Geophysical parameter estimation

- Application to denoising of seismological data

A Thesis

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Bhupendra Charan
20131140

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April, 2018

Supervisor: Prof. Shyam S. Rai

Certificate

This is to certify that this dissertation entitled 'Geophysical Parameter estimation – Application to denoising of seismological data' towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Bhupendra Charan at Indian Institute of Science Education and Research under the supervision of Shyam S. Rai, Professor, Department of Earth and Climate Science, during the academic year 2017-2018.

Prof. Shyam S. Rai

Committee:

Prof. Shyam S. Rai

Prof. Pravin K Gupta

This thesis is dedicated to my mentor
"Prof. Shyam S. Rai"

Declaration

I hereby declare that the matter embodied in the report entitled ‘Geophysical Parameter estimation – Application to denoising of seismological data’ are the results of the work carried out by me at the Department of Earth and Climate Science, Indian Institute of Science Education and Research, Pune, under the supervision of Prof. Shyam S. Rai and the same has not been submitted elsewhere for any other degree.

Bhupendra Charan

Acknowledgements

I would like to thank Prof. Shyam S. Rai for introducing me to various aspects of Inverse Theory and supporting me throughout the year in my fifth year Project. I would also like to thank my TAC member Prof Pravin K. Gupta for his guidance on this project. It wouldn’t have been possible without their constant help and support.

I am thankful to the “Earth and Climate Science Department” for making available all the necessary resources that were required in the thesis.

Abstract

Geophysical data modelling involves parameter estimation of the modelled system using mathematical relationship describing the physical process. Most of these relationships are inherently non- linear and requires solving them through a process of linearization or using any of the nonlinear search algorithm. Estimating model parameter from the geophysical data is not only unique, but also dependent on the initial model. Apart from these, the data error adds to the parameter estimation complexity.

Some of these issues have been addresses through robust statistics incorporating apriori information related to data and model covariance through their probability density function and search through global optimisation.

In this thesis, we look at the various denoising algorithm and implement the best approach to model gravitational field data from India in terms of Earth parameters.

Using this approach, we can retrieve the original signal from a noised signal upto a great extent with a decent signal to noise ratio.

Content

Abstract xi

1 Introduction 1

2 Preliminaries 4

3 Methods for Denoising 11

3.1 Singular Value Decomposition …………………………………………… 11

3.2 Least Square Fit ……………………………………………………………... 12

3.3 Digital filter ………………………………………………………………….. 13

3.4 Random Quick Response Denoising (rQRD) …………………………….. 14

3.5 Uncoil Random Quick Response Denoising (urQRD) …………………… 15

4 Application using seismological data 18

5 Results and discussions 32

6 Conclusion 34

Chapter 1

Introduction

The main focus of this thesis is the implementation of inverse theory to estimate geophysical parameters and analysis of various time series using inverse problems.

In 1929, The concept of inverse problems was first introduced and discovered by physicist "Viktor Ambartsumian".

An inverse problem can be structured as:

Data → Model Parameters

The inverse problem is the "inverse" to the forward problem which relates the model parameters to data:

Model Parameters → Data

The most common method to solve inverse problems is least square methods.

The purpose of Geophysical inversion is to look for models which explains geophysical observations. The solution of non-linear models are carried out by global optimisation. We have tried various global optimisation algorithms to find the "one" which can optimise the large data sets effectively and requires less time.

One can get a clear understanding of the context and a detailed study on inverse theory and its applications in the books titled *'Global optimisation methods in geophysical inversion'* by 'Sen M.K. and Stoffa P.L.' and *'Time series analysis and Inverse theory for geophysicists'* by 'David Gubbins'.

Inverse theory is applicable to various fields of geophysics including seismic attenuation, determination of velocity structure of the earth, ocean circulation, earthquake location, signal correlation etc.

Our topic of concern is, 'Signal correlation', in other words; to observe the similarities between the original signal and the signal which obtain after denoising.

