

Processing of noisy magnetotelluric time series from Koyna-Warna seismic region, India: a systematic approach

Ujjal K. Borah*, Prasanta K. Patro, Vaddeboina Suresh

CSIR- National Geophysical Research Institute, Magnetotellurics and Deep Resistivity Sounding Division, Hyderabad, India

Article history

Received October 27, 2014; accepted April 8, 2015.

Subject classification:

Magnetotellurics, Koyna, Time series processing, Cultural noise, Remote reference.

ABSTRACT

Rolling array pattern broad band magnetotelluric (MT) data was acquired in the Koyna-Warna (Maharashtra, India) seismic zone during 2012-14 field campaigns. The main objective of this study is to identify the thickness of the Deccan trap in and around the Koyna-Warna seismic zone and to delineate the electrical nature of the sub-basalt. The MT data at many places got contaminated with high tension power line noise due to Koyna hydroelectric power project. So, in the present study an attempt has been made to tackle this problem due to 50 Hz noise and their harmonics and other cultural noise using commercially available processing software MAPROS. Remote site was running during the entire field period to stand against the cultural noise problem. This study is based on Fast Fourier Transform (FFT) and mainly focuses on the behaviour of different processing parameters, their interrelations and the influences of different processing methods concerning improvement of the S/N ratio of noisy data. Our study suggests that no single processing approach can give desirable transfer functions, however combination of different processing approaches may be adopted while processing culturally affected noisy data.

1. Introduction

Magnetotelluric (MT) is a passive electromagnetic (EM) geophysical method to delineate the subsurface resistivity structure. MT has been widely used in different fields of geophysics, ranging from exploration surveys to deep crustal research, such that different resistivity bearing structures can be clearly identified. MT data processing plays an important role in achieving this goal. One of the key problems in MT data processing is to attenuate the cultural noise from the MT data as these noises often destructively interfere with primary information and lead to incorrect subsurface resistivity images.

MT being a natural EM method, the data gets affected by extraneous electric and magnetic field (cultural, meteorological), hence one cannot have the proper resistivity structure of the subsurface without

attenuating these noise. With the growth of the modern civilization it becomes difficult to get a site which is not contaminated by cultural noise. Many researchers [Kao and Rankin 1977, Gamble et al. 1979, Fontes et al. 1988, Egbert 1997, Yamane and Takasugi 1997, Nagata et al. 2012, and many others] have addressed this issue by developing innovative processing techniques to suppress the cultural noise. In the present study area, many of the MT stations were close to the high tension distribution power lines from Koyna and Warna hydroelectric power projects along with wind mills and were affected by other cultural noise such as wind, vehicle etc. Figure 1 shows an example of noisy time series

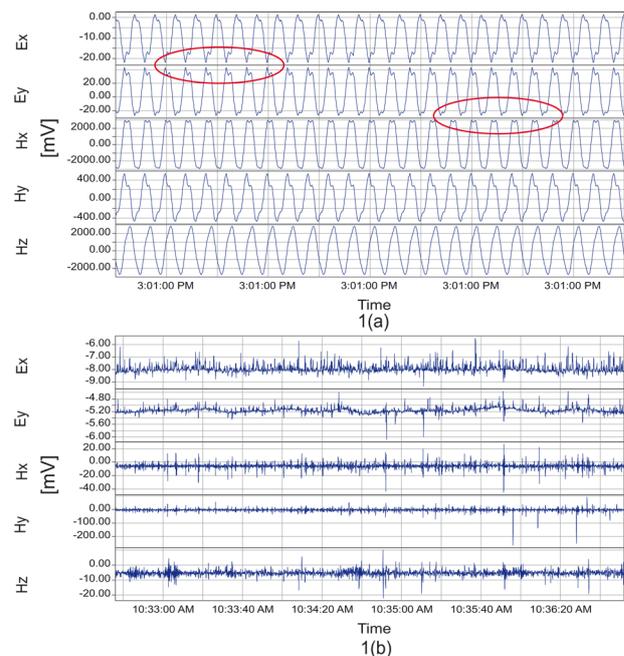


Figure 1. Time series of site K11 of Koyna profile contaminated with cultural noise. (a) Time series obtained from 4 KHz sampling frequency is affected by powerline noise (50 Hz), which shows the peak distortion (red indicator) prominently. (b) Time series from 32 Hz sampling frequency contaminated with heavy cultural noise.

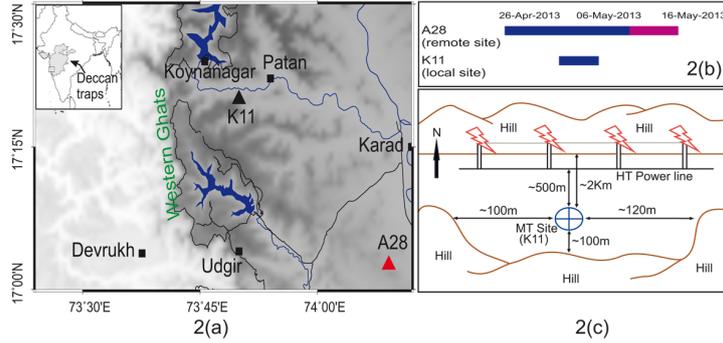


Figure 2. Local and remote sites and the surrounding topography. (a) Map showing the local site K11 (black triangle) and remote site A28 (red triangle). The remote site is in S-E direction at a distance of 50 km from the local site. The black colored squares represent the places in and around the Koyna-Warna region. (b) Data acquisition time plot between the local and the remote stations for the long period data (32 Hz sampling frequency). (c) A rough sketch of the site showing the topography and powerline distribution with approximate distances.

from the study area. Remote Reference (RR) processing technique can handle such kind of cultural noise upto certain limit [Jones et al. 1989]. But RR technique can improve the data quality if the cultural noise associated with the local and remote sites is uncorrelated [Clarke et al. 1983, Ritter et al. 1998, Uchida et al. 2005, Weckmann et al. 2005]. We have operated remote MT site (A28) nearly 50 km away from the local site (K11) for the total duration of the field campaign (Figure 2). The remote data was used for 32 Hz and lower sampling frequencies to improve the S/N ratio for the data below 10 Hz. We have found a better outcome with remote reference technique, but still, in some stations we have faced poor data quality following the same approach. However, for the short period data it was a challenge to process and derive a reasonable transfer function. In this present study, instead of looking for a new method we mainly concentrated on the methods which are already exist and try to study the behaviour of different processing parameters, their role and the interrelations among them regarding improvement of the S/N ratio using commercially available processing software, MAPROS [Friedrichs 2007]. During our study we have applied different processing approaches, viz. Selective

Stacking, Coherency Threshold, Stack All, Remote Reference with different combination of processing parameters (FFT window, FFT length, Parzen radius, etc). Finally we designed a work flow to process noisy MT data to derive an improved MT transfer function.

2. Magnetotelluric data

Broad band magnetotelluric data was acquired at more than 100 locations in and around the Koyna-Warna seismic zone, Maharashtra, western India [Gupta et al. 1969] during the 2012-14 field campaigns. Most of the region is undulating and covered with Deccan basalts. Koyna Hydroelectric Power project with capacity of 1960 MW is situated at Koyna nagar. This makes the surrounding region criss-crossed with a dense network of power lines. For the present study we have considered data from station K11 and the remote station A28 (shown in Figure 2). Data are collected at different sampling frequencies starting from 64 KHz to 32 Hz. The lower sampling frequencies (beyond 32 Hz) are obtained by filtering the 32 Hz data. Details of sampling frequencies and duration of recording for each sampling frequency is listed in Table 1.

3. MT transfer function

The basic theory of MT was proposed by a French geophysicist Louis Cagniard in the early 1950s. In MT natural electric and magnetic fields are recorded in three mutually perpendicular directions (Figure 3) to get the time series which is the basis for calculation of MT transfer function. Transfer function invokes an Earth model that describes a linear system with an input and a predicted output. The transfer function establishes a linear relationship between the physical properties that are measured in the field.

Transfer function is the ratio of output to input as a function of frequency. In MT it is represented by 'Z'. The orthogonal components of the horizontal electric (E_x and E_y) and magnetic (H_x and H_y) fields are related

Frequency	Time of acquisition	No. of acquisition
64 KHz	20 seconds	4
4 KHz	20 minutes	3
1 KHz	20 minutes	2
512 Hz	40 minutes	1
128 Hz	1 hour	1
32 Hz	3-4 days	1
Lower frequencies adopted are obtained from filtering of 32 Hz		
8 Hz, 2 Hz, 1 Hz, 2 S, 4 S, 8 S, 16 S, 32 S, 64 S, 256 S		

Table 1. Details of sampling frequencies acquired with duration of acquisition.

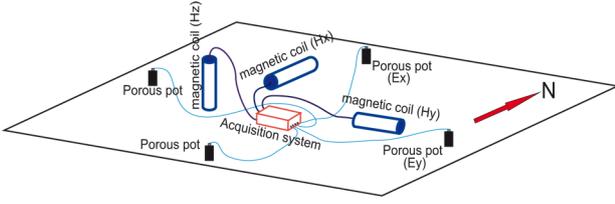


Figure 3. MT field layout. Three magnetic coils are placed in three mutually perpendicular directions (two horizontal and one vertical) to measure the natural magnetic fields. Two pairs of porous pots are put in mutually perpendicular directions (N-S and E-W) to measure the electric field induced by the magnetic field.

via a complex impedance tensor (Z) in frequency domain, which is nothing but the transfer function for the MT analysis

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} \begin{pmatrix} H_x \\ H_y \end{pmatrix} \quad (1)$$

where Z_{ij} being the components of the impedance tensor (Z) with $\{i,j\} = \{x,y\}$.

Z is complex, being composed of both real and imaginary parts. So, each component Z_{ij} , of Z has not only the magnitude but also the phase. The transfer function (impedance tensor) is usually represented by apparent resistivity Vs period and phase Vs period curves and these contain the same information as the response in the time domain. The transfer function (impedance tensor) is related to apparent resistivity and phase as follows [Cagniard 1953]

$$\rho(\omega) = (1/\omega\mu) |Z(\omega)|^2 \quad (2)$$

$$\varphi(\omega) = \tan^{-1}(\text{Im}(Z)/\text{Re}(Z)) \quad (3)$$

where ω is the angular frequency, μ is the magnetic permeability, ρ is the apparent resistivity, φ is the phase, Z is the impedance and $\text{Im}(Z)$ and $\text{Re}(Z)$ are the imaginary and real part of impedance tensor.

