

APPLICATION OF CLEAN ALGORITHM IN INCOMPLETE MICROTREMORS RECORDINGS

Hloupis G.^{1,3}, Moisiidi M.¹, Vallianatos F.^{1,2}, Makris J.P.^{1,2}, Stonham J.³ and Triantis D.⁴

¹ *Department of Natural Resources and Environment, Technological Educational Institute of Crete – Romanou 3 – 73133 – Chania, George.Hloupis@brunel.ac.uk, John.Stonham@brunel.ac.uk*

² *Center for Technological Research of Crete, Institute of Natural Resources and Natural Hazards, Greece, fvallian@chania.teicrete.gr, jpmakris@chania.teicrete.gr*

³ *Department of Electronic and Computer Engineering, Brunel University of London, Uxbridge – Middlesex – UB8 3PH – United Kingdom, John.Stonham@brunel.ac.uk*

⁴ *Department of Electronics, Technological Educational Institute of Athens, Greece, Triantis@ee.teiath.gr*

ABSTRACT

The reconstruction of gapped microtremors recordings with the aid of CLEAN algorithm, which originally developed for spectral analysis of unevenly spaced time series, is studied. As proved in previous relative studies the algorithm has the ability to "clean" the spectrum of a discrete signal from artifacts and additionally (through an Inverse DFT) to reconstruct the signal in time domain. We studied the algorithm's effectiveness of reconstruction in microtremors recordings. We saw that for continuous gap with length up to 10% of the total length of microtremor recording, the algorithm was able to reconstruct efficiently.

1 INTRODUCTION

The purpose of this study is to examine the successive estimation of HSVR applied in gapped microtremors recordings. Our approach is based on CLEAN algorithm. This algorithm originally developed for spectral analysis of unevenly spaced time series (Roberts et.al, 1987). Experimental studies (Gitterman et. al. 1996, Bour et.al. 1996, Giampiccolo et. al. 2001) used microtremors recordings in order to determine the response of a site as a precondition for a successful analysis of a local seismic hazard. The majority of the studies involved in situ recordings using three directional seismometers with appropriate data recorders. A problem that arises from in situ recordings is the missing data, which are depicted as gaps in recorded time-series. Common solutions such as interpolation usually produce spectral artifacts, which may be critical in final produced spectrum. As long as the common microtremors analysis techniques (Nogoshi et. al. 1978, Nakamura 1989) take place in frequency domain it is imperative to use a reconstruction technique that leaves the spectrum undominated.

The CLEAN algorithm performs a nonlinear deconvolution in frequency domain equivalent to a least-squares interpolation in the time domain. In this study, we used the CLEAN algorithm as described and modified (Baisch et al. 1999) in order to reconstruct microtremors recordings with continuous gap. As proved in previous relative studies (Ding et.al. 1999, Baisch and Bokelmann 1999) this algorithm has the ability to "clean" the spectrum of a discrete signal from artifacts and additionally (through an Inverse DFT) to reconstruct the signal in time domain. The idea of modified CLEAN algorithm is first to calculate the contaminated spectrum of gapped microtremor recording (which is known that is the convolution of harmonic contents of recording time series with the spectrum of sampling function) and then to find the strongest spectral component (i.e. amplitude and phase in a particular frequency) from this spectrum. Finally this component is subtracted

(including sidelobes) removing in this way its contribution to the spectrum. This process is iterative and every time we produced a residual spectrum "cleaner" than previous. After a predefined number of iterations we finally have a uncontaminated spectrum (which is acquitted from the spectrum of window function).

2 THE CLEAN ALGORITHM: FORMULATION

Let $x(t)$ a signal continuous in time and $D(f)$ its Fourier transform. Consider now the discrete case, where data are given at N times t_i of the above. Its Fourier transform $D(f_i)$, as known, is the convolution of the spectrum $D(f)$ with the spectrum $W(f_i)$ of the spectral window function as shown:

$$D(f_i) = D(f) \otimes W(f_i) \quad (1)$$

For convenience in the following, we named $D(f_i)$ as "contaminated" spectrum and $W(f_i)$ as "window" spectrum.

From equation (1) is obvious that in order to get the true spectrum we must deconvolve the contaminated spectrum from window spectrum preventing weak frequency peaks being masked from strong sidelobes of window spectrum. Especially in our case with the existence of gap in microtremors time series the spectrum $W(f_i)$ becomes more complex.

The CLEAN algorithm is based on a method proposed by Roberts et al. (1987). They utilized the knowledge of sampling function in order to perform an iterative deconvolution of the spectral window in frequency domain.

To illustrate the application of CLEAN algorithm we consider a single frequency cosine signal

$$f(t) = A \cdot \cos(2 \cdot \pi \cdot \hat{f} \cdot t + \theta) \quad (2)$$

For $t \rightarrow \infty$ the spectrum of $f(t)$ is:

$$F(f) = a \cdot \delta(f - \hat{f}) + a^* \cdot \delta(f + \hat{f}) \quad (3)$$

Where \hat{f} denotes peak frequency, $a = \frac{A}{2} e^{i\theta}$ and a^* the complex conjugation of a .