During the Progress of thesis, we will study various algorithms used for denoising the signal which will give us an insight about their advantages and limitations. So that we would be able to consider the appropriate algorithm according to nature of different signals.

The seismological data which is used in this project to test various algorithms, is taken from the IRIS (Incorporated Research Institutions for Seismology).

The central goal of the project is to contribute an effective approach to handle and remove the noise which gets developed by certain factors viz. background data, measurement error, spurious readings that can corrupt the seismological data which might result in loss of information.

Chapter 2

Preliminaries

Some of the important definitions and concepts used repetitively in the Project:

1. Relation between model parameter and data:

Data: $d = [d_1, d_2, d_3, d_4, \ldots\ d_N]^T$

Model Parameters: $m = [m_1, m_2, m_3, m_4, \ldots\ m_M]^T$

$$Gm = d$$

Where, G = data kernel

A standard linear inverse problem is often represented by the above mentioned equation.

Model parameters & data have analogs which are the continuous functions represented as $m(x)$ & $d(x)$, where x is an independent variable. ‘Continuous inverse theory’ exists between these two ends having a continuous model function and discrete data.

Different representations of this relation:

(i) ‘Continuous inverse theory’:

$$d_i = \int G_i(x)\ m(x)\ dx$$

(ii) ‘Discrete inverse theory’:

$$d_i = \sum_{j=1}^{M} G_{ij} m_j$$

(iii) ‘Integral equation theory’:

$$d(y) = \int G(y,x)\ m(x)\ dx$$

The applicability of these theories depend upon whether the data "d" and model parameter "m" are discrete parameters or continuous functions.

Least square solution of the linear inverse problem:

Consider the linear inverse problem:

$$d = Gm$$

$$d - Gm = E$$

$$(d - Gm)^T(d - Gm) = e^T e = E$$

If G is not a square matrix then we take,

$$m^{est} = [G^T G]^{-1} G^T\ d$$

2. Well posed problem:

This term was given by a mathematician, ‘Jacques Hadamard’. According to him, a system of equation is said to be well posed if it satisfies the following properties:

(i) Existence of solution,

(ii) Uniqueness of solution,

(iii) Behaviour of the solution alters continuously with initial conditions.

If any of the condition is violated then the problem is said to be “ill posed”.

3. Singular value decomposition:

Assume A is a $m \times n$ matrix whose entries comes from the field of real numbers or complex numbers, let's call it K, Then there exists a factorization which is defined as a 'singular value decomposition' of A having the form as follows:

$$A = U\Sigma V^*$$

Where,

U is an $m \times m$ unitary matrix (if $K = \mathbb{R}$, unitary matrices are orthogonal matrices).

Σ is a diagonal $m \times n$ matrix with non – negative real numbers on the diagonal.

V is an $n \times n$ unitary matrix over K, and

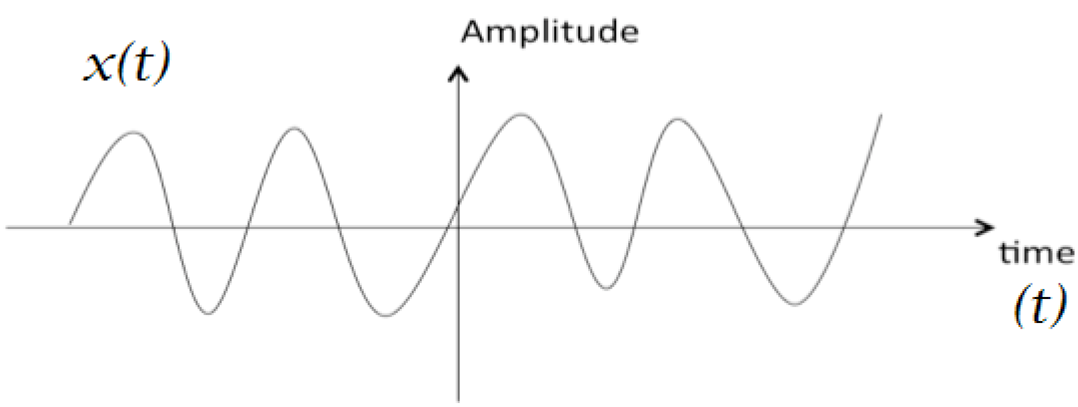
V^* is the conjugate transpose of V.

