

# Specific signal processing for high frequency range in Audio-Magnetotelluric

Author: Albert Malet Planas

Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain\*.

**Abstract:** The usual methodology used in magnetotelluric signal processing (MT) has been studied and applied to the case of audio-magnetotellurics (AMT). The results obtained have proved not to be good enough in order to obtain a good subsurface resistivity model. Therefore, some changes have been made to the methodology of signal processing so that it suits audio-magnetotellurics. These changes have applied to the selection of time windows and Fourier transform, to the frequency target list (its own code in Matlab), to Parzen radius and to the method of optimal transfer function estimation.

## I. INTRODUCTION

Magnetotellurics is a passive geophysical method that uses the phenomenon of electromagnetic induction to obtain information on the distribution of electrical resistivity in the subsurface. This induction takes place when an electromagnetic field of the type called *primary* reaches the Earth's surface, which behaves as a conductor, causing a distribution of charges and currents, which will in turn result in a *secondary field*. A measuring device on the surface records the electromagnetic components resulting from the superposition of primary and secondary fields.

Magnetotelluric transfer functions (MTFs) - which typify electric resistivity in the deep subsurface - can be obtained from these field components.

These primary waves that enter the subsurface undergo exponential decay which varies depending on soil conductivity and the wave period (or frequency). This decay results in a penetration depth ( $\delta$ ) (skin depth), which is defined as the depth at which the amplitude of the incident wave decreases in an e factor:

$$\delta(T) = \left( \frac{T}{\pi\mu\sigma} \right)^{1/2} \quad (1)$$

In comparison with other subsurface exploration methods, the main advantage of this method lies in the wide range of depths it enables to study thanks to the existing diversity of frequencies in electromagnetic fields, which enter the subsurface reaching varying depths according to equation (1). Primary fields, which constantly reach the Earth's surface, have got a great diversity of origins and can be either natural or artificial. Naturally originating low frequencies (<10Hz) are mainly related to solar wind while high frequency ones (>10Hz) to electrical storms. Between these two zones there is a zone with very low natural signal amplitude which is known as dead band. Therefore, it will only be possible to study this depth by means of artificial frequencies [1]. Artificial fields have got a wide range of frequency diversity, such as 50Hz and its harmonics coming from high-voltage power lines.

Although the term magnetotellurics refers to the method, in terms of frequency range, it refers to the study of the frequency range comprising  $10^2$  Hz and  $10^{-5}$  Hz. Thus being audio-magnetotellurics (AMT) the study between  $10^5$  Hz and 10Hz and radio-magnetotellurics (RMT) from  $10^6$ Hz to  $10^4$  Hz. Generally, the boundaries between the zones under study

are not well defined in that there is slight overlapping and/ or variation depending on the sources.

The commercial software developed has been devised especially with view to MT range, due to the interest in the study of deep geological strata (mining, oil drilling, geothermal energy...). Consequently the established methodology to process data is devised for MT, therefore being not appropriate for processing signals in the AMT range.

This TFG (Bachelor's Degree Final Paper) is aimed at obtaining a data processing methodology appropriate for AMT and is organized as follows: Section II deals with MT signal processing, accounting for the parameters that are unsuitable for AMT and proposing an optimization. In section III the results obtained using the modified approach explained in the previous section and those obtained using a non-modified approach are listed, compared and discussed. Finally, IV presents the conclusions arrived at in this TFG.

## II. SPECIFIC SIGNAL PROCESSING FUNDAMENTALS

Signal statistical processing includes all the steps implemented in order to obtain MTFs [2]. There are multiple transfer functions, the most widely used of which is known as impedance tensor ( $Z(f)$ ), which relates horizontal components of electric and magnetic fields to each of the frequencies according to the formula below:

$$\begin{bmatrix} E_x(\omega) \\ E_y(\omega) \end{bmatrix} = \begin{bmatrix} Z_{xx}(\omega) & Z_{xy}(\omega) \\ Z_{yx}(\omega) & Z_{yy}(\omega) \end{bmatrix} \begin{bmatrix} H_x(\omega) \\ H_y(\omega) \end{bmatrix} \quad (2)$$