4. Approach for MT time series processing

In order to alleviate the problems due to cultural noise during the processing of MT time series we designed a work flow as given in the flow chart (Figure 4). In the foregoing sections we explain the steps followed in the flow chart.

4.1. Power line frequency treatment

The power line noise is associated with the signal having sampling frequency greater than 50 Hz. This interaction of power line noise can be seen in stacked

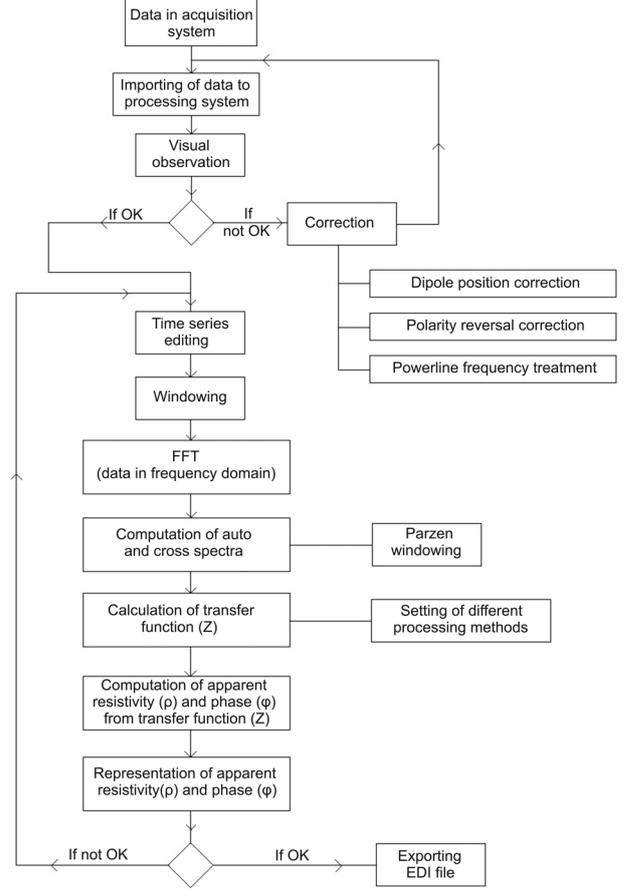


Figure 4. Flowchart followed during processing the data from Koyna-Warna region. Such work flow for culturally noisy MT data yielded desirable transfer function.

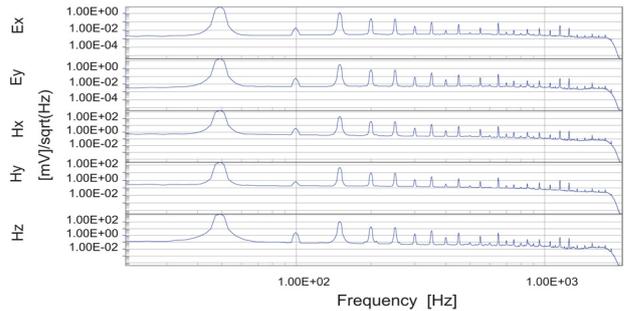


Figure 5. Stacked spectra of 4 KHz data set showing the presence of powerline frequency (50 Hz) and its harmonics (150 Hz, 250 Hz, ...) on the data.

spectra of a particular data set (Figure 5). To eliminate it we used FIR (Finite Impulse Response) notch filter with

$$\begin{aligned} \text{filter length } (L) &= N * \text{Sampling rate} + 1 \quad (4) \\ \text{where } N &= 2^n = 2, 4, 8, \dots \quad (n = 1, 2, 3, \dots) \end{aligned}$$

and tried to suppress the 50 Hz frequency and its harmonics taking a filter bandwidth, $\Delta f (= f_h - f_l)$ for notch filtering according to

$$f_c = \sqrt{(f_l \cdot f_h)} \quad (5)$$

f_c is the centre frequency (50 Hz and its harmonics such as 150 Hz, 250Hz ...) and f_l, f_h are the lower and upper cut off frequencies. Figure 6a shows the stacked spectra of data set with 4 KHz sampling frequency after notch filtering at 50 Hz. Sometimes the higher harmonics of 50 Hz (i.e. 150 Hz, 250 Hz ...) had strong influence then we had gone for notch filtering for these harmonics stepwise. Figure 6b shows the stacked spectra of the same after applying notch filtering at 50 Hz and its harmonic 150 Hz. This notch filter was used for each frequency band with frequency greater than 50 Hz. Figures 7a and 7b shows the effect of FIR notch filtering on 4 KHz data (4096 sampling frequency) of site K11. The effectiveness of the notch filter can be realised through the coherency plot of the same data set (Figures 7c and 7d). From Figures 7c and 7d it is seen that for the high frequency portion (between 1000 Hz and 100 Hz) the coherency is improved prominently after notch filtering, while the coherency of the low frequency portion (below 100 Hz) is not improved significantly, still curves are improved. The scattering of points in low frequency part is due to the overlapping

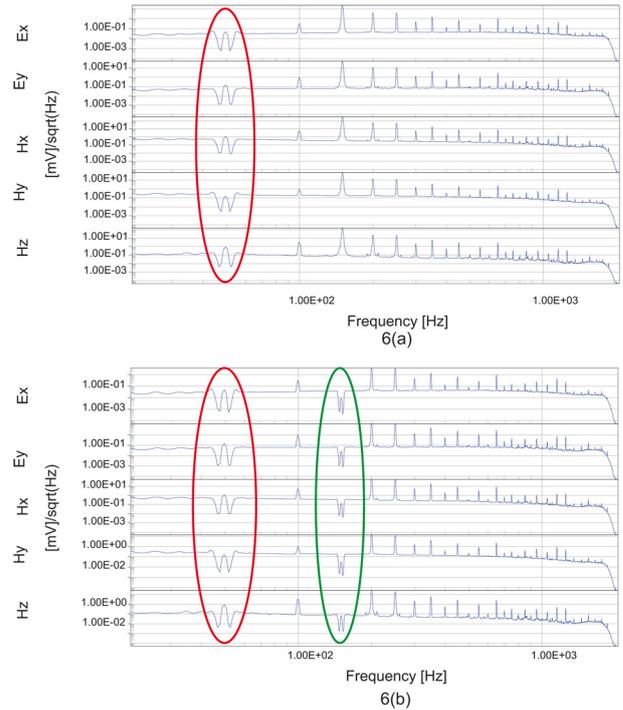


Figure 6. Stacked spectra of 4 KHz data set after notch filtering. (a) Stacked spectra showing the effect of FIR notch filtering on 50 Hz frequency which is indicated by red mark. (b) Stacked spectra showing the effect of FIR notch filtering on 50 Hz and its harmonic 150 Hz frequencies which are shown by red and green mark respectively.

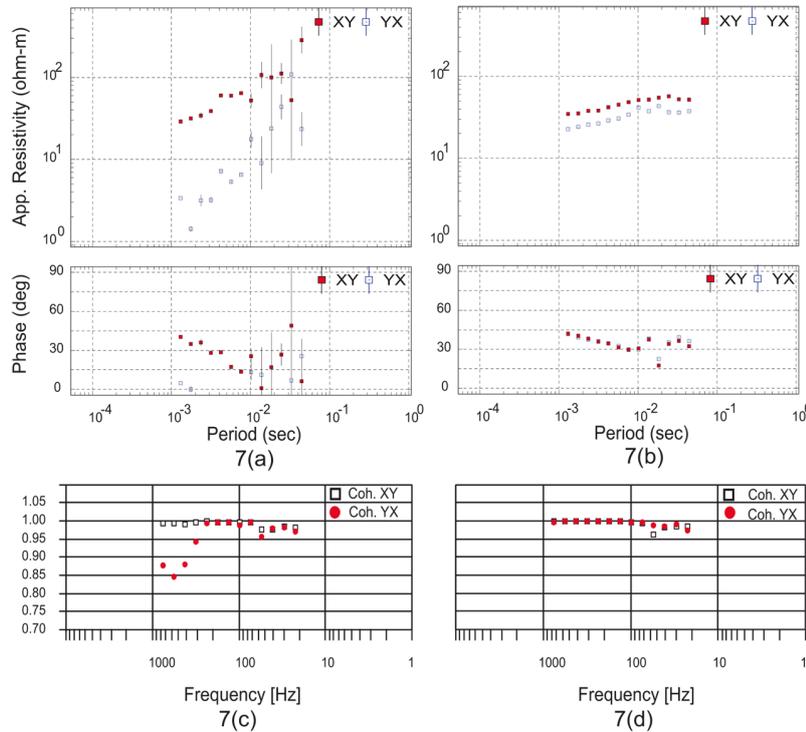


Figure 7. Influence of FIR (Finite Impulse Response) notch filter on data with 4096 Hz sampling frequency of site K11 from Koyna profile. The filter length applied here is 16,385 (with $N = 4; 4 \times 4096 + 1$) and the filter bandwidth is 10Hz with lower and upper cut off frequencies 45 Hz and 55 Hz ($\Delta f = 10 \text{ Hz} = 55 \text{ Hz} - 45 \text{ Hz}$) to suppress the power line frequency at 50 Hz. (a) Apparent resistivity (top panel) and phase (bottom panel) curves obtained from the data before notch filtering. (b) Apparent resistivity (upper panel) and phase (lower panel) curves of the same data obtained after notch filtering showing an improved result. The red colored boxes represent XY components and the blue colored boxes represent YX components of apparent resistivity an phase. The effectiveness of the notch filter can be realised from the coherency plots. (c) Coherency plot of the same data set before notch filtering. (d) Coherency plot after notch filtering shows an improvement in coherency of the components. Here boxes represent the coherency of XY components and the red circles represents the coherency of YX components.