If $f(t)$ sampled at N discrete times the spectrum will be:

$$F(f) = a \cdot W(f - \hat{f}) + W(f + \hat{f}) \quad (4)$$

In equation (4) we have is a "contaminated" spectrum, which contains real peaks at $\pm \hat{f}$ and false peaks, which generated by discretization of the signal as well as the sidelobes of window function $W(f)$.

At peak frequency \hat{f} assuming $W(0) = 1$ we have:

$$F(\hat{f}) = a + W(2\hat{f}) \quad (5)$$

Knowing \hat{f} we can determine a as Ding et al. (1999), calculated:

$$a(\hat{f}) = \frac{F(f) - F^*(\hat{f}) \cdot W(2\hat{f})}{1 - |W(2\hat{f})|^2} \quad (6)$$

The idea of CLEAN algorithm is to use equation (6) to find the first spectral component (i.e. amplitude and phase in a particular frequency) and then subtract it (including sidelobes) from D using equation (4) removing in this way its contribution to the spectrum. This is done by choosing the largest peak in the contaminated spectrum. This procedure is iterative and the iteration scheme as given by Roberts et al (1989) is:

1. Compute the contaminated spectrum $F(f)$
2. Start the iteration with the initial residual spectrum $R^0 \equiv F$

3. On the K'th iteration find the maximum frequency f_{peak} in the previous residual spectrum R^{K-1} and calculate its complex amplitude $a(f_{peak})$ using equation (6).
4. Use equation (4) to calculate the contribution of $a(f_{peak})$ to the contaminated spectrum and form the residual spectrum R^K by subtracting a fraction g ($0 < g < 1$) of the result from R^{K-1} :

$$R^K = R^{K-1} - g \left(\left(a^K \cdot W(f - f_{peak}) + (a^K)^* \cdot W(f + f_{peak}) \right) \right)$$

5. Store the subtracted fraction $g \cdot a^K$ to an uncontaminated component array at locations f_{peak} and $-f_{peak}$
6. After the iteration, convolve the uncontaminated component array with a Gaussian function to obtain reasonable frequency resolution. Finally add the residual spectrum of the last iteration

Using the above iteration scheme we calculated the spectrum of gapped microtremors time series, which is known to be a convolution of harmonic contents of time series with the spectrum of sampling function. At the end of the iterations, we add the uncontaminated spectrum the residual spectrum and we apply an inverse DFT to this sum in order to reconstruct the signal in time domain.

3 APPLICATION TO MICROTREMORS

Microtremors measurements recorded from various sites at Crete (Greece). Rec 1 recorded at old Venezian city of Chania (11/7/2002), Rec 2 recorded at Ancient Aptera (24/1/2004) and Rec 3 recorded at Castello (21/10/2003). The microtremor data used have been recorded with a tridirectional sensor Lennartz 3D-Lite (1Hz period), connected with a 24-bit digital acquisition unit and a personal computer. The sensor has the same characteristics on the three axes. All the equipment, sensor and power supply included, is mounted in a single aluminum case whose base is stiffened by a 1cm thick steel plate. This mounting enables the avoidance of any interference (mechanical or electrical) caused by external wiring. The sampling rate was 125Hz and each recording was 112500 points long provides 15 min of recording. Each time series were corrected for base-line and anomalous trends and then a band pass filter applied with range 0.3-20Hz following the methodology described by Bour et al.(1996) and Mucciarelli (1998). The result from three recordings depicted in fig.1, 2 and 3.