The diagonal entries σ_i of Σ are known as the singular values of A. The diagonal matrix, Σ, is uniquely determined by A.

4. Signals and systems:

(i) Signals:

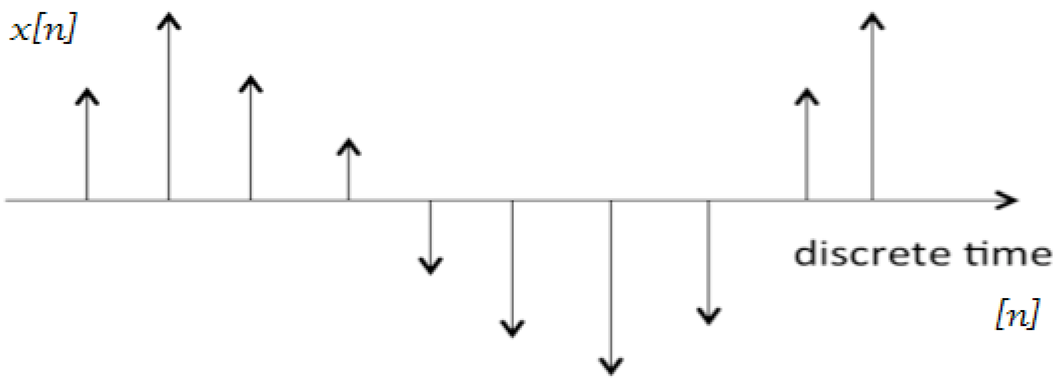
(a) Continuous time signal (speech signal) :



The signals which are defined for every instants of time are termed as continuous time signal.

Eg. $f(t) = \sin t$

(b) Discrete time signal:



The signals which are defined at only certain instants of time or discrete instants of time are termed as discrete time signal.

Eg. Stock market index

- Multi-dimensional signal

$$X[n, m]$$

Where, n - array no. in horizontal direction

m – array no. in vertical direction.

(ii) Systems:

Systems processes signals by producing an output signal in response to an input signal.

→ For continuous time:

$$x(t) \longrightarrow \boxed{System} \longrightarrow y(t)$$

→ For discrete time:

$$x[n] \longrightarrow \boxed{System} \longrightarrow y[n]$$

Types of systems:

- Linear $\longrightarrow$ Time invariant

- Non – Linear ⟶ Time – varying

Domains for analysis and representation:

- *Time domain:*

→ $X(t)$

→ $X[n]$

- *Frequency domain:*

→ Fourier transform

→ Laplace transform

→ z - transform

Continuous – time sinusoidal signal:

$$X(t) = A\cos(\omega_0 t + \phi)$$

Where, A – Amplitude

ω_0 – Frequency

ϕ – Phase

Periodic:

$$X(t) = x(t + T_0)$$

$$A\cos(\omega_0 t + \phi) = A\cos(\omega_0 t + \omega_0 T_0 + \phi)$$

$$T_0 = \frac{2\pi m}{\omega_0}$$

$$Period = \frac{2\pi}{\omega_0}$$

Unit impulse function: Discrete time

$$\delta[n] = u[n] - u[n-1]$$

Where, $\delta[n]$ – unit impulse

$u[n]$ – unit step

Unit step function: Continuous time

$$u(t) = \begin{cases} 0, & t < 0 \\ 1, & t > 0 \end{cases}$$

∗ Note – Impulse function is derivative of step function:

$$\delta(t) = \frac{du(t)}{dt}$$

$$\delta\Delta(t) = \frac{du_{\Delta}(t)}{dt}$$

$$\delta(t) = \delta\Delta(t) \; as \; \Delta \rightarrow 0$$

5. Convolution Sum:

$$\left\{ y[n] = \sum_{k=-\infty}^{+\infty} x[k]\, h[n-k] \right\} = x[n] * h[n]$$