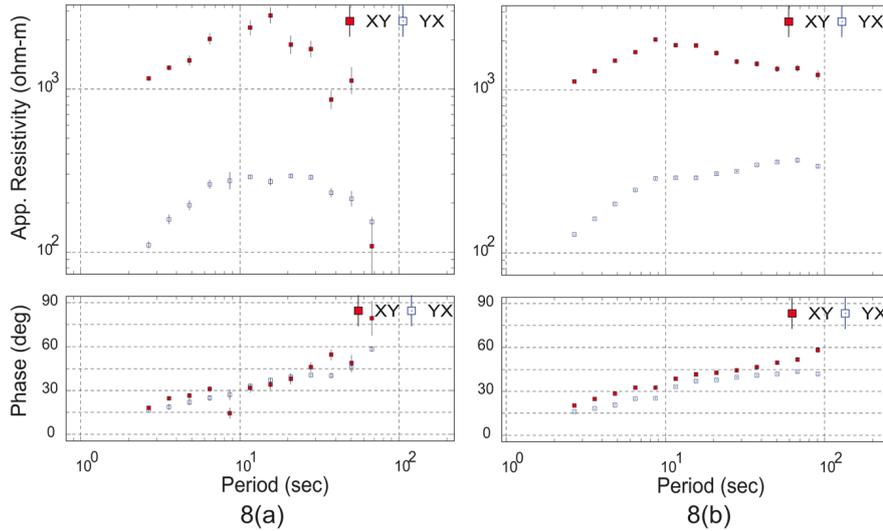


Figure 8. Effect of editing on 2 Hz data of site K11 from Koyna profile. (a) Apparent resistivity (top panel) and phase (bottom panel) curves obtained from the data before editing the time series. (b) Apparent resistivity (upper panel) and phase (lower panel) curves of the same data obtained after editing the time series provide a better outcome. The red colored boxes indicate XY components and the blue colored boxes indicate YX components of apparent resistivity and phase.

of power line noise and so, it is improved after notch filtering. But, there may be some coherent noise associated with this frequency band. So, the coherency does not show a proper improvement with the improvement of the curves.

4.2. Time series editing

Another pre-processing technique is manual time series editing. It is the simplest and most important way to handle with the noisy data. In manual editing we selected the noise burst, spikes and uncorrelated portion (which are not in trend with the entire time series) from the time series and not allowed these selected portions during processing. We did not replace these bad data portion with some other values or manipulation for it. Figure 8 describes the influence of editing on 2 Hz data.

After adopting above procedure one can get a better outcome up to certain level but for further improvement one should take the help of different processing parameters and processing methods.

4.3. Influences of different processing parameters

4.3.1. Window function (windowing)

Our study is based on Fast Fourier Transform (FFT). FFT assumes time series to be stationary. But the source of MT is related to the transient EM fluctuation of ionosphere and in general, it provides non stationary signals. Non-stationary signals are generally non-periodic. Besides this, non periodicity of a signal becomes more prominent when a random noise interacts with the signal (Figure 9). When a non-periodic signal is treated with FFT then the out coming frequency spectrum is affected by leakage. Due to leakage signal en-

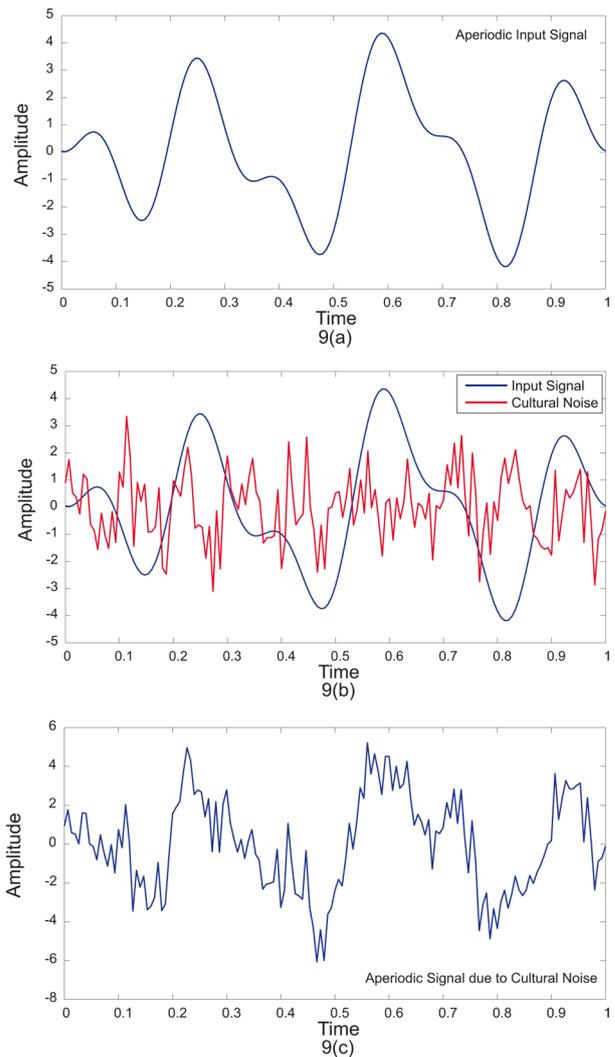


Figure 9. Effect of cultural noise on periodicity of input signal. (a) Aperiodic input signal. (b) Interaction of cultural noise with the input signal. (c) Aperiodicity of the signal increases after interaction of cultural noise.

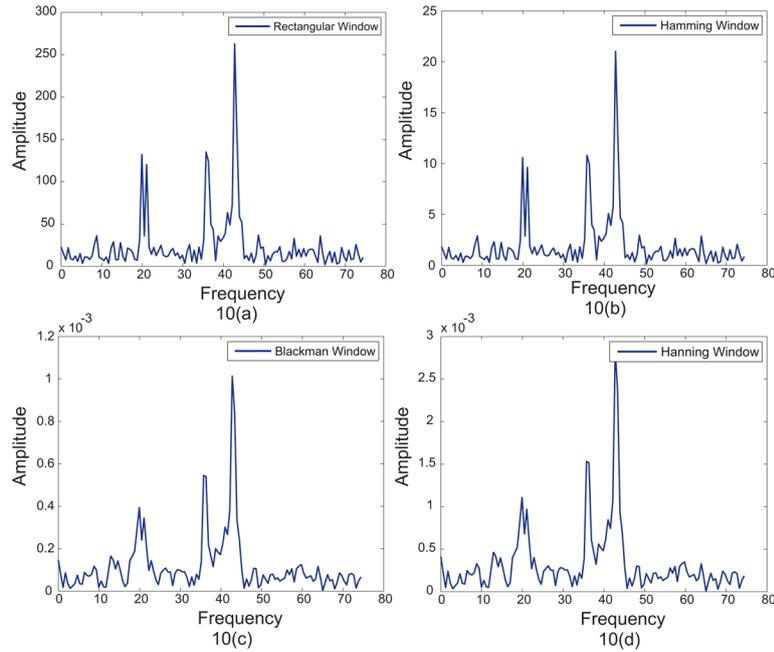


Figure 10. Effect of windowing on a signal containing nearby interfering frequency component (cultural noise). (a) Effect of Rectangular window on the signal with frequency components 43 Hz, 36 Hz, 20 Hz which interact with random noise of frequency 21 Hz. (b) Effect of Hamming window on the same data set. (c) Effect of Blackman window on the same data set. (d) Effect of Hanning window on the same data set. Rectangular and Hamming windows have more frequency resolution between the signal frequency component of 20 Hz and noise of 21 Hz.

ergy spread out over a wide frequency range after FFT, but the energy should be in a narrow frequency range. So we have to minimize this leakage during FFT. Windowing plays an important role in this regard. Window function is a mathematical function with a zero value outside some specific interval and a particular shape in between. Windowing is just multiplication of the window with the time series forcing it to be periodic to minimize the leakage effect.

4.3.1.1. Selection of window function

We have used four window functions - namely Rectangular, Hamming, Hanning and Blackman - and studied the characteristics of each window and made a comparison among them.

There is no any specific approach to select a window. Every window has its own characteristics based on which different windows are used for different purpose. To select a better window one should have an idea about the signal frequency content. If the signal is contaminated with interfering frequency (cultural noise) which is near to the frequency of interest then frequency resolution is important and in this regard a window with a narrow main lobe and low maximum side lobe level (e.g. Rectangular and Hamming) will give a desirable outcome (Figure 10). If the signal carries interfering frequency component (cultural noise) which is away from the frequency of interest then a window with a high side lobe roll off rate is better so that after FFT it will not interfere with signal frequency component significantly. But in case of large frequency

separation windows have not much effect on frequency resolution after FFT (Figure 11). So, it is better to take a window with narrow main lobe, low main side lobe level and high side lobe roll off rate to get distinction between different frequency components after FFT. Here, we have found that if the nature of the interfering component is not known then it is better to start with Hamming window (as its characteristics are suitable for most of the situations) so that we can have a better frequency resolution and can avoid the nearby noise frequency during stacking for a particular signal frequency. Figure 12 describes the effect of different windows on data with 8 Hz sampling frequency from site K11. If the data is not contaminated by cultural noise then windows do not have much effect and different window produce almost similar result. Figure 13 shows the effect of different windows on data with 1 Hz sampling frequency from the remote site A18, which is free from cultural noise. The result shows a similar outcome from different windows.

4.3.2. Fast Fourier Transform (FFT)

After the time series has been treated with windowing the FFT is applied to the data which transformed the time series into the frequency domain.

4.3.2.1. Selection of FFT length

FFT length or FFT size refers to the number of data points contained in the window during FFT. Let us consider a FFT with data points (N) 1024 (i.e. with FFT length 1024) which is sampled at a frequency (sam-

MT DATA PROCESSING

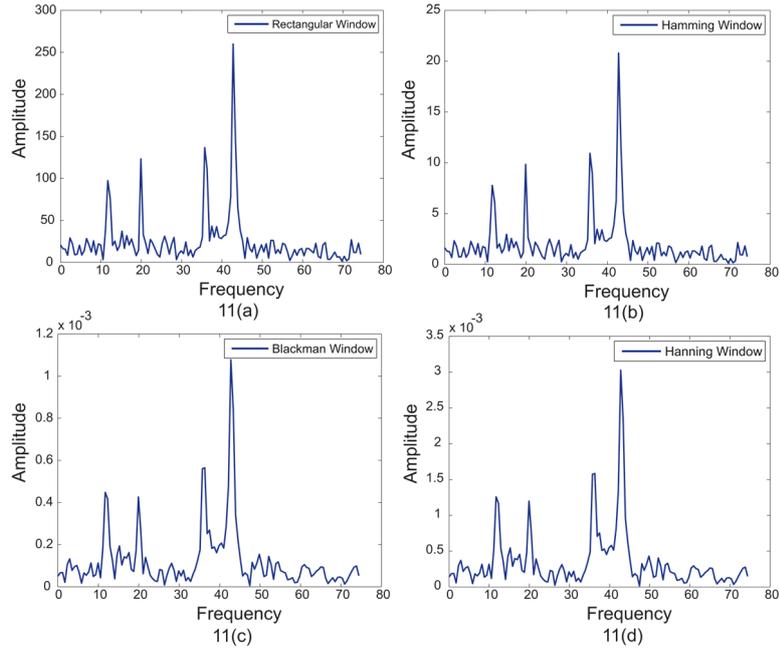


Figure 11. Effect of windowing on signal containing far interfering frequency component (cultural noise). (a) Effect of Rectangular window on a signal with frequency components 43 Hz, 36 Hz, 20 Hz which interact with a random noise of frequency 12 Hz. (b) Effect of Hamming window on the same data set. (c) Effect of Blackman window on the same data set. (d) Effect of Hanning window on the same data set. All windows have similar frequency resolution for this far interfering frequency after FFT.