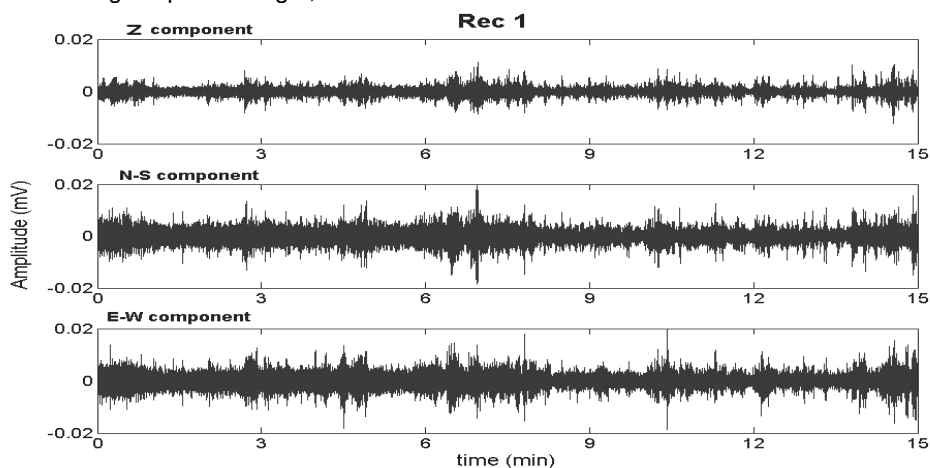


Figure 1: Recording from old Venezian city of Chania, Greece

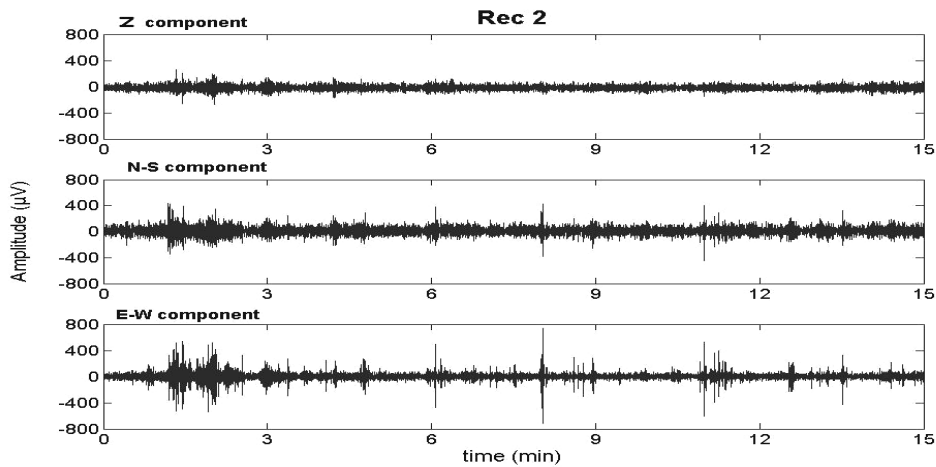


Figure 2: Recording from Ancient Aptera, Greece

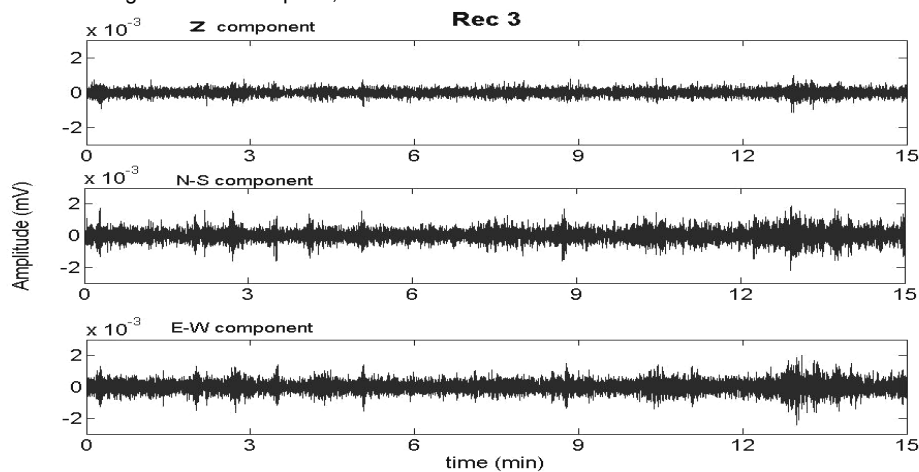


Figure 3: Recording from Castello site, Chania, Greece

The CLEAN algorithm as implemented by Baisch and Bokelmann (1999) requires two crucial parameters: Number of iterations and gain factor. The number of iterations (K) controls how many times the algorithm is repeated until it reaches the final value of residual spectrum. The next parameter, gain factor (g), controls the portion of peak frequency that subtracted from the spectrum because it proved that we couldn't subtract in each iteration the whole frequency peak that we find. Therefore, in every iteration we subtract from contaminated spectrum the $g \cdot a$ portion with $0 < g \leq 1$.

In previous studies, (Roberts et al. 1987, Baisch and Bokelmann 1999, Ding et al. 1999) g usually was 0.1 until 0.5 but these values referred to data with noise up to 31%. We tested the microtremors with these values of g but we did not receive successful reconstructions. The main reason for this seems to be the nature of microtremors, which is pure noise. Based on this, we expect a spectrum with rich frequency content without excessive peaks. In order to choose the appropriate values for g and K parameters we calculate the spectrums for each component for several values of g, K (i.e. Rec 1-NS component spectrums for several values of g and $K=100$, at Fig.4). Then we compare each calculated spectrum with the original one and we define the r^2 coefficient between them. Assuming successive reconstruction for $r^2 > 0.95$ and as plotted at Fig.5 and Fig.6 we chose $g=0.05$ in order to remove small fractions of smeared frequencies in every iteration and $K=100$ iterations.