From this tensor apparent resistivity ( $\rho_a$ ) and the phase ( $\varphi$ ) can be obtained:

$$\rho_{a_{ij}}(\omega) = \frac{1}{2\pi\mu\omega} |Z_{ij}(\omega)| \quad (3)$$

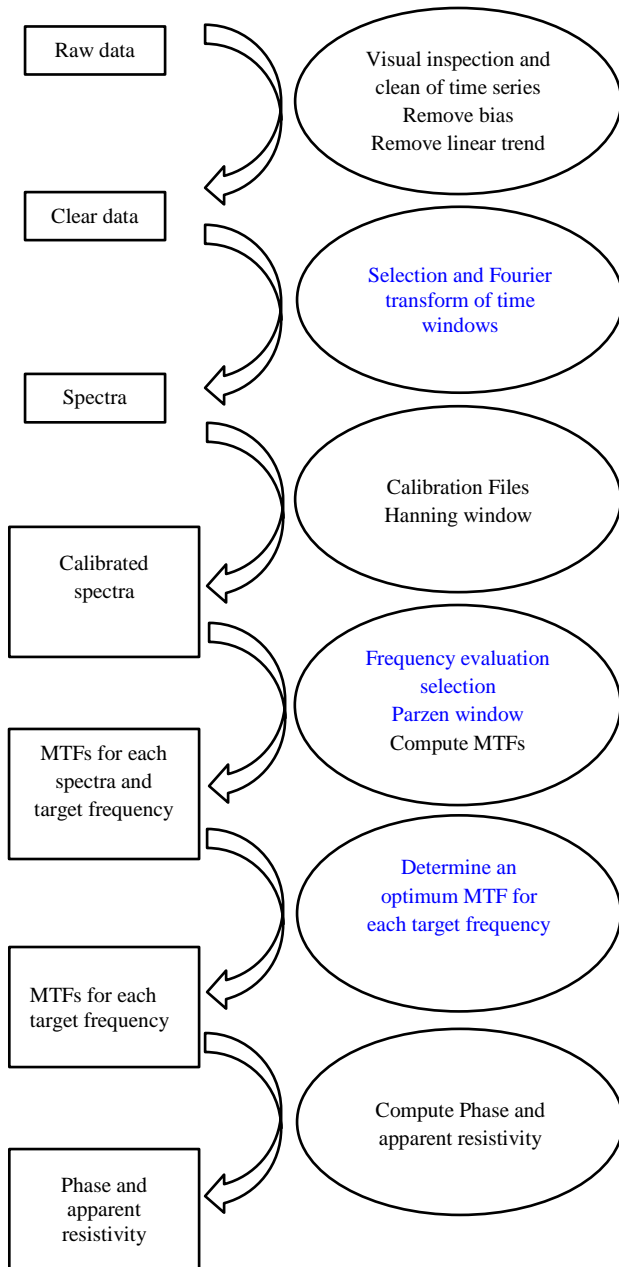
$$\varphi_{a_{ij}}(\omega) = \arctan \left( \frac{\text{Im } Z_{ij}(\omega)}{\text{Re } Z_{ij}(\omega)} \right) \quad (4)$$

where  $ij = xx, xy, yx, yy$ .

This processing is summarized in the flowchart below, where the steps directly related to the range of frequencies applied are marked in blue, and on which my work has been focused in order to obtain an optimization that can work in AMT.

---

\* amaletpl7@alumnes.ub.edu



**FIG. 1:** Flowchart summarising the steps involved in MT processing data.

Consequently, below is an analysis of these parameters, which mean an improvement in processing time series collected at high frequencies (AMT).

### A. SELECTION AND FOURIER TRANSFORM OF TIME WINDOWS

The first step in time signal processing consists of dividing the time series in a specific number of windows so that all of them have the same amount of data. This amount of data in each window is known as half of FFT longitude (LFFT/2) and given a sampling frequency ( $F_s$ ) and collecting time ( $t$ ), it defines the number of windows ( $M$ ) in which the series is divided as:

$$M = \frac{F_s}{LFFT} \cdot t \quad (5)$$

Then the Fourier transform is calculated for each of these windows, obtaining  $M$  windows for each electric and magnetic field. As the length of the FFT is the amount of data used in the transform, this is a parameter closely related to statistical error and the resolution obtained from the process. So that the lower the LFFT, the higher the number of windows, but the less amount of data in each window, which will lead to less error. However, that will also cause a decrease in resolution. Consequently, the higher the LFFT, the higher the resolution but also the error increases.