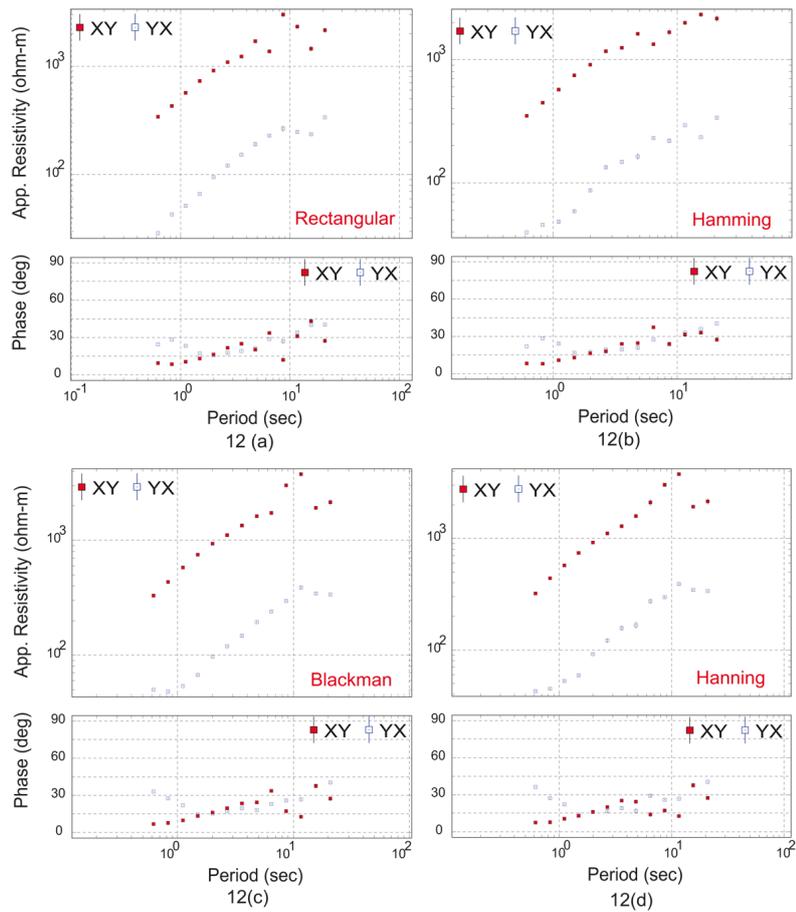


Figure 12. Result obtained after windowing with a fixed window length of 1024 from different window functions on 8 Hz data of site K11 from Koyna profile. (a) Apparent resistivity (upper panel) and phase (lower panel) curves obtained using Rectangular window. (b) Apparent resistivity (top panel) and phase (bottom panel) curves obtained applying Hamming window. (c) Apparent resistivity (top panel) and phase (bottom panel) curves obtained from Blackman window. (d) Apparent resistivity (upper panel) and phase (lower panel) curves obtained using Hanning window. Hamming window provides a better result than the others which is followed by the output of Blackman window. The red colored boxes represent XY components and the blue colored boxes represent YX components of apparent resistivity and phase.

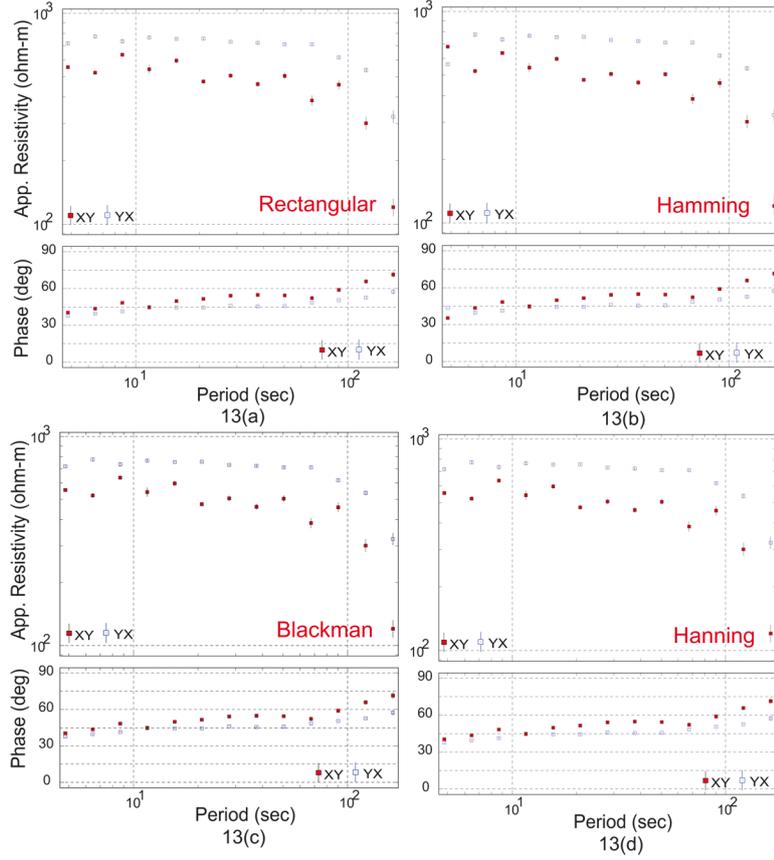


Figure 13. Result obtained after windowing with a fixed window length of 1024 from different window functions on 1 Hz data of Remote Site A28. (a) Apparent resistivity (upper panel) and phase (lower panel) curves obtained using Rectangular window. (b) Apparent resistivity (top panel) and phase (bottom panel) curves obtained applying Hamming window. (c) Apparent resistivity (top panel) and phase (bottom panel) curves obtained from Blackman window. (d) Apparent resistivity (upper panel) and phase (lower panel) curves obtained using Hanning window. All the windows have almost similar outcome. The red colored boxes represent XY components and the blue colored boxes represent YX components of apparent resistivity and phase.

pling frequency, f_s) 1024 Hz. Then its Nyquist frequency will be 512 Hz. So, after FFT, the spectrum will contain $N/2$ spectral lines with frequency value below 512 Hz. So, bandwidth for this spectrum will be 512 Hz. Now, the frequency resolution for the spectra is

$$\text{Frequency resolution } (\Delta f) = \frac{\text{Sampling frequency } (f_s)}{\text{FFT length } (N)} \quad (6)$$

So, if we use 1024 FFT length it will produce 512 spectral lines evenly spaced at 1 Hz apart. Again, if we apply 512 and 256 FFT lengths then it will produce 256 and 128 spectral lines evenly spaced at 2 Hz and 4 Hz apart upto 512 Hz respectively. So, for these three FFT lengths, for a particular sampling frequency, bandwidth remain the same but the resolution will change. So, a large FFT length gives high frequency resolution than a low FFT length for the target frequency.

Now, in case of data set affected by cultural noise, for a particular sampling frequency, the number of stack should be high so that the events are highlighted from the random noise. Also we should take large FFT length to get high frequency resolution so that we can

make distinction between signal and noise in frequency domain. But for a fixed number of data points if we take a large FFT length then it will decrease the number of stacks. Again, to increase the number of stacks if we take small FFT length then it will decrease the frequency resolution. So, one should have a criteria for optimal FFT length on the basis of number of stacks for a data set affected by cultural noise. During our study we have established empirical criteria for optimal FFT length on the basis of number of stacks which is related to the available data points of a particular sampling frequency and the FFT length to be used. The condition for optimal FFT length is as follows

$$\text{No of stacks} \geq 20 \quad (7)$$

i.e. minimum number of stacks should be 20 in our case. But to get the optimal number of stacks which is related to the optimal FFT length, we generalised the no of stacks by multiplying a multiplier with 20, i.e.

$$\text{No of stacks} \geq 20n \quad (8)$$

where $n = 1, 2, 3, \dots$ is the weight on the minimum number of stacks

$$\text{or } \left(\frac{\text{Total number of data points}}{\text{FFT length}} \right) \geq 20n \quad (9)$$

$$\text{or } \text{FFT length} \leq \left(\frac{\text{Total number of data points}}{20n} \right) \quad (10)$$

Now, the optimal FFT length can be expressed in terms of minimum FFT length to be taken. So, we have

$$\text{Optimal FFT length} = \text{Minimum FFT length} \times 2^n \quad (11)$$

where $n = 1, 2, 3, \dots$

So, Equation (10) can be written as

$$\text{Optimal FFT length} = \text{Minimum FFT length} \times 2^n \leq \left(\frac{\text{Total number of data points}}{20n} \right) \quad (12)$$

So, for a fixed number of data points of a particular data set the optimal FFT length will be obtained when the above condition (Equation 12) is satisfied for a particular 'n' value.

In our study we used 7 FFT lengths viz. 16384, 8192, 4096, 2048, 1024, 512, and 256 (minimum FFT length). Let us first consider a small data set with 19,303 data points of 16S sampling frequency from site K11. For this data set we have

$$\text{Optimal FFT length} = 256 \times 2^1 \leq \left(\frac{19303}{20 \times 1} \right) \text{ for } n = 1 \quad (13)$$

512 < 965.15 (inequality condition is satisfied)

$$\text{Optimal FFT length} = 256 \times 2^2 \leq \left(\frac{19303}{20 \times 2} \right) \text{ for } n = 2 \quad (14)$$

1024 < 482.575 (inequality condition is violated)

So, for this data set available FFT lengths which will provide number of stacks more than 20 are 512 and 256. The inequality condition breaks after FFT length 512. So, we should take FFT length below 1024. As large FFT length gives better frequency resolution so for this data set 512 will be the optimal FFT length. Figure 14 shows the effect of FFT lengths on this small data set.