6. Convolution integral:

$$\left\{ y(t) = \int_{-\infty}^{+\infty} x(\tau) h(t-\tau) d\tau \right\} = x(t) * h(t)$$

Where,

$$x(\tau) = input$$

$$h(t-\tau) = impulse\ response$$

7. Properties of convolution:

(i) Commutative:

$$x[n] * h[n] = h[n] * x[n]$$

$$x[t] * h[t] = h[t] * x[t]$$

(ii) Associative:

$$x * \{h1 * h2\} = \{x * h1\} * h2$$

(iii) Distributive:

$$x * \{h1 + h2\} = x * h1 + x * h2$$

∗ Note – I have referred [2], [3], [4], [9], [10], [11] for definitions mentioned in preliminaries.

Chapter 3

Methods of Denoising

Numerous denoising methods have been proposed which are used according to the nature of signal or data sets. We will give a glance at some of the algorithms which I came across during my project.

3.1 Singular Value Decomposition (SVD)

In the preliminaries, I've already given an insight about the singular value decomposition and its computation.

Methodology:

Let the original signal be $x(n)$ and the noise signal be $y(n)$.

Take a portion of noise signal of N length

Consider $h(n) = x(n) + y(n)$

Where, $n = 0,1,2, \dots , N-1$

Now, write it in the form of Hankel matrix of order $(N - M + 1) \times M$

$$H = \begin{bmatrix} h(0) & \cdots & h(N-M) \\ \vdots & \ddots & \vdots \\ h(M-1) & \cdots & h(N-1) \end{bmatrix}$$

We can apply singular value decomposition on the H matrix.

$$H = U\Sigma V^*$$

Where,

U is an orthogonal matrix of order $(N - M + 1) \times M$

V is an orthogonal matrix of order $M \times M$

Σ is diagonal matrix of order $M \times M$

The entries of Σ are $\Sigma_0 \geq \Sigma_1 \geq \dots \geq \Sigma_{M-1}$

Now, construct $\tilde{\Sigma}$ by removing the lower power (noise portion of the signal) or keeping only the largest K entries of Σ

Define:

$$\tilde{X} = U\tilde{\Sigma}V^*$$

It can be considered as the 'Hankel matrix' of the estimated signal $\hat{x}(n)$

On taking average of the entries of $\tilde{X}$ over the diagonal, we get

$$\hat{x}(j) = \sum_{i=0}^{j} \tilde{X}_{i,j-1}$$

Limitations:

(i) This method works on the assumption that original signal $x(n)$ and noise signal $y(n)$ should have low 'cross correlation'.

(ii) The noise added to the signal should be 'white Gaussian noise'.

(iii) It doesn't work on large datasets as the size of the matrix (which is supposed to be decomposed using SVD) goes to square of data length.

(iv) It is relatively slow with order (mn^2) .

3.2 Least Square Fit

It is the most common method used in regression analysis. The main application of least square method is to minimize the sum of squared residuals in order to achieve the best fit in data fitting.

Let the noisy signal be $x(n)$ and $y(n)$ be the desired signal after denoising.

We need to minimise:

$$min_x \; \|x - y\|_2^2 + \lambda \, \|Dy\|_2^2$$

The first order difference for $y(n)$ is:

$$x(n) = y(n) - y(n-1)$$

After taking the differential of the above equation, we get the second order differential of this signal which is given by,

$$x(n) = y(n-1) - 2y(n) + y(n-1)$$

S is a second order differential square matrix that can be defined using the above equation.

Similarity between $y(n)$ and $x(n)$ is attained when the term $\|x - y\|_2^2$ is minimized and smoothness of $y(n)$ is attained when the term is $\|Dy\|_2^2$ is minimised.

Smoothness depends on the controlling parameter, $\lambda > 0$

At $\lambda = 0$,

$$x(n) = y(n)$$

For higher values of λ, the signal will be denoised more efficiently.

