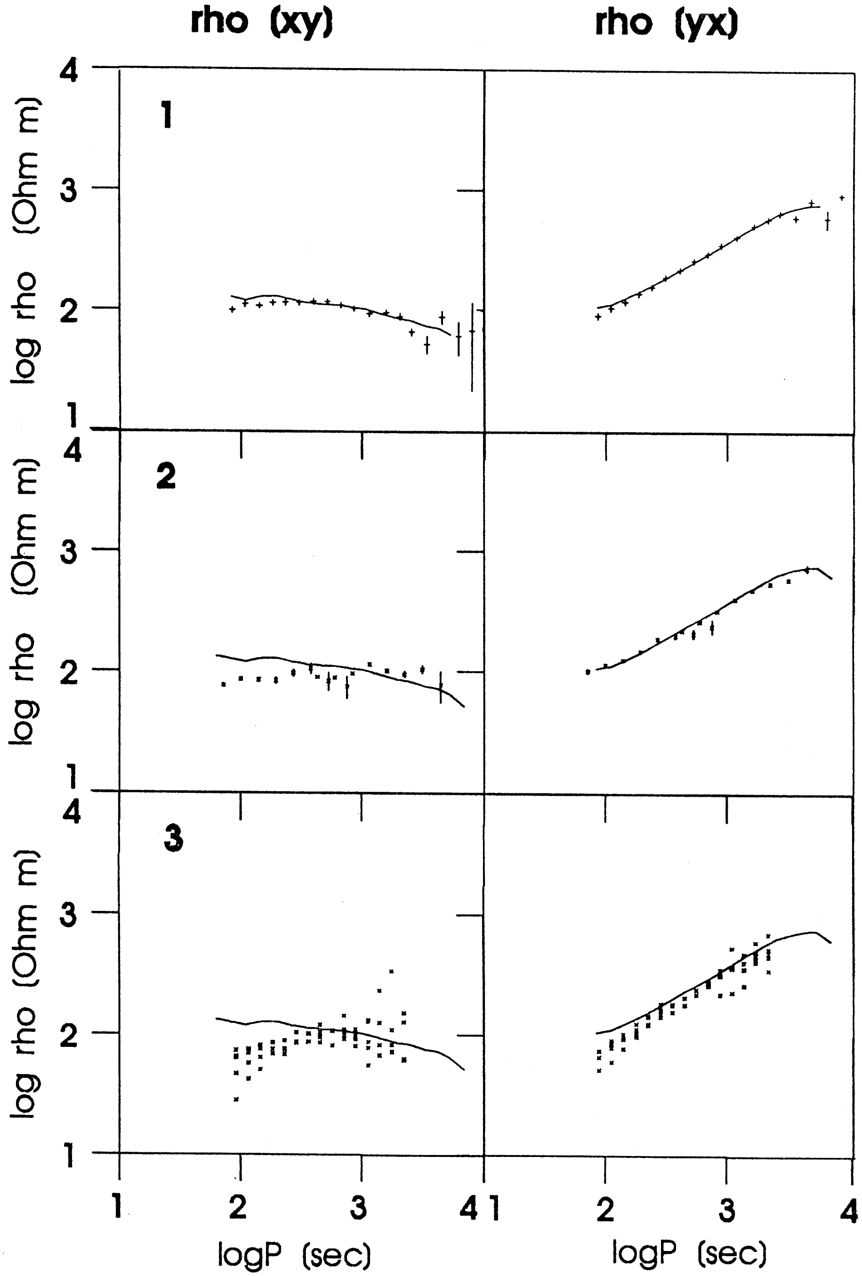

Fig. 1. Analyses of 5 day EMSLAB data during geomagnetic active interval. Apparent resistivity curves obtained with 3 methods: 1 = robust processing, Egbert and Booker (1986) (note that the *ρxy* curve has been shifted downward by one decade); 2 = present method; 3 = time domain processing after Spagnolini (1994). Results are compared to the analysis of 10 weeks of data obtained from remote reference robust analysis (continuous curve).

Table II. Brief summary of the methods used in the test of the performance of different impedance calculation schemes.