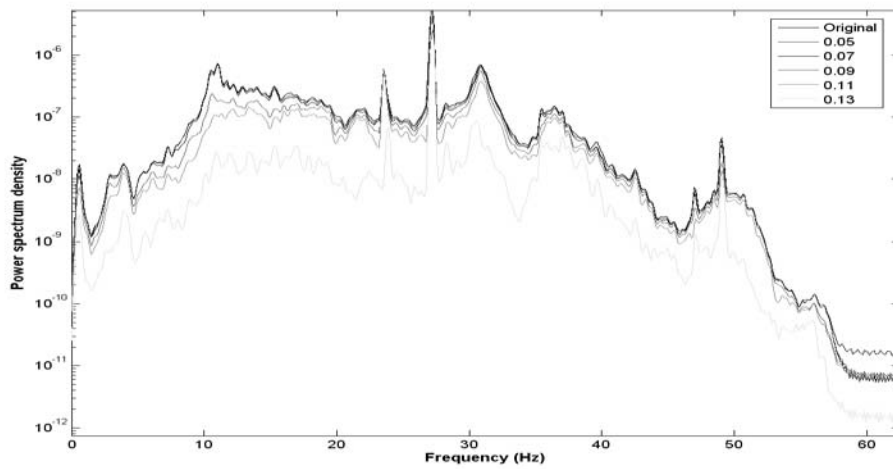


Figure 4: Calculated spectrums for various values of g and $K=100$ (Rec 1 – NS component)

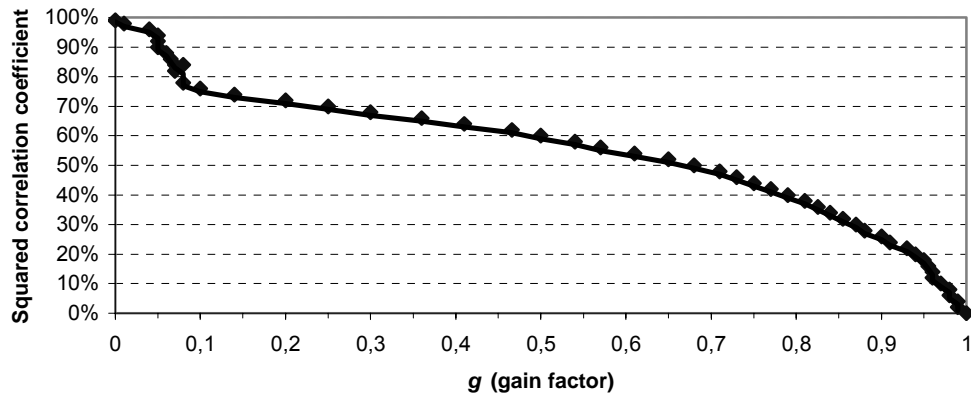


Figure 5: r^2 vs. g parameter for $K=100$ iterations

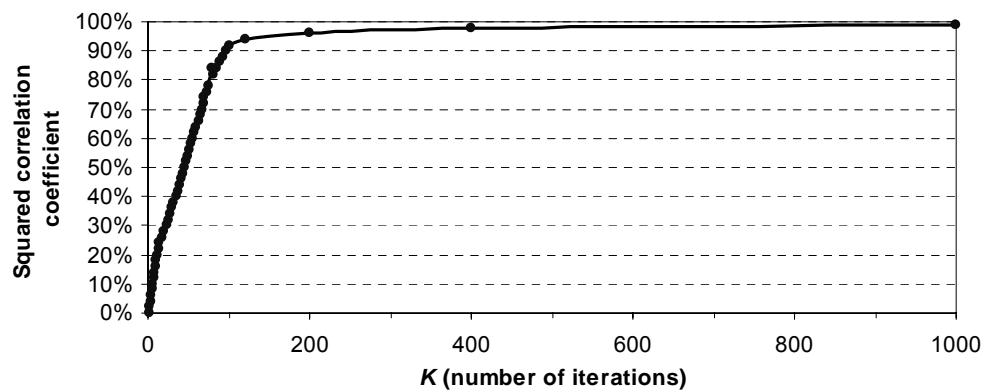


Figure 6: r^2 vs. K parameter for fixed $g=0.005$

For each recording we produce gapped microtremors signals with gaps from 2% to 30% of the total time series length. From each gapped microtremor signal, we produced two reconstructed

signals: One with the aid of CLEAN algorithm and the other with interpolation in time domain. For each triple of the signals (Original, reconstructed by CLEAN, reconstructed by interpolation) we calculate their spectrums (fig.7 - fig 15).

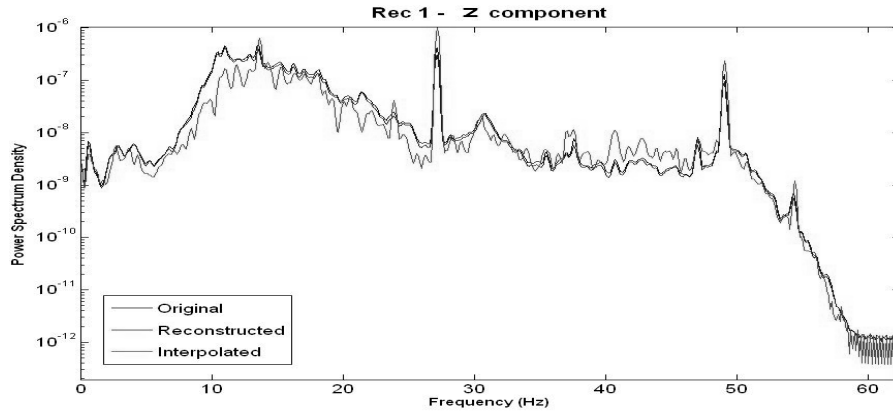


Figure 7: Original, CLEANed reconstructed and interpolated spectrums for 10% gap of Rec 1 (UP)