Now consider a medium data set with 154,587 data points of 2S sampling frequency from site K11. For this data set we have

$$\text{Optimal FFT length} = 256 \times 2^1 \leq \left(\frac{154587}{20 \times 1} \right) \text{ for } n = 1 \quad (15)$$

512 < 7729.35 (inequality condition is satisfied)

$$\text{Optimal FFT length} = 256 \times 2^2 \leq \left(\frac{154587}{20 \times 2} \right) \text{ for } n = 2 \quad (16)$$

1024 < 3864.675 (inequality condition is satisfied)

$$\text{Optimal FFT length} = 256 \times 2^3 \leq \left(\frac{154587}{20 \times 3} \right) \text{ for } n = 3 \quad (17)$$

2048 < 2576.45 (inequality condition is satisfied)

$$\text{Optimal FFT length} = 256 \times 2^4 \leq \left(\frac{154587}{20 \times 4} \right) \text{ for } n = 4 \quad (18)$$

4096 < 1932.3375 (inequality condition is violated)

So, for this data set available FFT lengths which will provide number of stacks more than 20 are 4096, 2048, 1024, 512 and 256. But the inequality condition breaks after FFT length 2048. So, we should take FFT length below 4096. As large FFT length gives better frequency resolution so for this data set 2048 will be the

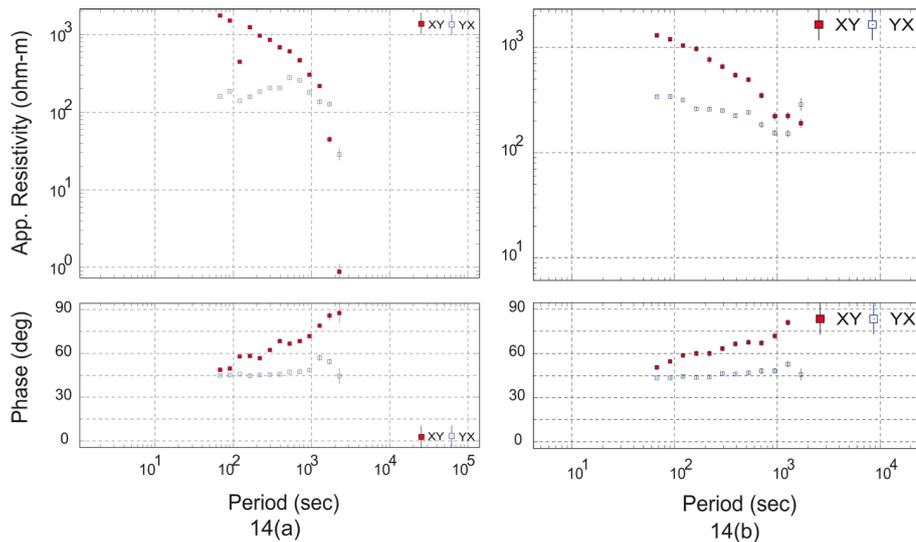


Figure 14. Effect of FFT length on a small culturally affected noisy data set of 16S sampling frequency of site K11 with 19,303 data points. (a) Apparent resistivity (upper panel) and phase (lower panel) curves obtained from the data using a FFT length of 1024. (b) Apparent resistivity (top panel) and phase (bottom panel) curves obtained applying a FFT length of 512 on the same data set provides a better outcome as this FFT length can give optimal number of stacks for this small disturbed data set. The red colored boxes represent XY components and the blue colored boxes represent YX components of apparent resistivity and phase.

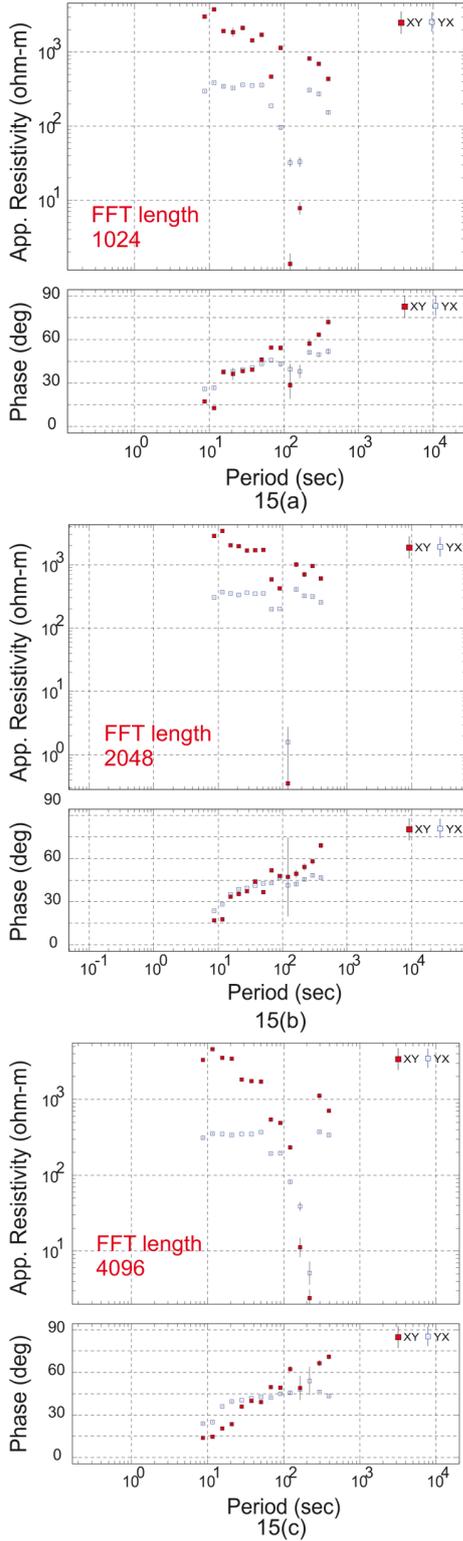


Figure 15. Effect of FFT length on a medium sized culturally affected noisy data set of 2S sampling frequency of site K11 with 154,587 data points. (a) Apparent resistivity (upper panel) and phase (lower panel) curves obtained from the data using a FFT length of 1024. (b) Apparent resistivity (top panel) and phase (bottom panel) curves obtained applying a FFT length of 2048 on the same data set. (c) Apparent resistivity (top panel) and phase (bottom panel) curves obtained applying a FFT length of 4096. FFT length 2048 provides a better outcome as this FFT length can give optimal number of stacks for this medium sized disturbed data set. The red colored boxes represent XY components and the blue colored boxes represent YX components of apparent resistivity and phase.

optimal FFT length. Figure 15 shows the effect of FFT lengths on this medium data set.

Now consider a large data set with 2,473,774 data points of 8 Hz sampling frequency from site K11. So, for this data set we have

$$\begin{aligned} \text{Optimal FFT length} = \\ 256 \times 2^1 \leq \left(\frac{2473774}{20 \times 1} \right) \text{ for } n = 1 \quad (19) \\ 512 < 123688.7 \text{ (inequality condition is satisfied)} \end{aligned}$$

$$\begin{aligned} \text{Optimal FFT length} = \\ 256 \times 2^6 \leq \left(\frac{2473774}{20 \times 6} \right) \text{ for } n = 6 \quad (20) \\ 16384 < 20614 \text{ (inequality condition is satisfied)} \end{aligned}$$

$$\begin{aligned} \text{Optimal FFT length} = \\ 256 \times 2^7 \leq \left(\frac{2473774}{20 \times 7} \right) \text{ for } n = 7 \quad (21) \\ 32768 < 17669.81 \text{ (inequality condition is violated)} \end{aligned}$$

So this large data set allows a large FFT length for stacking. As 16384 is the highest FFT length we had used, so, it will be the optimal FFT length for this data set. Figure 16 shows the effect of FFT length on this large data set.

If the data is not contaminated by cultural noise then the FFT lengths above and below the optimal FFT length have not much effect on the output; as during stacking only signals are get stacked. So, in this case stacking is not an issue. The main thing is the frequency resolution. So, to get a proper resolution we should use large FFT length in case of noise free data. Figure 17 explains the effect of FFT lengths on 2S sampling frequency data from a noise free site A28 (Remote Reference station).

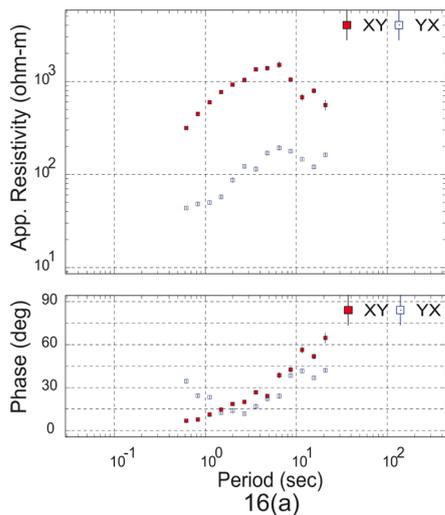
Besides taking this optimal FFT length we followed overlapping time series processing which helped us to increase the number of stacks without relating to the FFT length. We used 40% of the FFT length for overlapping and obtained a bit effective result. Figure 18 shows the effect of overlapping on 1 Hz data set.

4.4. Computation of auto and cross spectra

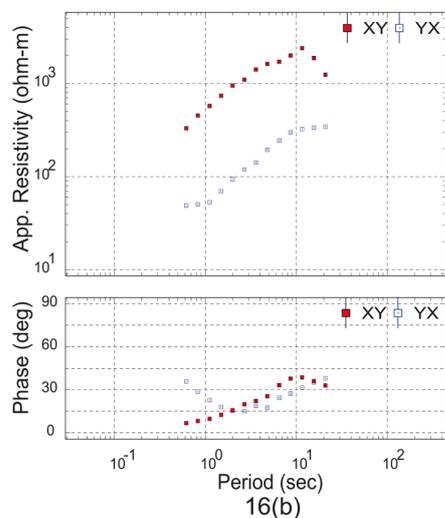
The orthogonal components of the horizontal electric and magnetic fields are combined to get the impedance tensor, Z . So, to determine the transfer functions, we need two reference signals (say 'A' and 'B'). 'A' and 'B' are combined with input and output and get an equation which is in terms of auto and cross power. Auto and cross spectra can be generated in the frequency domain by multiplying electric and magnetic spectra by their complex conjugates.

4.5. Parzen window

For each target frequency cross and auto spectra are calculated from frequency spectra. During the calculation, each value of cross and auto spectra for a par-



16(a)



16(b)

Figure 16. Effect of FFT length on a large data set of 8 Hz of site K11 with 2,473,774 data points. (a) Apparent resistivity (upper panel) and phase (lower panel) curves obtained from the data using a small FFT length 512. (b) Apparent resistivity (top panel) and phase (bottom panel) curves obtained applying a large FFT length of 16384 from the same data set provide a better resolution than the small one. The red colored boxes represent XY components and the blue colored boxes represent YX components of apparent resistivity and phase.