The expression for denoising of signal using 'Least square' is given by,

$$y = (I + \lambda \, S^T S)^{-1} \, x$$

Where, size of $I(Identity\ matrix) = S$

Limitations:

(i) This method is highly sensitive towards outliers. Few outliers in the data sets can highly affect the result obtained by this analysis.

(ii) With the increase in data ranges, it becomes complicated to get a linear model for inherently nonlinear processes that fits the data effectively.

(iii) If the data is not normally distributed then test statistics can come up as erratic.

3.3 Digital filter

Consider an input signal x and output signal (denoised signal) y.

It filters or smoothens the input signal using a rational transfer function in $Z-$ Transform domain which is defined by,

$$y(z) = \frac{b(1) + b(2)z^{-1} + \; ... + b(n_b + 1)z^{-n_b}X(z)}{1 + a(2)z^{-1} + \; ... + a(n_a + 1)\, z^{-n_a}}$$

Where,

$n_a -$ feedback filter order

$n_b -$ feedforward filter order

$b -$ Numerator Coefficient

$a -$ Denominator Coefficient

After getting the numerator and denominator coefficient, the denoised signal can be obtained by using the following syntax in MATLAB :

$$y = filter(b, a, x)$$

It worked efficiently on the seismological data, I have used for removing the noise as compare to other mentioned methods. It is also relatively easier to apply and smoothens the large data sets in less time.

3.4 rQRD (Random Quick response Denoising)

Algorithm *(as proposed by L. Chiron et al)*:

Consider a time series T with rank P and order M that returns $\tilde{T}$ a denoised approximation of T.

Require: $T, P, M \;\; P \leq M \; \leq length(T)/2$

Require: Function $RANDOM: n, p \mapsto \; \Omega$ → Ω a $\mathcal{N}(0,1)\; n \times p \; matrix$

Require: Function $QR: A \mapsto Q, R$ → the QR decomposition of A

$L \leftarrow LENGTH(T)$

$N \leftarrow L - M + 1$

for $i \leftarrow 1, M \quad j \leftarrow 1, N$ do

$H_{ij} \leftarrow T_{i+j-1}$ → H is a $M \times N$ matrix

end for

$\Omega \leftarrow \; \mathrm{RANDOM(N, P)}$

$Y \leftarrow H\Omega$

$\mathrm{(Q, R) \leftarrow QR(Y)}$

$\widetilde{H} \leftarrow \mathrm{Q}Q^{*}\mathrm{H}$

for $l \leftarrow 1, L$ do

$\widetilde{T_l} \leftarrow \langle H_{ij} \rangle_{i+j=l+1}$

end for

return $\tilde{T}$

The matrices H and $\widetilde{H}$ are the largest objects stored in memory. This represents a memory burden proportional to $O(MN) \leq O(L^2)$.

The slowest step is the computation of $\widetilde{H} = QQ^{*}H$ in $O(PMN)$ while the computation of $\tilde{T}$ is in $O(LM)$. This results in a theoretical time dependence in $O(PMN + LM)$.

3.5 urQRD (Uncoil Random Quick Response denoising)

Algorithm *(as proposed by L. Chiron et al)*:

Consider a time series T with rank P and order M that returns $\tilde{T}$ a denoised approximation of T.

Require: $T, P, M \;\; P \leq M \leq length(T)/2$

Require: Function $RANDOM: n, p \mapsto \Omega$ → Ω a $\mathcal{N}(0,1)$ $n \times p$ $matrix$