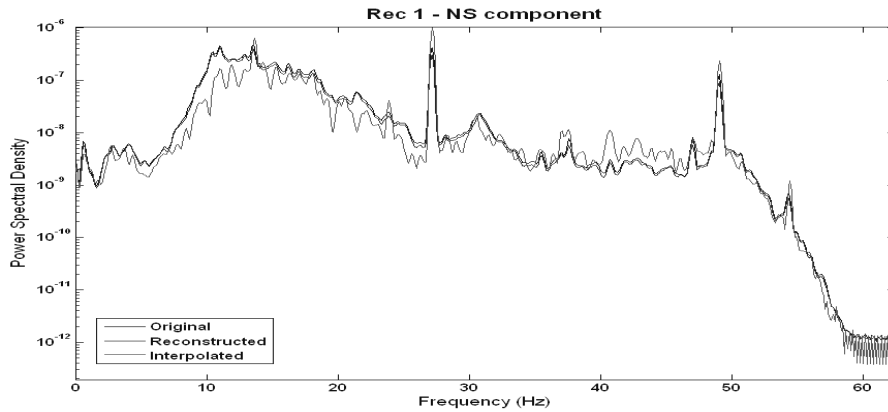


Figure 8: Original, CLEANed reconstructed and interpolated spectrums for 10% gap of Rec 1 (N-S)

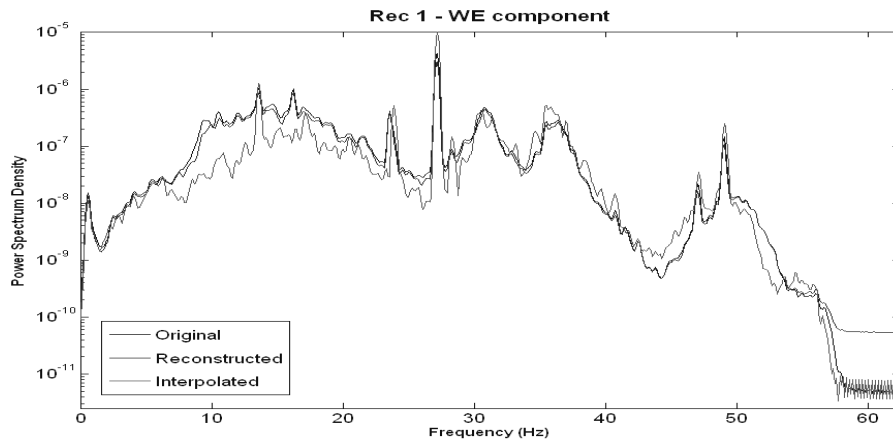


Figure 9: Original, CLEANed reconstructed and interpolated spectrums for 10% gap of Rec 1 (W-E)

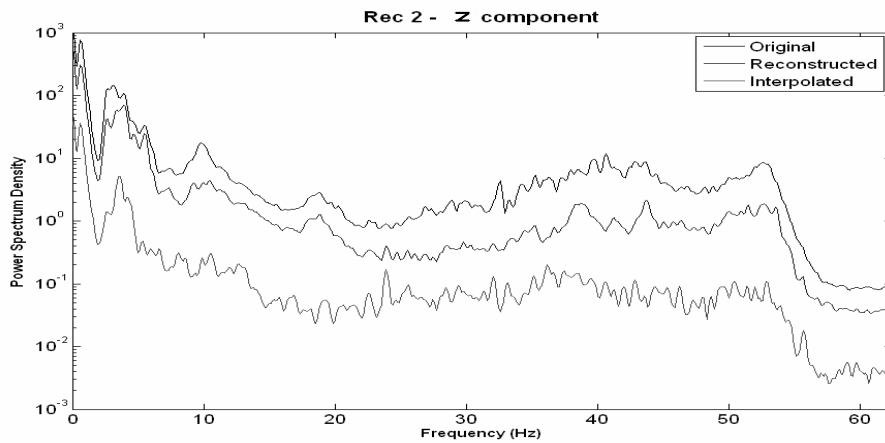


Figure 10: Original, CLEANed reconstructed and interpolated spectrums for 10% gap of Rec 2 (UP)