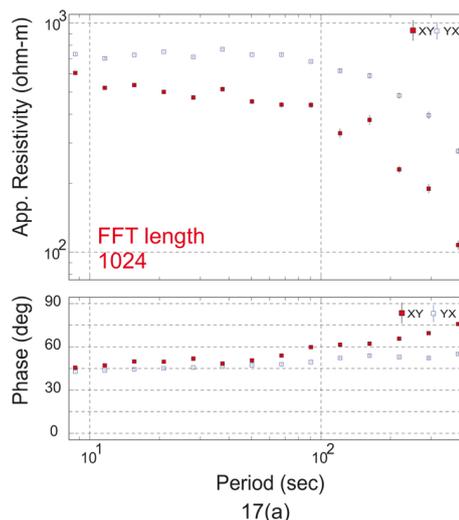
ticular target frequency obtain from the summation of all the spectral lines which fall under a window taking the target frequency at the middle. Here, we have used the Parzen window for smoothing the spectra. Figure 19 shows the Parzen window over the target frequency f_t .

If f_t is the target frequency and f_r is the radial frequency then we have [Friedrichs 2007]

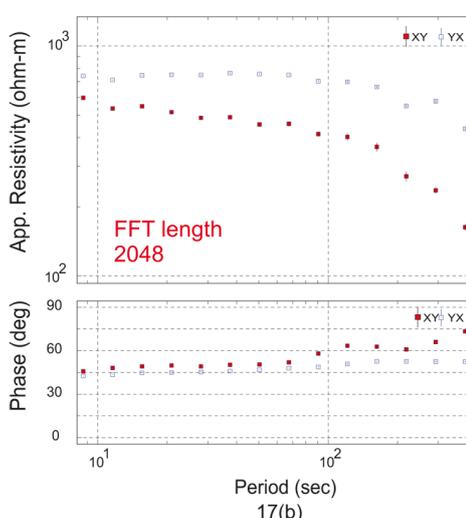
$$u = \pi |f_t - f| / f_r \quad (22)$$

$$\text{parzen}(f) = \begin{cases} 1 & \text{if } |f_t - f| = 0 \\ \text{or } \{\sin(u)/u\}^4 & \text{if } 0 < |f_t - f| < f_r \\ \text{or } 0 & \text{if } |f_t - f| \geq f_r \end{cases} \quad (23)$$

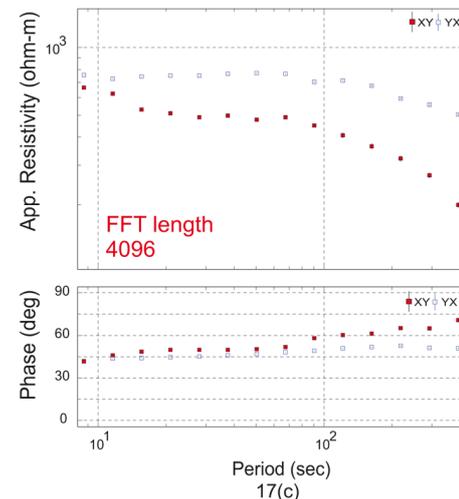
where, f_t = target frequency; f_r = radial frequency which determines the width of the window.



17(a)



17(b)



17(c)

Figure 17. Effect of FFT length on a data set of 2S sampling frequency of noise free remote site A28 with 659,399 data points. (a) Apparent resistivity (upper panel) and phase (lower panel) curves obtained from the data using a FFT length of 1024. (b) Apparent resistivity (top panel) and phase (bottom panel) curves obtained applying a FFT length of 2048 on the same data set. (c) Apparent resistivity (top panel) and phase (bottom panel) curves obtained applying a FFT length of 4096. As this data is free from random noise so during stacking only signals get stacked and so all the FFT lengths provide almost similar outcome. The red colored boxes represent XY components and the blue colored boxes represent YX components of apparent resistivity and phase.

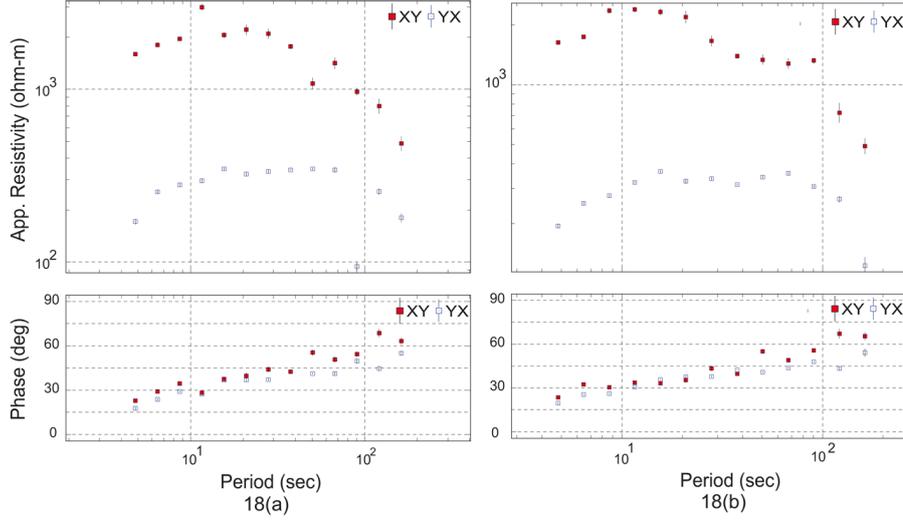


Figure 18. Effect of overlapping processing on a data set of 1 Hz sampling frequency of site K11. (a) Apparent resistivity (upper panel) and phase (lower panel) curves obtained from the data without overlapping processing. (b) Apparent resistivity (upper panel) and phase (lower panel) curves obtained from the data with overlapping processing provides an improvement of the curves between 10^1 to 10^2 sec period.

The target frequency and the radial frequency are related as

$$f_r = f_t C_r \quad (24)$$

where, C_r is the Parzen radius.

4.5.1. Selection of Parzen radius

Parzen radius (C_r) to be selected is related to the frequency resolution and the radial frequency (f_r) of the Parzen window. As for a particular target frequency (f_t) we averaged it by summing the spectral lines nearby. So, for a target frequency we have to take the minimum radial frequency greater than the frequency resolution so that at least one spectral line on both side of the target frequency can come under the Parzen window. So, we have

$$f_r > \text{frequency resolution} \quad (25)$$

$$f_r > \left(\frac{\text{sampling frequency}}{\text{FFT length}} \right) = f_s / N \quad (26)$$

Again we have $f_r = f_t C_r$

$$\text{So,} \quad f_t C_r > f_s / N \quad (27)$$

$$\text{or} \quad C_r > f_s / (f_t N) \quad (28)$$

Now, if we consider the highest sampling frequency 64 KHz (in the order of 10^4 Hz) that we have, then it will be related to the target frequency in the order of $(\frac{1}{2}) 10^4$ Hz.

So, we have

$$f_s / (f_t N) = \frac{64000}{(\frac{1}{2}) 10000 N} = \frac{12.8}{N} \quad (29)$$

Now for $N = 16384$ (highest FFT length)

$$f_s / (f_t N) = \frac{12.8}{16384} = 0.00078 \quad (30)$$

and for $N = 256$ (lowest FFT length)

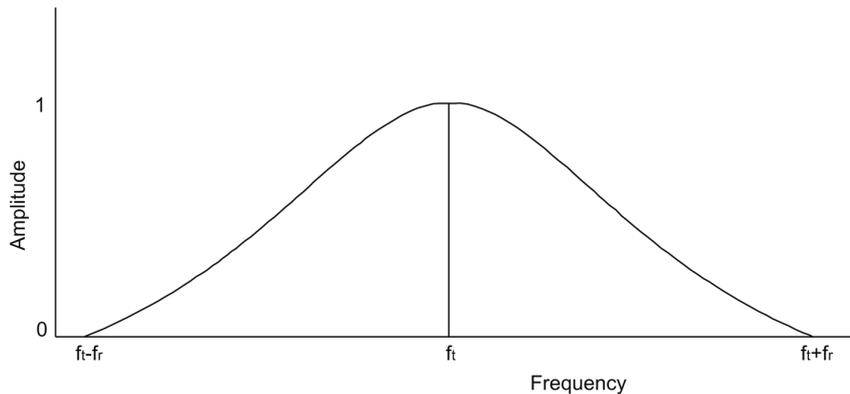


Figure 19. Parzen window over the target frequency ' f_t ' in frequency domain. f_t is the radial frequency which determines the value of Parzen radius.

$$f_s / (f_t N) = \frac{12.8}{256} = 0.05 \quad (31)$$

Again, if we consider a low sampling frequency, say 128 S or 0.0078 Hz (in the order of 10^{-3} Hz) then it will be related to the target frequency in the order of $(\frac{1}{2}) 10^{-3}$ Hz. Then we have

$$f_s / (f_t N) = \frac{0.0078}{(\frac{1}{2})_{0.001} N} = \frac{15.6}{N} \quad (32)$$

So, for $N = 16384$ (highest FFT length)

$$f_s / (f_t N) = \frac{15.6}{16384} = 0.00095 \quad (33)$$

And for $N = 256$ (lowest FFT length)

$$f_s / (f_t N) = \frac{15.6}{256} = 0.06 \quad (34)$$

So, we have found that value of Parzen radius not exceeds 1

$$\text{i.e.} \quad C_r < 1. \quad (35)$$

So, we can conclude from Equations (28) and (35) that the range of Parzen radius is

$$1 > C_r > f_s / (f_t N) \quad (36)$$

A high sampling frequency data set gives more number of data points than a low sampling frequency data set. During our study we have established an empirical relation regarding optimal Parzen radius to be taken for data affected by cultural noise. We studied this situation in two cases:

- a) for the data sets having sampling frequency less than 1 Hz
- b) for the data sets having sampling frequency greater than 1 Hz.

In case of long period data (low sampling frequency < 1 Hz), if the sampling frequency (f_s) is in the order of 10^{-n} Hz then we should take the radial frequency (f_r) for Parzen windowing in the order of $k10^{-(n+1)}$ (with $k = 1, 2, 3, \dots$). Then we have the Parzen radius in the order of:

$$k10^{-(n+1)} \approx (1/2) 10^{-n} C_r \quad (37)$$

$$\text{or} \quad C_r \approx 2k10^{-1} \quad (38)$$

where k should be taken according to Equation (36). So, $k = 1, 2, 3, 4$

$$\begin{aligned} \text{for } k = 1; \quad C_r &\approx 2 \times 1 \times 10^{-1} = 0.2 \\ \text{for } k = 4; \quad C_r &\approx 2 \times 4 \times 10^{-1} = 0.8 \end{aligned}$$

So, range of Parzen radius for low sampling fre-

quency less than 1 Hz is 0.2 to 0.8.