Require: Function $QR: A \mapsto Q, R$ → the QR decomposition of A

Require: Function $F_{HV}: H, M, T \mapsto Y$

Require: Function $F_{HM}: H, M, A \mapsto B$

$L \leftarrow LENGTH(T)$

$N \leftarrow L - M + 1$

$\Omega \leftarrow RANDOM(N, P)$

$Y \leftarrow F_{HM}(X, \Omega)$

$(Q, R) \leftarrow QR(Y)$

$U \leftarrow [F_{HM}(T, Q^{*})]^{*}$

for $p \leftarrow 1, P$ do

$Q^{(p)} \leftarrow \{Q_{1,p}, \dots, Q_{M,P}\}$

$U'^{(p)} \leftarrow \{U_{p,N}, U_{P,N-1} \dots, U_{P,1}\}$

$$W^{(p)} \leftarrow \{\underbrace{0,\dots,0}_{N-1 \text{ values}}, Q_1^{(p)}, \dots, Q_M^{(p)}, \underbrace{0,\dots,0}_{N-1 \text{ values}}\}$$

$$Z^{(p)} \leftarrow \{F_{HV}(W^{(p)}, U'^{(p)}\}$$

end for

$$Z \leftarrow \sum_{k=1}^{K} Z^{(k)}$$

for $1 \leftarrow 1, L$ do

$$\tilde{T}_l \leftarrow \alpha_l Z_l \ with \ \alpha_l = \begin{cases} 1/l & 1 \le l \le M \\ 1/M & M < l < N \\ \frac{1}{L-l+1} & N \le l \le L \end{cases}$$

end for

return $\tilde{T}$

The matrices Y, Q and U are the largest objects stored in memory. This represents a memory burden proportional to $O(PL)$.

The slowest step is the loop on P for the computation of $\tilde{T}$ and its processing time is proportional to $O(PL\log(L))$.

The algorithms ‘rQRD’ and ‘urQRD’ differs only on the grounds of processing speed and implementation but they are similar in terms of results. Due to additional complexity, ‘urQRD’ is slower for small datasets but ‘urQRD’ easily works on less computer memory footprint and can be applied to a much larger datasets which makes this algorithm more robust and fast.

Chapter 4

Applications using Seismological data

I have applied the denoising algorithms after adding noise on the seismological data taken from IRIS to obtain a better signal to noise ratio.

Signal to Noise Ratio is defined as,

$$snr = 10 \times \log_{10}\left(\frac{P_{signal}}{P_{noise}}\right)$$

Where,

$P_{signal} = \text{Power of signal} = \text{rms(signal)}^2$

$P_{noise} = \text{Power of noise} = \text{rms(noise)}^2$

All simulations were carried out in MATLAB.

The filter I used for this purpose is Digital filter because of its reliability over large datasets.

Computation:

1. Station: L01_2014.082.18.20.01.e

s = readsac('L01_2014.082.18.20.01.e');

→ Open the sac file

z = awgn(s.DATA1,4,'measured');

→ Adding noise

window size = 3;

b = (1/window size)*ones(1,window size);

a = 1;

denoised signal = filter(b,a,y);

→ Removal of noise using moving average filter

Moving average filter works on filter function which is defined by,

$$y(n) = \frac{1}{window\ size}(x(n) + x(n-1) + \ ... + x(n-(window\ size - 1)))$$

Where,

$y = output$

$x = input$

By assigning the window size, we evaluate the ‘numerator coefficient’ and ‘denominator coefficient’ of ‘Rational Transfer Function’.

figure(1)

subplot(4,1,1),plot(s.DATA1),xlim([4*10^4,5*10^4])

title('Original Signal from station L01 2014.082.18.20.01.e')

→ Plotting the original signal

subplot(4,1,2),plot([s.DATA1 z]),xlim([4*10^4,5*10^4])

title('Noisy Signal')