Therefore, it is a matter of finding an LFFT that induces minimum error but keeps a high enough resolution in order to distinguish the different subsurface resistivity structures.

Although this is a parameter which implicitly depends on the data to process, for significantly good MT data (without excessive noise), it is usually appropriate to use a value of 4096Hz [3]. This value is excessively high for as high a sampling frequency as used in AMT due to the fact that being at such high frequencies the data obtained are more contaminated by noise, therefore it is better to use a lower LFFT.

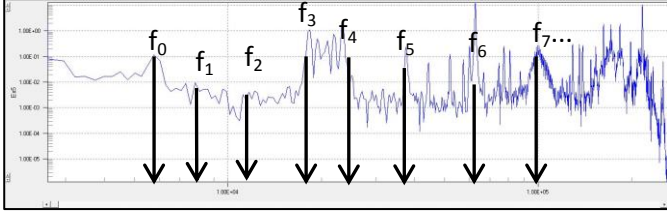
### B. FREQUENCY EVALUATION SELECTION

Subsequently the selected target frequencies will be the values used in calculating transfer functions, considering the distribution of the surrounding frequencies through statistical estimation, which will be explained in section II.C. Therefore being important for these target frequencies to be as precise values as possible (minimum noise in the measurement).

Frequency lists used in MT processing are created using three parameters: initial frequency, final frequency and total number of equidistant frequencies distributed in the logarithmic space [3], [4]. In order to improve the results obtained in AMT, these lists of frequencies will not be used. Instead, a specific selection of these frequencies will be used.

Our proposal in this section is to select peak frequencies (relative maximums) in the spectral range, as it is in those peaks where the signal-noise ratio will be higher. Nevertheless, outliers must be avoided, therefore being very important to conduct a previous visual examination of the data. Selecting those peaks manually from all the fields is a costly, complex and time-consuming process due to the length of the time series (a lot of measurements per second). Moreover, it is also imprecise due to spectral density. In order to show how complex the manual process can be, the following equation is presented (5). In this case, the measurement with a sampling frequency of 524 KHz has duration of 10 seconds, so dividing the time series in FFT lengths of 4096, a number of windows that equals 1279 is obtained.

Therefore, it is clear that this process has to be machined by means of programming. The code I have developed in Matlab analyses the spectrum of electric and magnetic field components in order to determine those frequencies presenting amplitude peaks and to produce a file in text format (.txt) listing them. Although it has been designed to optimize AMT processing, the program is broad enough to allow working in any frequency range without the need to modify any parameters.



**FIG. 2:** Image calibrated spectrum of the field Ex where some relative maximum ( $mV/\sqrt{Hz}$ ) are indicated by black arrows. These frequency values (Hz) could comprise the list of frequencies.

Some considerations that have been taken into account when designing this code are listed below:

i. Time series reading

The reading code of the file coming from instrumentation has been supplied by the workgroup EXES from the Geodynamics and Geophysics Department of UB's Faculty of Geology.

ii. Equidistant frequencies in logarithmic space

Equidistance in logarithmic space proves appropriate in order to obtain homogeneous distribution of the values all over the space, which is necessary in order to obtain a good map of subsurface resistivity. That is, obtaining smooth resistivity curves that enable subsequent data inversion.

A part from the visual effect, which is very important in representing data, there is another reason for selecting equidistant frequencies. Equidistance in logarithmic space makes relative penetration error constant [2]:

$$\frac{dp}{p} = \frac{1}{2} \frac{dt}{t} \quad (6)$$

iii. Number of frequencies per decade

Weidelt's dispersion relation forecasts that frequencies close to one another in logarithmic space give the same results [5],

$$\varphi(\omega) = \frac{\pi}{4} - \frac{\omega}{\pi} \int_0^{\infty} \log \left\{ \frac{\rho_a(x)}{\rho_0(x)} \right\} \frac{dx}{x^2 - \omega^2} \quad (7)$$

where  $\rho_a$  is the apparent resistivity and  $\rho_0 = 1/\sigma(0)$

In accordance with Weidelt's dispersion relation, 6-10 points per decade are recommended, since a higher number of points are unnecessary and a lower number of frequencies may induce exaggerated extrapolation in resistivity curves, which result in incorrect data.