Again, in case of high frequency sampling frequency data sets (frequency > 1 Hz) we considered two conditions:

1) If the order of the sampling frequency is 10^n (where $n = 3, 4$) then we should take the radial frequency in the order of $k10^{n-2}$. So, the order of Parzen radius will be

$$k10^{n-2} \approx (1/2) 10^n C_r \quad (39)$$

$$\text{or} \quad C_r \approx 2k10^{-2} \quad \text{with } k = 1, 2, 3, 4 \quad (40)$$

$$\text{for } k = 1; \quad C_r \approx 0.02$$

$$\text{for } k = 4; \quad C_r \approx 0.08$$

So, range of Parzen radius for high frequency (in the range $10^3, 10^4$ Hz) is 0.02 to 0.08.

2) If the order of sampling frequency is 10^n (where $n = 1, 2$) then the radial frequency should be in the order of $k10^{n-1}$. So, the order of Parzen radius will be

$$k10^{n-1} \approx (1/2) 10^n C_r \quad (41)$$

$$\text{or} \quad C_r \approx 2k10^{-1} \quad \text{with } k = 1, 2, 3 \quad (42)$$

$$\text{for } k = 1; \quad C_r \approx 0.2$$

$$\text{for } k = 3; \quad C_r \approx 0.6$$

So, range of Parzen radius for high frequency (in the range $10^1, 10^2$ Hz) is 0.2 to 0.6.

4.6. Relation between FFT length and Parzen radius

The FFT length and the Parzen radius are very important parameters and they are inversely related (Equation 28).

Our main aim is to increase the resolution of the target frequency. If we use high FFT length it will produce distinct spectral lines with small frequency separation. So, while smoothing with Parzen window we should take small Parzen radius so that we can average out the target frequency from the spectral lines which are close to the target frequency. But in case of small data set we have to use small FFT length and this will produce spectral lines with large frequency separation. Because of this the resolution becomes poor. So, during the smoothing with Parzen window we should use large Parzen radius such that neighbouring spectral lines can come under the window for averaging the target frequency. But in this case if we take a small Parzen radius then we may not get spectral lines for averaging the target frequency. So, during our study we optimised the FFT length and Parzen radius for a particular data

set on the basis of number of data points (Equation 12) and sampling frequency (Equations 36, 38, 40, 42) respectively. Figures 20 and 21 show the interrelationship between FFT length and Parzen radius in a small data set and in a large data set respectively.

4.7. Different methods for estimation of transfer function

In this present study we focused on four types of processing methods namely- Selective Stacking, Coherency Threshold, Stack All and Remote Reference. These methods are explained below. Later section how these methods are used for our dataset is discussed.

4.7.1. Selective Stacking

Selective Stacking try to minimize the outlier problem by excluding a number of matrices from taking an average. This method is based on standard deviation of the calculated spectral matrices. Actually, standard deviation shows how much variation from the average or mean exists. Selective Stacking adds those matrices which are falling inside the standard deviation and the extreme matrices are eliminated.

Suppose we have ‘n’ spectral matrices ($e_1, e_2, e_3, \dots, e_n$). Then, we have

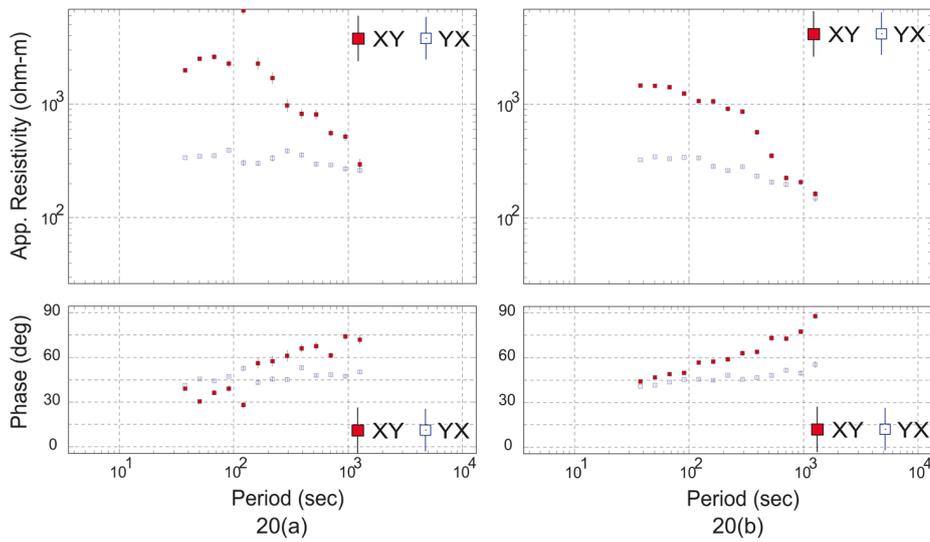


Figure 20. Interrelation between FFT length and Parzen radius for a small data set of 8S of site K11 with 38,628 data points. (a) Apparent resistivity (top panel) and phase (bottom panel) curves obtained from the data using a FFT length of 2048 and a Parzen radius of 0.13 (b) Apparent resistivity (upper panel) and phase (lower panel) curves obtained using a FFT length of 512 and a Parzen radius of 0.7 from the same data. Regarding small data set a small FFT length and a big Parzen radius provide an acceptable outcome. The red colored boxes represent XY components and the blue colored boxes represent YX components of apparent resistivity and phase.

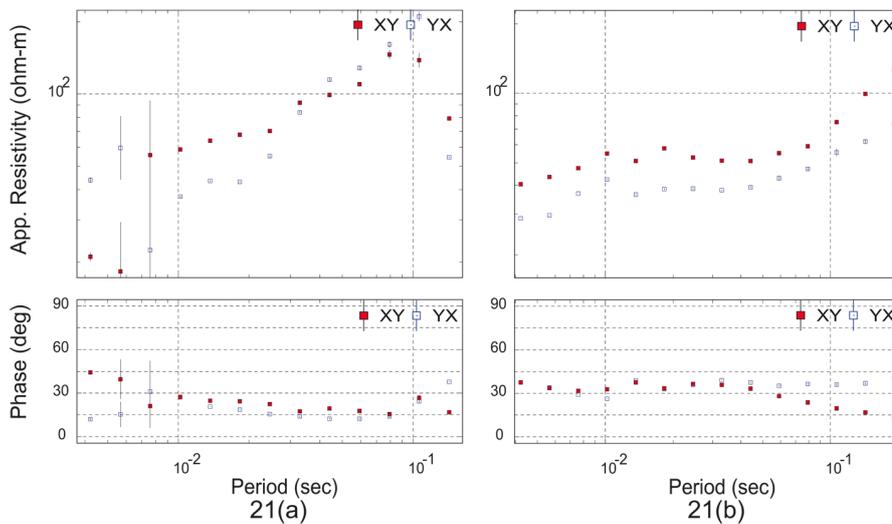


Figure 21. Interrelation between FFT length and Parzen radius in case of a large data set of 1 KHz of site K11 with 1,228,800 data points. (a) Apparent resistivity (upper panel) and phase (lower panel) curves obtained from the data using a FFT length 512 and a Parzen radius 0.6 (b) Apparent resistivity (top panel) and phase (bottom panel) curves obtained applying a FFT length 16384 and a Parzen radius of 0.05 on the same data set. Considering a large data set, a large FFT length and a small Parzen radius provide a better outcome. The red colored boxes indicate XY components and the blue colored boxes indicate YX components of apparent resistivity and phase.

$$\text{mean } (\bar{e}) = (1/n) \sum_{i=0}^n e_i \quad \text{and} \quad (43)$$

$$\text{standard deviation } (\sigma) = \sqrt{[(1/n) \sum_{i=0}^n (e_i - \bar{e})^2]} \quad (44)$$

Now, consider

$$P(i) = \begin{cases} 1 & \text{if } \bar{e} - \sigma \leq e_i \leq \bar{e} + \sigma \\ 0 & \text{otherwise} \end{cases} \quad (45)$$

So, for the final stacked matrix

$$E = \sum_{i=0}^n P(i) \cdot e_i \quad (46)$$

4.7.2. Coherency Threshold

Coherency means similarity or linear dependence between two or more channels and threshold is the lower limit that will produce a particular phenomenon. So, coherency threshold is a method of noise suppression that filters out the coherent portion from different spectral matrices on the basis of a threshold value and stack it i.e. it emphasizes the coherent events.

Suppose, for a particular set of spectral matrices ($s_1, s_2, s_3, \dots, s_n$), threshold value is ' t_h '. Then, consider [Ellinghaus 1997]

$$P(s) = \begin{cases} 1 & \text{if } |\text{coh}(s)| \geq t_h \\ 0 & \text{if } |\text{coh}(s)| < t_h \end{cases} \quad (47)$$

Then, the final stack matrix will be

$$S = \sum_{i=0}^n P(i) \cdot s_i \quad (48)$$

4.7.3. Stack All

In this method all the spectral matrices are stacked together. Hence, similar portion of the data are added up and dissimilar portion of the data are cancelled out.

If we have a set of spectral matrices $m_1, m_2, m_3, \dots, m_n$ then, by Stack All we will have the final stack matrix as [Ellinghaus 1997]

$$M = \sum_{i=0}^n m_i \quad (49)$$

4.7.4. Remote Reference

This method is basically used to suppress the effect of noise in the local input channels. This method uses input data (magnetic field) from a site (denoted as 'remote') located at some distance from the MT site of interest (denoted as 'local') assuming that the input natural field is coherent over spatial scales of many kilometres.