Method	Description	Author
1	Single station robust processing. Outliers cleaning with median and median absolute filter. Cascade decimation. Data selection based on minimum and maximum acceptable *B*-field energy. Band and section averaging of spectra using regression *M*-estimate.	Egbert and Booker (1986)
2	Signal to noise enhancing iterative processing. Fixed time window of 720 points with different sampling rates. Data cleaned for spikes prior to decimation. Signal to noise enhancing iteration. Weighted impedance averaging: weights given by SNR of *E*-field; impedances with large residuals from average are downweighted. Jackknife errors.	Present paper
3	Time domain processing. Iterative technique for impulse response estimation of impedance tensor based on least mean square. Baye's criterion for data classification, into noisy and adequate data. Single impedances obtained on 5 overlapping data windows are presented.	Spagnolini (1994)

and window lengths given in table I. Prior to decimation, the data were despiked and smoothed, using a moving window on which the central value is reduced if it exceeds the standard deviation from the median. This operation was necessary in geomagnetic active periods, where sudden storms otherwise appear as spikes in the decimated sequence. Data efficiency was almost complete for the long period curve estimation, whereas data selection was applied for the short period part of the curve, excluding the low *B*-field energy sequences from the analysis. This was accomplished introducing a *B*-field index based on the mean *B*-field energy distribution over the entire 10 weeks data set. The selection resulted in a 70% data efficiency. Our response curves (method 2) are smooth over the entire frequency range. The estimates obtained from the 20 s sampling (4 h intervals) and those obtained after decimation to the sampling of 2 min (24 h intervals) agree well in the frequency range of superposition, which comprises periods from 540 s to 1100 s. Our method reliably retrieves the resistivity curves with small random errors and low bias from the 5 days of data, competing with the results of the robust processing of Egbert and Booker (method 1). Only in the short period part of the ρ_{xy} curve do we obtain down biased estimates of resistivity. This was observed in several of the other methods, as also in Spagnolini (1994). We have undertaken a test on the stationarity of the response curves as a function of an index of *B*-field energy integrated over the entire bandwidth from 100-10000 s, as introduced in Egbert and Booker (1986). The ρ_{yx} curve is found hardly to vary, whereas the ρ_{xy} has been observed to have different bias depending on geomagnetic activity index. We thus maintain that the observed bias reflects the particular way of solving the data selection, rather than the methodology of impedance calculation. Phases are shown in fig. 2 for methods 2 and 3 only, as Jones *et al.* (1989) do not show the performance of the different methods on phase determination for the 5 day analysis. Our method retrieves phases well for the short period range of the φ_{xy} curve, giving values that agree with the all data curve. In the long period range and for the φ_{yx} curve, our method and method 3 agree in a substantiable bias with respect to the all data values. It is possible though, that the 5 days of data are not sufficient to retrieve the phase curves, which are even more sensitive than resistivity to noise in the data.