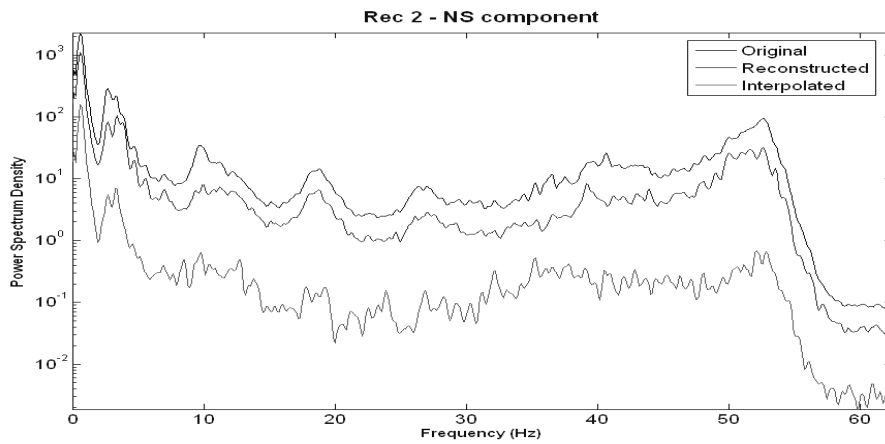


Figure 11: Original, CLEANed reconstructed and interpolated spectrums for 10% gap of Rec 2 (N-S)

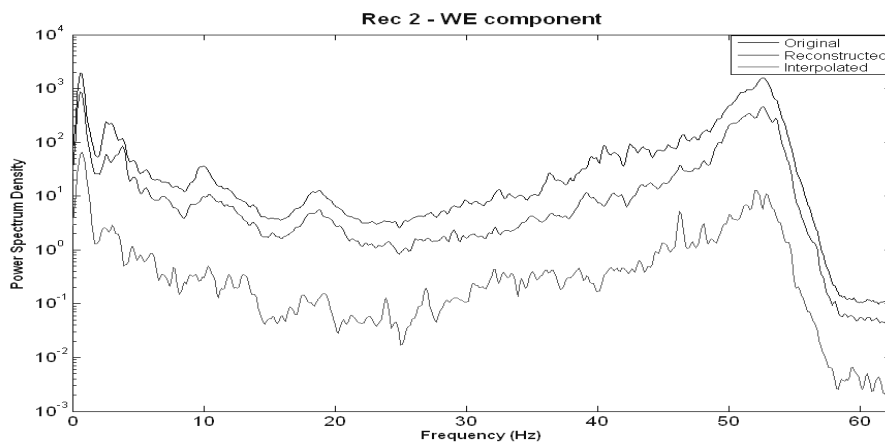


Figure 12: Original, CLEANed reconstructed and interpolated spectrums for 10% gap of Rec 2 (W-E)

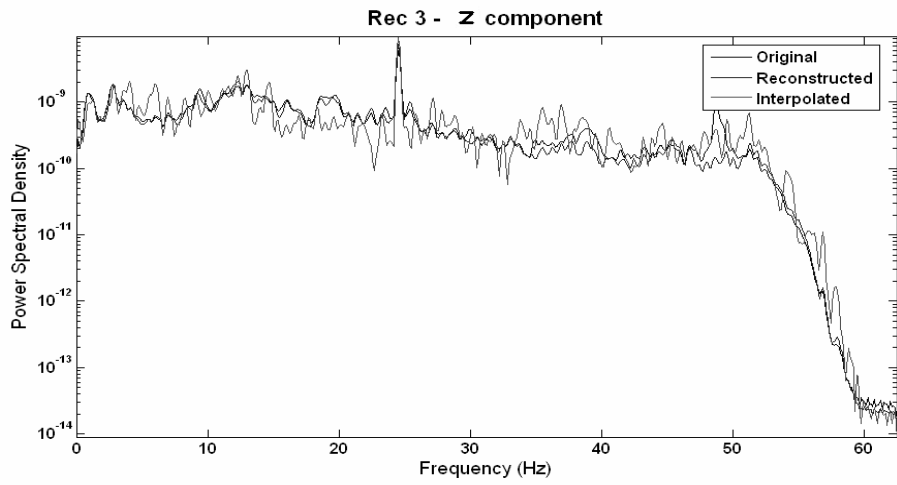


Figure 13: Original, CLEANed reconstructed and interpolated spectrums for 10% gap of Rec 3 (UP)