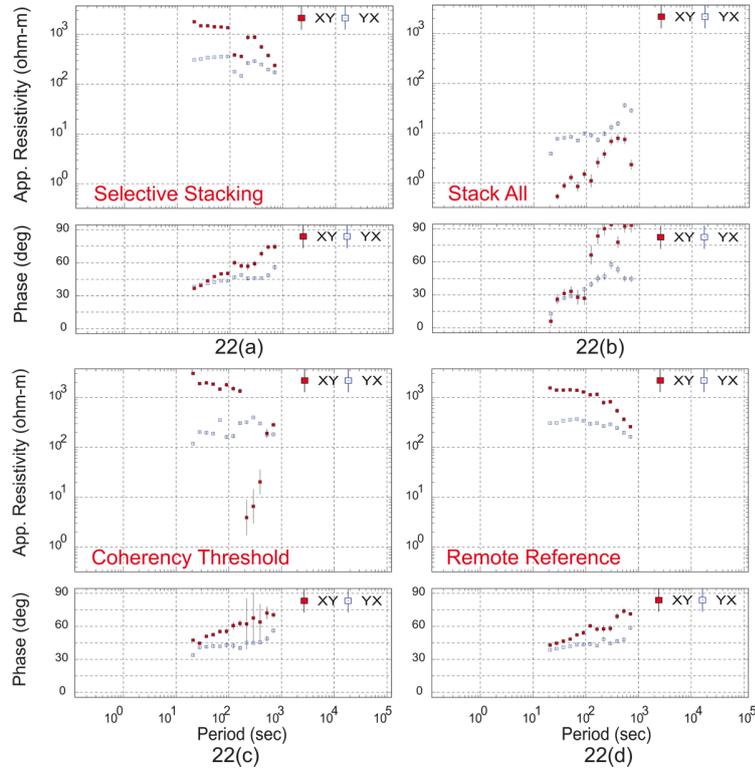


Figure 22. Outcome of 4S sampling data of site K11 from Koyana profile with different processing methods keeping the other parameters constant. (a) Apparent resistivity (upper panel) and phase (lower panel) curves obtained from the data using Selective Stacking. (b) Apparent resistivity (top panel) and phase (bottom panel) curves of the same data obtained from Stack All. (c) Apparent resistivity (top panel) and phase (bottom panel) curves from Coherency Threshold for the same data set. (d) Apparent resistivity (upper panel) and phase (bottom panel) curves obtained from the data applying Remote Reference provide a better result than the others. If remote site is not available then Selective Stacking is more acceptable which is followed by Coherency Threshold (with carefully chosen threshold values) and Stack All. The red colored boxes represent XY components and the blue colored boxes represent YX components of apparent resistivity and phase.

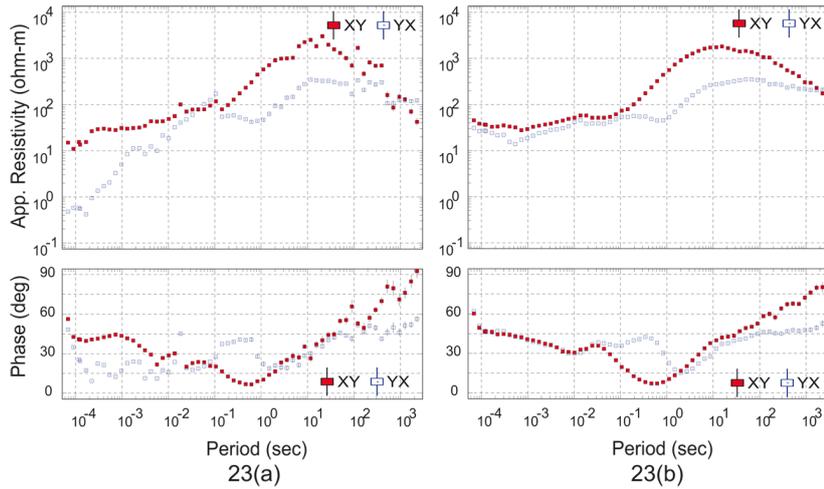


Figure 23. Comparison between processing approaches that are applied on broad band MT data of site K11 from Koyna profile. (a) Result (apparent resistivity (top panel) and phase (bottom panel) curves) obtained from a single processing method (Selective Stacking) with a single set of processing parameters. (b) Result (apparent resistivity (upper panel) and phase (lower panel) curves) obtained from a combination of processing methods with combinations of different sets of processing parameters which indicates a better output while processing culturally affected noisy data. The red colored boxes represent XY components and the blue colored boxes represent YX components of apparent resistivity and phase.

This method removes the biasing effect of random noise [Gamble et al. 1979].

The transfer functions contain auto powers of electric field and magnetic fields. Since any component (Magnetic or electric) is coherent with itself, so, if any noise is present in it then it will be amplified in the auto power which causes transfer function (Z) to be biased. Using Remote Reference technique on magnetic data we can handle this problem.

4.8. Selection of processing method

So far we discussed different processing methods, their physical meaning and mathematical form.

Remote Reference provides a better output in case of disturbed data set than the other methods if a noise free remote site is available. If such remote site is not available then we should proceed with the other processing methods. During our study we have found that if the data contains a small amount of incoherent noise then ‘Selective Stacking’ can be applied. Otherwise it will average out the signal which will carry the noise also. Because as soon as the average value become high, the standard deviation will also go high and allow the noise to come under the signal. In case of aperiodic random noise ‘Stack All’ is better, because it will emphasize the event and diminish or cancel out the noise. If the data is contaminated with heavily disturbed incoherent noise then ‘Coherency Threshold’ will be better with carefully chosen threshold values.

During our study remote site was available only for long period data. So, for the short period data, we have gone through Selective Stacking and Coherency Threshold mainly with proper threshold values, and

partly Stack All on the basis of the nature of the data set. Figure 22 describes the influences of different processing methods on 4S sampling data.

5. Application of new approach to the real data

Figure 23 provides a comparison between the outputs from different processing approaches. Figure 23a shows the output from a single processing method Selective Stacking with single FFT length 1024. Same processing parameters were applied to each frequency band in this case. Figure 23b shows the output from a combination of processing parameters and methods. Here, for windowing Hamming window was applied for entire frequency bands. For the high frequency bands i.e. 64 KHz, 4 KHz and 1 KHz Selective Stacking was used after notch filtering at each frequency band. For the low frequency bands i.e. 32 Hz and beyond Remote Reference technique was used. For each frequency band the required FFT length and Parzen radius were decided on the basis of empirical relations that were discussed earlier. It clearly shows that in case of disturbed data set adoption of systematic approach with different sets of processing parameters and methods provide improved transfer function.

6. Conclusion

To obtain a desirable MT transfer function from time series which is highly contaminated with cultural noise was our prime objective. Remote reference technique can handle this situation until the remote site is free from any kind of noise. During our study we had remote site for long period data only, for short period data we used the data recorded simultaneously from

nearby stations along the profile as remote site. However, as expected we did not get an improved transfer function for the short period data as the nearby stations have the same noise source as that of original station. Though there were remote stations for long period data, at some stations the transfer functions did not improve. Hence, we mainly focused systematically on the processing parameters to improve the S/N ratio. According to our study, not a single method or a single set of processing parameters can give a desirable transfer function in case of culturally affected noisy data. So, we have to follow a systematic approach depending on the data set. Hamming window provides acceptable outcome in case of disturbed data after editing and notch filtering while Blackman window may be the second choice. We should take FFT length as much larger as we can to increase the frequency resolution according to the optimal FFT length criteria. But in case of small data set the required optimal FFT length is small to get a sufficient number of stacks to highlight the signal. A large FFT length and small Parzen radius is more worthy. But in case of small data set, a small FFT length and large Parzen radius is useful. Regarding the processing method Remote Reference provides better output than the others if the remote site is not associated with any kind of noise. If such remote site is not available then Selective Stacking produces a better result followed by Coherency threshold. When applied to the real data the derived transfer function is quite encouraging.

Acknowledgements. We acknowledge to Ministry of Earth Sciences, Govt. of India for funding this study through the project GAP-632-28(NPR) to CSIR-NGRI, Hyderabad. We are thankful to K.C. Reddy, N. Babu, Kashi Raju, Dushyant Rathore and Sivamalliah for participating MT field campaign. A profound sense of gratitude is expressed to Dr. Bernhard Friedrichs for insightful discussion on processing methods. Thanks to Harsh. K. Gupta, Mrinal K. Sen and N. P. Rao for their support and encouragement during the study. Constructive suggestions from both the anonymous reviewers have improved the clarity of the paper.

References

- Cagniard, L. (1953). Basic theory of the magnetotelluric method of geophysical prospecting, *Geophysics*, 18, 605-635.
- Clarke, J., T.D. Gamble, W.M. Goubau, R.H. Koch and R.F. Miracky (1983). Remote reference magnetotellurics: Equipment and procedures, *Geophys. Prospect.*, 31, 149-170.
- Egbert, G.D. (1997). Robust multiple-station magnetotelluric data processing, *Geophys. J. Int.*, 130, 475-496.
- Ellinghaus, A. (1997). PROCMT, Metronix Measurement Instrument and Electronics Ltd., 1-100.
- Fontes, S.L., T. Harinarayana, G.J.K. Dawes and V.R.S. Hutton (1988). Processing of noisy magnetotelluric data using digital filters and additional data selection criteria, *Phys. Earth Planet. In.*, 52, 30-40.
- Friedrichs, B. (2007). MAPROS, Metronix Measurement Instrument and Electronics Ltd., 1-126.
- Gamble, T.D., W.M. Goubau and J. Clarke (1979). Magnetotellurics with a remote magnetic reference, *Geophysics*, 44, 53-68.
- Gupta, H.K., H. Narain, B.K. Rastogi and I. Mohan (1969). A study of the Koyna earthquake of December 10, 1967, *B. Seismol. Soc. Am.*, 59, 1149-1162.
- 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., and D. Rankin (1977). Enhancement of Signal-to-Noise Ratio in Magnetotelluric Data, *Geophysics*, 42, 103-110.
- Nagata, H., H. Mikada, T.N. Goto, J. Takekawa and T. Kasaya (2012). Data processing of magnetotelluric survey data in time domain using digital filter - An Example at the Nankai trough, Extended Abstract, 21st EM Induction Workshop, Darwin, Australia, 1-4.
- Ritter, O., A. Junge and G.J.K. Dawes (1998). New equipment and processing for magnetotelluric remote reference observations, *Geophys. J. Int.*, 132, 535-548.
- Uchida, T., Y. Song, T.J. Lee, Y. Mitsuhashi, S.K. Lim and S.K. Lee (2005). Magnetotelluric Survey in an Extremely Noisy Environment at the Pohang Low-Enthalpy Geothermal Area, Korea, Proceeding World Geothermal Congress, Antalya, Turkey, 1-9.
- Weckmann, U., A. Magunia and O. Ritter (2005). Effective noise separation for magnetotelluric single site data processing using a frequency domain selection scheme, *Geophys. J. Int.*, 161, 635-652.
- Yamane, K., and S. Takasugi (1997). Data Processing Procedures for Minami-Kayabe Magnetotelluric Soundings, *J. Geomag. Geoelectr.*, 49, 1697-1715.

Corresponding author: Ujjal K. Borah,
CSIR- National Geophysical Research Institute, Magnetotellurics
and Deep Resistivity Sounding Division, Hyderabad, India;
email: ujjal.borah6@gmail.com.

© 2015 by the Istituto Nazionale di Geofisica e Vulcanologia. All rights reserved.