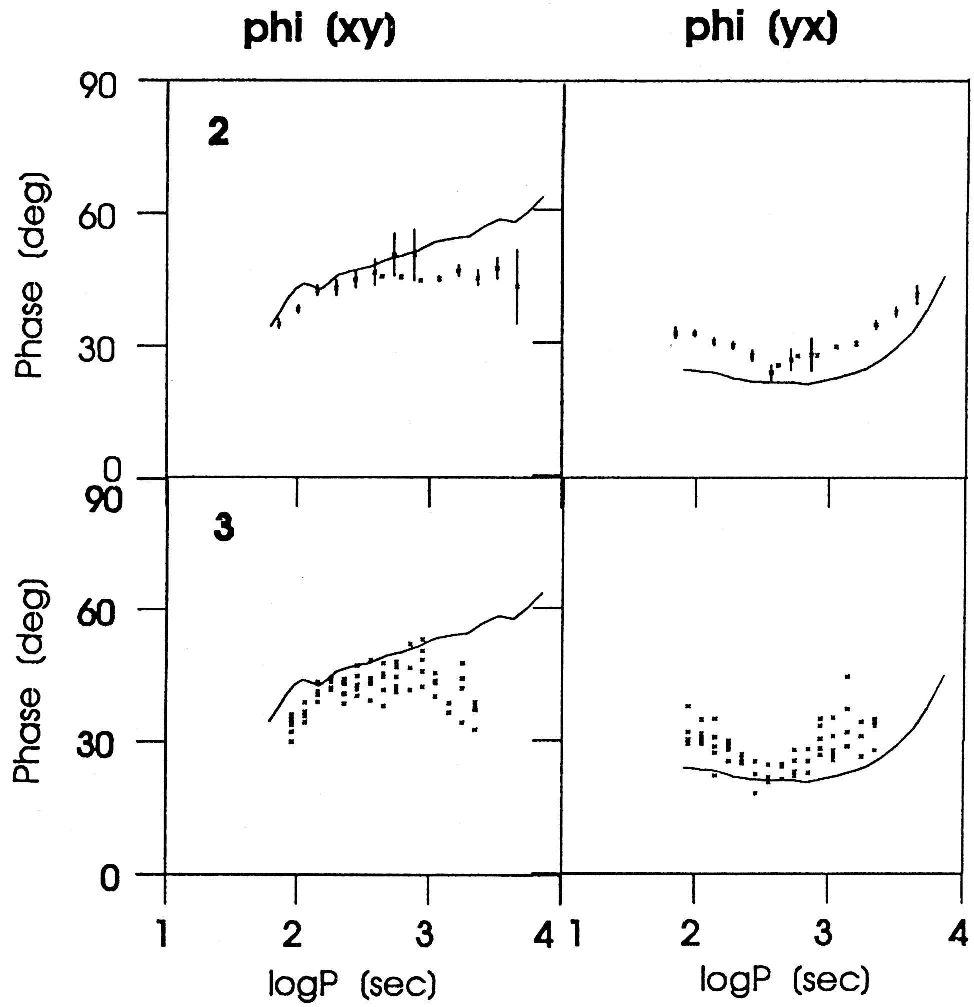

Fig. 2. Analyses of 5 day EMSLAB data during geomagnetic active interval. Phase curves obtained with the methods: 2 = present method; 3 = time domain processing after Spagnolini (1994). Results are compared to the analysis of 10 weeks of data obtained from remote reference robust analysis (continuous curve).

6. Conclusions

The present note documents the methodology we use for calculating impedance, as a reference for any future MT-campaigns. The opportunity given by Alan Jones of distributing a MT-data series free to use to test the calculation schemes adopted is excellent to present, test, and also refine the proper methods. In this respect all preliminaries such as data selection, data cleaning and data decimation are included, as the EMSLAB data are not free of disturbances.

Compared to other methods presented in Jones *et al.* (1989) our method meets the standards set by the robust methodologies of Egbert and Booker (1986), which seemed to have performed best among the non-remote reference methods. Our method has one advantage though, which is intrinsic to the fact that averaging is done on impedances obtained on single subintervals: this controls the time variability of impedance, which may be done as a function of the *B*-field energy distribution or known industrial disturbance sources (in Italy the railways).

Acknowledgements

The research was supported partially by MURST40% R94 Cod.1823 and 40% R93 Cod.1506 and CNR 94.00143.05.

REFERENCES

BAHR, K. (1991): Geological noise in magnetotelluric data: a classification of distortion types, *Phys. Earth Planet. Inter.*, **66**, 24-38.

BRAITENBERG, C. and M. ZADRO (1990): The magnetotelluric campaign in eastern Alps, NE-Italy: regional and local 2-D responses, *Boll. Geofis. Teor. Appl.*, **32**, 141-156.

CHAVE, A.D. and D.J. THOMSON (1989): Some comments on magnetotelluric response function estimation, *J. Geophys. Res.*, **94**, 14215-14225.

CLAERBOUT, J.F. and F. MUIR (1973): Robust modelling with erratic data, *Geophysics*, **38**, 826-844.

EGBERT, G.D. and J.R. BOOKER (1986): Robust estimation of geomagnetic transfer function, *Geophys. J. R. Astron. Soc.*, **87**, 173-194.

GROOM, R.W. and K. BAHR (1992): Corrections for near surface effects: decomposition of the magnetotelluric impedance tensor and scaling corrections for regional resistivities: a tutorial, *Surv. Geophys.*, **13**, 341-379.

JONES, A.G., A.D. CHAVE, G. EGBERT, D. AULD and K. BAHR (1989): A comparison of techniques for magnetotelluric response function estimation, *J. Geophys. Res.*, **94**, 14201-14213.

KAO, D.W. and D. RANKIN (1977): Enhancement of signal to noise ratio in magnetotelluric data, *Geophysics*, **42**, 103-110.

SPAGNOLINI, U. (1994): Time domain estimation of MT impedance tensor, *Geophysics*, **59**, 712-721.

WANNAMAKER, P.E., J.R. BOOKER, J.H. FILLOUX, A. JONES, G.R. JIRACEK, A.D. CHAVE, P. TARITS, H.S. WAFF, G.E. EGBERT, C.T. YOUNG, J.A. STODT, M.G. MARTINEZ, L.K. LAW, T. YUKUTAKE, J.S. SEGAWA, A. WHITE and A.W. GREEN JR. (1989): Magnetotelluric observations across the Juan de Fuca subduction system in the EMSLAB project, *J. Geophys. Res.*, **94**, 14111-14125.