→ Plotting the Noisy signal

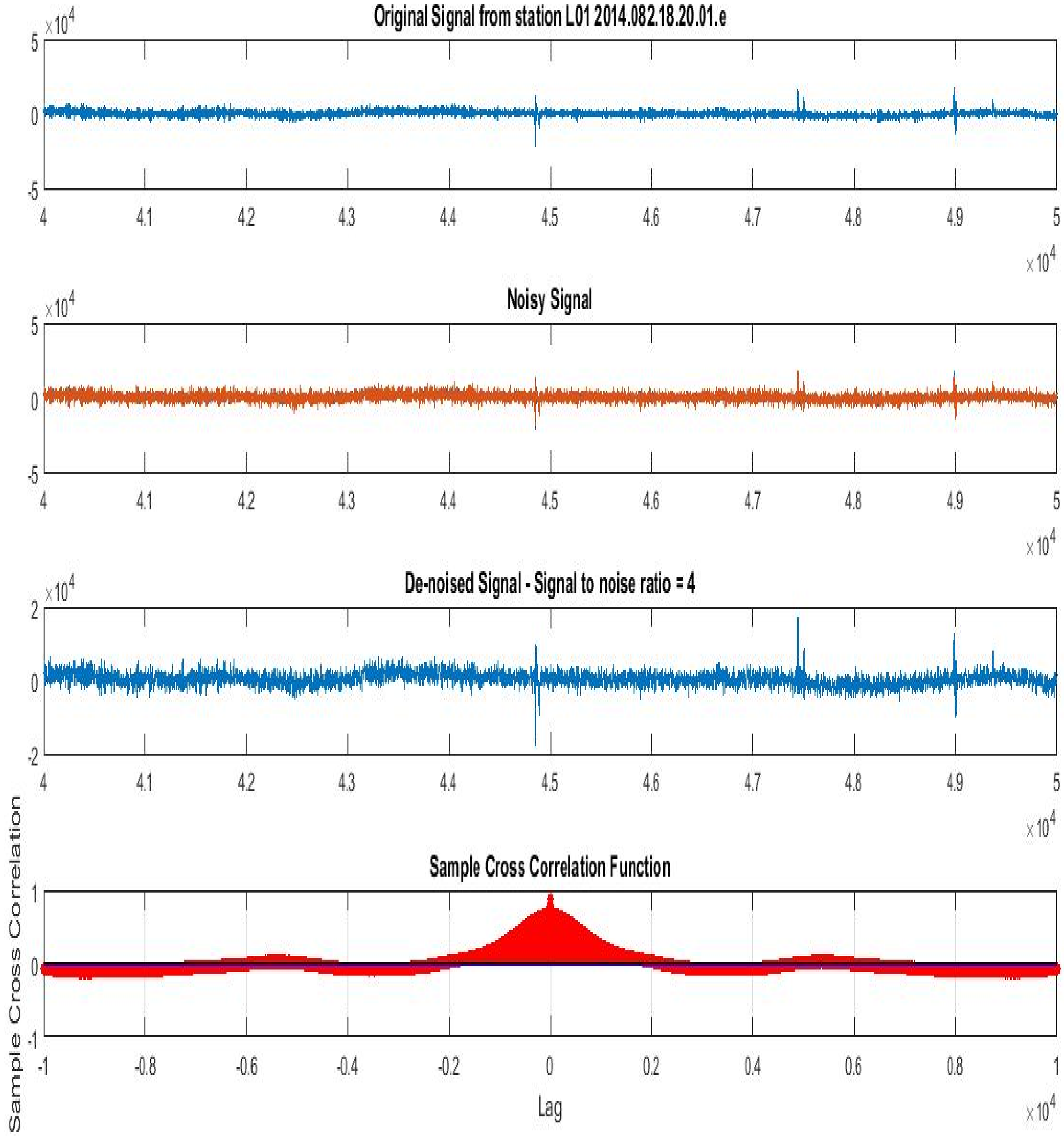
subplot(4,1,3),plot(denoised signal),xlim([4*10^4,5*10^4])

title('De-noised Signal - Signal to noise ratio = 4')

→ Plotting the Denoised signal

subplot(4,1,4),crosscorr(s.DATA1,denoised signal, 1*10^4)

→ Plotting the Cross Correlation between Original signal and denoised signal.