iv. Nyquist Teorem

Nyquist Theorem states that any signal limited by a bandwidth  $f_0$  can be sampled without any information loss providing that the sampling frequency is higher than twice the bandwidth ( $2 f_0$ ). Not meeting this condition can result in aliasing. Applying this theorem to our case, given a sampling frequency  $f_s$ , all frequencies sampled higher than half the

sampling frequency  $\frac{f_s}{2}$  can cause aliasing. Consequently, they must be eliminated in order not to use corrupted data.

Thus, a significant contribution of this project is the production of currently non-existing list of frequencies (relative maximum ones in the spectrum), and which will allow for a better processing of high frequencies in particular.

### C. PARZEN WINDOW

Once the frequencies appropriate to audio-magnetotellurics have been determined, Parzen window is applied. This is a probability density estimation method of a value with relation to those neighbouring on it, which reduces variance in estimating its value. On the one hand, in considering various neighbouring frequencies we are taking into account various depths and therefore, various resistivities. Therefore, applying the window results in a decrease in the resolution of resistivity structures. On the other hand, having these statistical estimates makes error decrease in the selected value. Specifically, we use Parzen Gaussian window as it is better in considering neighbouring contributions [6]:

$$P(f) \begin{cases} 1 & \text{if } (f_z - f) = 0 \\ (\sin(u)/u)^2 & \text{if } |f_z - f| < f_R \\ 0 & \text{if } |f_z - f| \geq f_R \end{cases} \quad (8)$$

where  $u = \pi|f_z - f|/f_R$ ,  $f_z$  is the windows centre and  $f_R$  is the width of windows. We choose the values of  $f_z$  while  $f_R$  value is given by the next relation:  $C_R = \frac{f_z}{f_R}$ , where  $C_R$  is know how Parzen radio or Parzen parameter and we have to choose their value.

As shown in the previous equation(8), the higher the Parzen radius, the greater the number of neighbours that affect the value of the selected frequency, therefore being the obtained resolution lower, but minor the error, too (the curve will be smoother). Thus, the perfect Parzen radius will be that of a lower value, as long as the error is negligible (smooth enough).

For the same reason mentioned in the previous section (high level of signal noise), it will be interesting to perform smoothing in the resistivity curve, being a high Parzen radius more appropriate for the processing.

However, the strict selection of frequencies performed in section II.B should allow the selection of a lower Parzen radius than when processing with a list by default, thereby increasing resolution.

### D. DETERMINE AN OPTIMUM MTF FOR EACH TARGET FREQUENCY

Finally, it must be noted that there exist a wide range of methods for estimating the optimal MTF for each target frequency, but they are not all implemented in commercial software. For this reason, the focus will be on only two of the most commonly used main methods.

- Stack All: It is calculated by means of the arithmetic mean of all values.

- Selective Stack: A selection criterion of MTFs is used for calculating the arithmetic mean (We use the Mapros criterion [4]).

Overall, the results obtained in MT data processing through Selective Stack are usually better. However, both methods will be tackled in section III as applied to audio-magnetotellurics using the parameters analysed and modified in this project.

### III. RESULTS AND DISCUSSION

This section shows the implementation of the method for obtaining transfer functions which has been developed in this project using the data collected in an AMT station located in Vallès Oriental close to the town of Cardedeu, and the instrumentation used is by Metronix, ADU07e, and it is ICGC property.

Using the analysis of the parameters that affect time series processing (section II), multiple processing has been performed on real data collected in Vallès. The data were collected by means of three frequency bands that use 524 KHz, 131 KHz and 16 KHz sampling frequencies in order to cover full AMT range. The aim is to check whether the changes proposed in signal processing improve the results.

The tables below show the different processes performed, which are classified using different colours according to the quality of the  $\rho_{a_{XY}}(\omega)$  i  $\rho_{a_{YX}}(\omega)$  resulting graphics. The criterion used in assessing the graphics is as follows: a valid graphic (in green) is that which presents apparent, continuous phase, smooth resistivity curves, which present a coherence curve with values close to 1. Conversely, a non-valid graphic (in red) is that which does not meet the above criteria. A graphic that is clearly in a middle point between these two ends will be considered as acceptable (in brown).