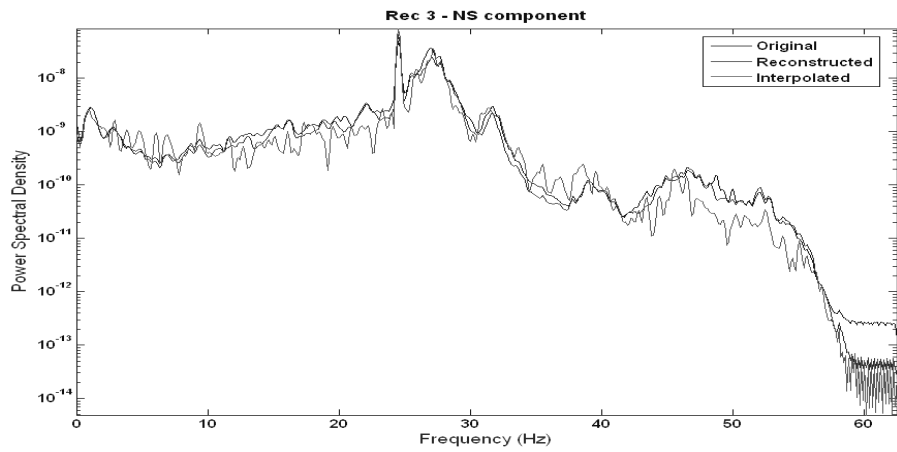


Figure 14: Original, CLEANed reconstructed and interpolated spectrums for 10% gap of Rec 2 (N-S)

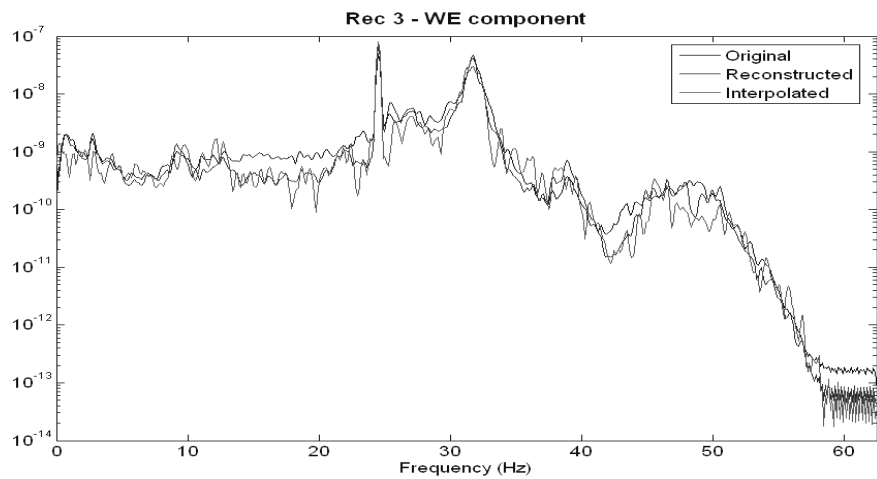


Figure 15: Original, CLEANed reconstructed and interpolated spectrums for 10% gap of Rec 3 (W-E)

In order to investigate the effectiveness of CLEAN algorithm in HVSR calculation we produce HVSR plots using original recording, gapped recording and recording reconstructed by CLEAN for various gaps and for each recording. At fig.16 we present representative HVSR plots for 10% gap for each recording.

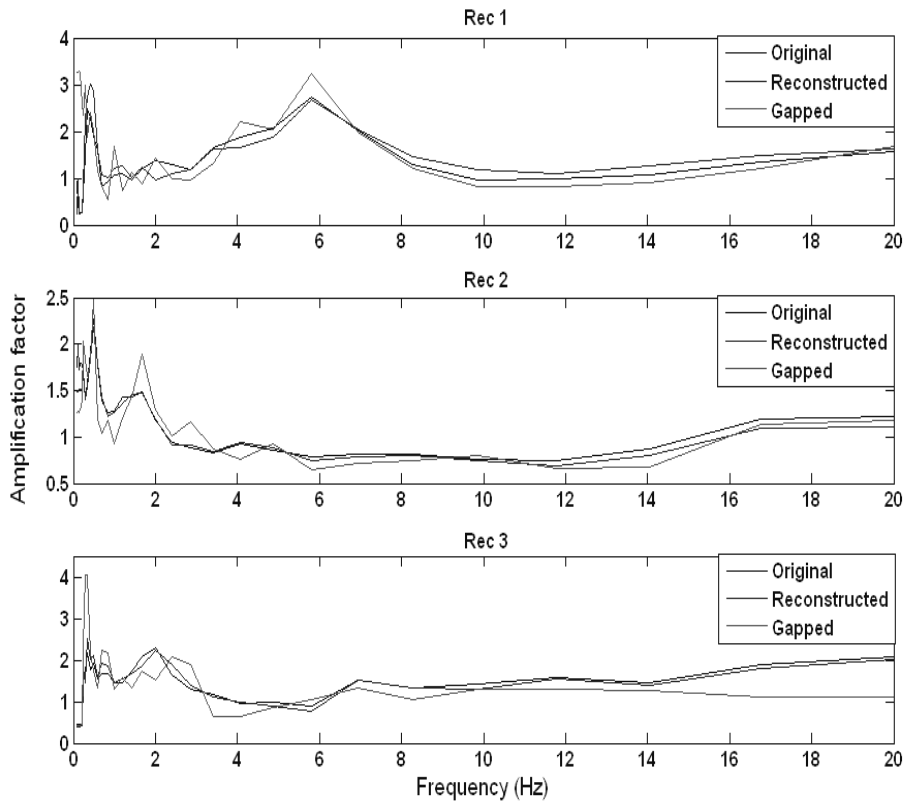


Figure 16: HVSR plots for original, gapped and reconstructed by CLEAN recordings for 10% gap