Method	LFFFT	C <sub>R</sub>	C <sub>R</sub>	C <sub>R</sub>
Stack All	256	0,25	0,5	0,75
	512	0,25	0,5	0,75*
	1024	0,25	0,5	0,75
	2048	0,25	0,5	0,75
	4096	0,25	0,5	0,75
Selective Stack	256	0,25	0,5	0,75
	512	0,25	0,5	0,75
	1024	0,25	0,5	0,75
	2048	0,25	0,5	0,75
	4096	0,25	0,5	0,75

TABLE I: Processed table with  $F_s = 524\text{KHz}$  and our target frequency list

As shown in table I, for 524 KHz signal processing, the Stack All method clearly renders better results than the Selective Stack one. Furthermore, the best results are obtained through low LFFFT and high Parzen radius, results that are consistent with the approach in section II.

As shown in the next table (table II) the general tendency in the results rendered for 131 KHz sampling frequency are very similar to those obtained for 524 KHz. Therefore, the most effective approach is Stack all combined with low LFFFT and high Parzen radius.

Method.	LFFFT	C <sub>R</sub>	C <sub>R</sub>	C <sub>R</sub>
Stack All	256	0,25	0,5	0,75*
	512	0,25	0,5	0,75
	1024	0,25	0,5	0,75
	2048	0,25	0,5	0,75
	4096	0,25	0,5	0,75
Selective Stack	256	0,25	0,5	0,75
	512	0,25	0,5	0,75
	1024	0,25	0,5	0,75
	2048	0,25	0,5	0,75
	4096	0,25	0,5	0,75

TABLE II: Processed table with  $F_s = 131\text{ KHz}$  and our target frequency list

Finally, the results obtained for a 16 KHz sampling frequency (table III) are different from those presented above. There is a tendency change, the Selective Stack method combined with a high LFFFT begins to render better results. Despite this improvement, the best results obtained follow the tendency as shown in the previous tables.

Method	LFFFT	C <sub>R</sub>	C <sub>R</sub>	C <sub>R</sub>
Stack All	256	0,25	0,5	0,75
	512	0,25	0,5	0,75
	1024	0,25	0,5	0,75*
	2048	0,25	0,5	0,75
	4096	0,25	0,5	0,75
Selective Stack	256	0,25	0,5	0,75
	512	0,25	0,5	0,75
	1024	0,25	0,5	0,75
	2048	0,25	0,5	0,75
	4096	0,25	0,5	0,75

TABLE III: Processed table with  $F_s = 16\text{ KHz}$  and our target frequency list

Below is a graphic showing processing with default parameters used in low frequencies, MT (figure 3), applied to high frequency actual data, AMT. This figure shows the reason why this project has been undertaken, since by means of usual processing that renders good results in MT, we obtain an apparent resistivity curve and phase curves that present highly abrupt changes in the case of high frequencies, AMT, especially in the phase.

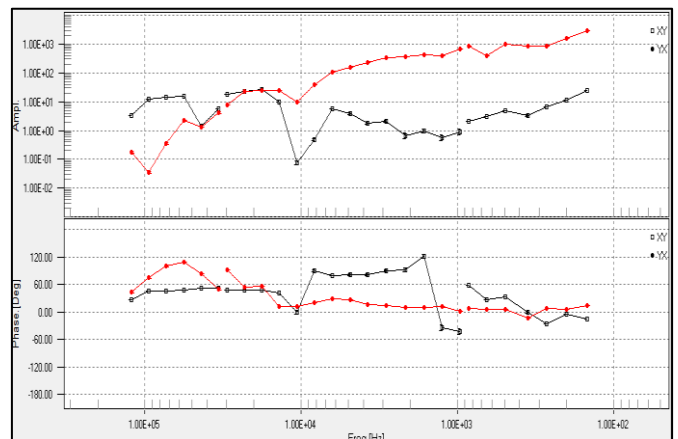
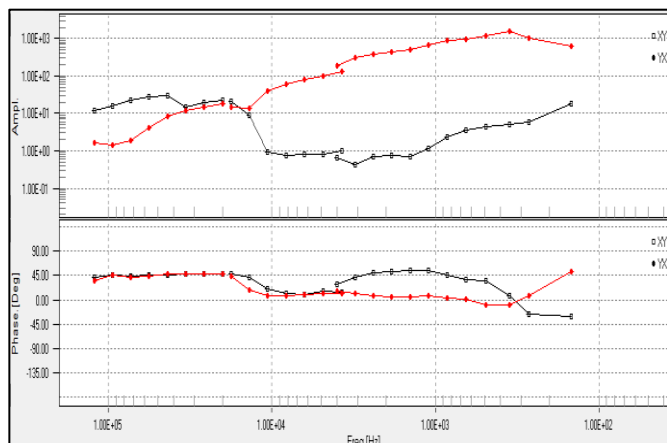


FIG. 3: Component XY and YX of apparent resistivity ( $\Omega\cdot\text{m}$ ) and phase ( $^\circ$ ) using habitual MT processing parameters: Selective Stack, default frequency list, LFFFT= 4096 and  $C_R=0,5$ .



The second graph shows the results rendered applying optimal parameters for the three AMT frequency bands (previous tables). The three frequency bands are: the leftmost part is 524Hz, the central part is 131 KHz and that to the right of the figure is 16 KHz. As mentioned before, and as shown in figure 4, using a low LFFT, a high Parzen radius, and applying a determined frequency list for each band contribute to significantly smoothing both the apparent resistivity curve and the phase curve. This will allow for better result interpretation.



**FIG. 4:** Component XY and YX of apparent resistivity ( $\Omega\cdot m$ ) and phase ( $^{\circ}$ ) using processing parameters marked with \* in tables.

The results obtained as shown in figure 4 are much better than those obtained in figure 3 because it shows apparent resistivity curves and phase curves that are smoother and with no discontinuities. Thus, when subsequently implementing inversion on these curves, programs will render a better subsurface model

#### IV. CONCLUSIONS

After in-depth analysis of time series processing obtained through the magnetotelluric method, the following conclusions can be drawn:

1) Using standard MT methodology for processing high frequency time series, AMT is not recommended. Because

the results obtained show curves presenting too abrupt variations, which hinder the production of a subsurface electrical resistivity model. Better curves are obtained through the use of the specific program implemented in Matlab, which provides a list of specific frequencies (II.B), and using some specific processing parameters, which are listed below.

Fs	524KHz	131KHz	16KHz
LFFT	512	256	1024
C <sub>R</sub>	0,75	0,75	0,75
Method	Stack All	Stack All	Stack All

**TABLE IV:** Processing parameters with optimum results.

2) Observing the graphics it can be inferred that the bands which are more sensitive to the methodology implemented for time signal processing are high frequency bands (524 KHz and 131 KHz).

3) In producing this project, measurements at 3 different sampling frequencies have been used, which have been collected in a single specific geographic location. The data can be greatly conditioned by geographic location. Therefore, it can be concluded that in spite of the good results obtained, it would be appropriate to implement the proposed method in various areas.

4) Thanks to its good results, the code produced and the methodology used in this TFG will become part of the AMT data processing techniques that are used by the ICGC geophysics team.

#### ACKNOWLEDGMENTS

I would like to thank Dr Juanjo Ledo (UB) for accepting to be my tutor and Dra. Anna Gabàs (ICGC), for accepting to be my project director. I am extremely grateful for all their help. Finally, my thanks to Lena Escalas (UB) for her collaboration.

[1] P. Piña-Varas, "Aplicación del Método Magnetotelúrico a la Caracterización de Reservorios: Anticlinal de El Hito (Cuenca y Sistema Geotérmico de Tenerife)", Thesis, 2014.

[2] F.Simpson and K.Bahr, "Practical Magnetotellurics", Cambridge University Press, 2005

[3] A.Ellnghausm, "PROCMT User's guide: Offline Processing Software For Magnetotelluric Time Series", Rev. 2.1, 2002.

[4] B. Friedrichs, "MAPROS User's guide", Ver 0.87b, 2007.

[5] Weidelt, " The inverser problem of geomagnetic inducton", Geophys 38: 257-289, 1972.

[6] C.Manoj, "Magnetotelluric Data Analysys Using Advances In Signal Processing Techinques", National Geophysical Research Institute, Hyderabad, 2003.