We can clearly see that the CLEAN algorithm provides better reconstruction than simple interpolation as the calculated HVSR is quite similar than original signal's HVSR. We can also see that as the gap of the missing data increases then the reconstruction became unsuccessful. In order to investigate the exact behavior of CLEAN we calculated the correlation coefficient r^2 between original and reconstructed signal for gap length 2% till 30%, for the microtremors recordings. The results plotted in fig.17 where we can clearly see that if we assume successful reconstruction when $r^2 > 0.95$ then we have successful reconstruction for gap lengths below 10%. We can also state that after 15% gap we have a dramatic failure in reconstruction which shows us that here lied the limit of the algorithm for these type of microtremors signals.

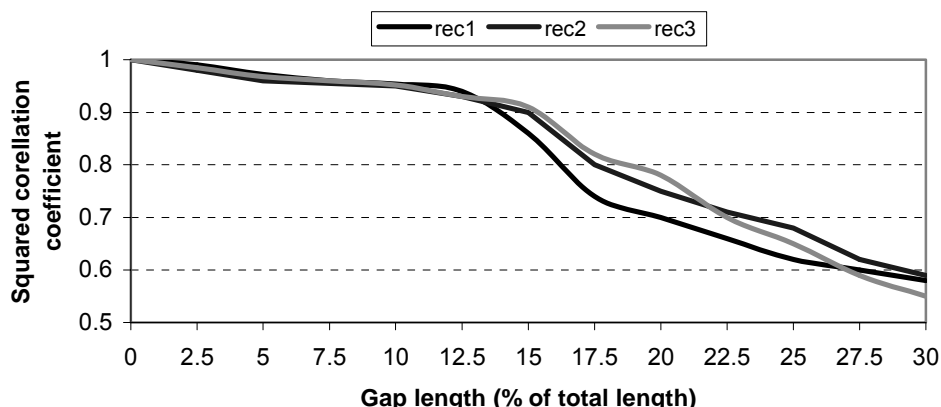


Figure 17: Squared correlation coefficient vs. gap length for the three recording microtremors signals.

4 CONCLUSIONS

The application of CLEAN algorithm for reconstruction of incomplete sampled microtremors signals is studied. Even the algorithm originally developed for spectral analysis we used its ability to uncontaminate the spectrum of a discrete signal and after that with a back transformation to time domain (through an Inverse DFT) to reconstruct the gapped signal from its uncontaminated spectrum.

We applied the algorithm in five gapped microtremors recordings and we examined the successive reconstruction in relation with gap length. We show that for continuous gap lengths below 10% of total length of the signal. Additionally we see that for performing the best results with CLEAN algorithm we must know the nature of our signal in order to select the appropriate value of parameter g .

ACKNOWLEDGEMENTS

Research of the author GH is funded by Greek National Foundation for Scholarships (IKY). The work of FV, JM and MM was supported by "ARCHIMEDES" project "Seismic Hazard in the frontal part of the Hellenic Arc", Greek Ministry of Education and CRINNO project, Region of Crete

REFERENCES

- Baisch, S., Bokelmann, G., 1999, "Spectral analysis with incomplete time series: an example from seismology", *Computer and Geosciences*, 25, pp.739-750
- Bour, M., Fouissac, D., Dominique, P., Martin, C., 1996, On the use of microtremor recording in seismic microzonation. *Soil Dynamics and Earthquake Engineering*, vol.17, issue 7-8, pp.465-474
- Ding, Y., Li, Z., Diwu, Y., 1999, The CLEAN method of spectral analysis and its application to the oscillations of the cataclysmic variable TT Ariets", *Chinese Astronomy and Astrophysics*, 23, pp.484-492
- Giampiccolo, E., Gresta, S., Mucciarelli, M., de Guidi, G., Gallipoli, M., 2001, Information on subsoil geological structure in the city of Catania from microtremors measurements. *Annali di Geofisica*, Vol.44, N.1
- Gitterman, Y., Zaslavsky, Y., Shapira, A., Shtivelman, V., 1996, Empirical site response evaluations: case studies in Israel. *Soil Dynamics and Earthquake Engineering*, vol.15, pp.447-463
- Mucciarelli, M., 1998, Reliability and applicability of Nakamura's technique using microtremors: an experimental approach, *Journal of Earthquake Engineering*, Vol.2, No.4, pp.625-638.
- Nakamura, Y., 1989, A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface, *Rep. Railway Tech. Inst. Jpn.*, 30, 1, pp.25-33.
- Nogoshi, M., 1978, On fundamental nature of microtremors and its application. *J.Min. Coll. Akita Univ. Jpn*, Ser.A 5 3, pp.1-51
- Roberts, D., Lehar, J., Dreher, J., 1987, Time series analysis with CLEAN. I: Derivation of a spectrum, *Astronomical Journal*, 93 (4), pp.968-989