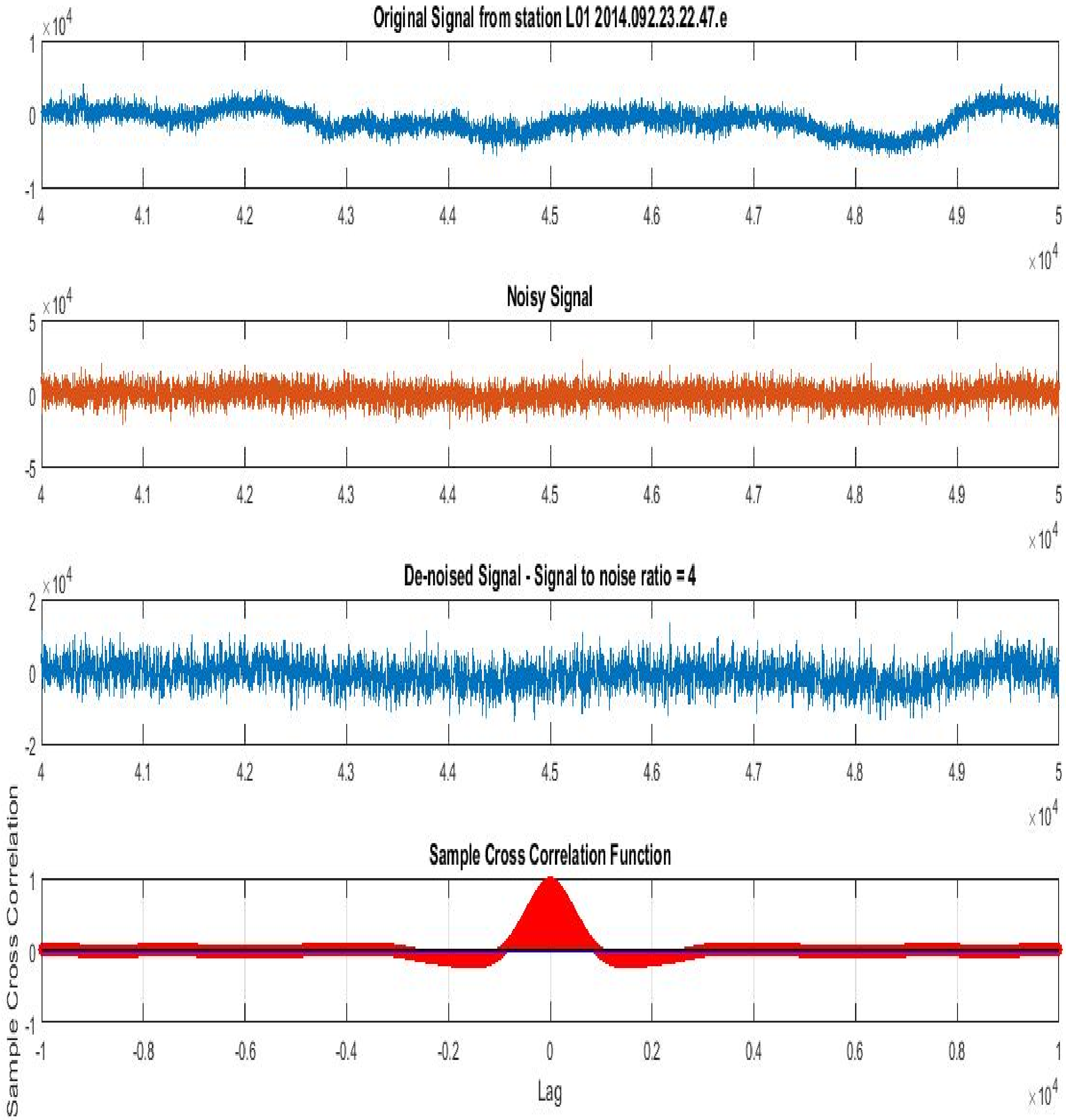
Calculating lag between two signals :

```
[c, lag]=xcorr(s.DATA1,denoised signal);
[maxC,I] = max(c);
lag = lag(I);
```

The “lag” between ‘s.DATA1’ and ‘denoised signal’ turns out to be -1

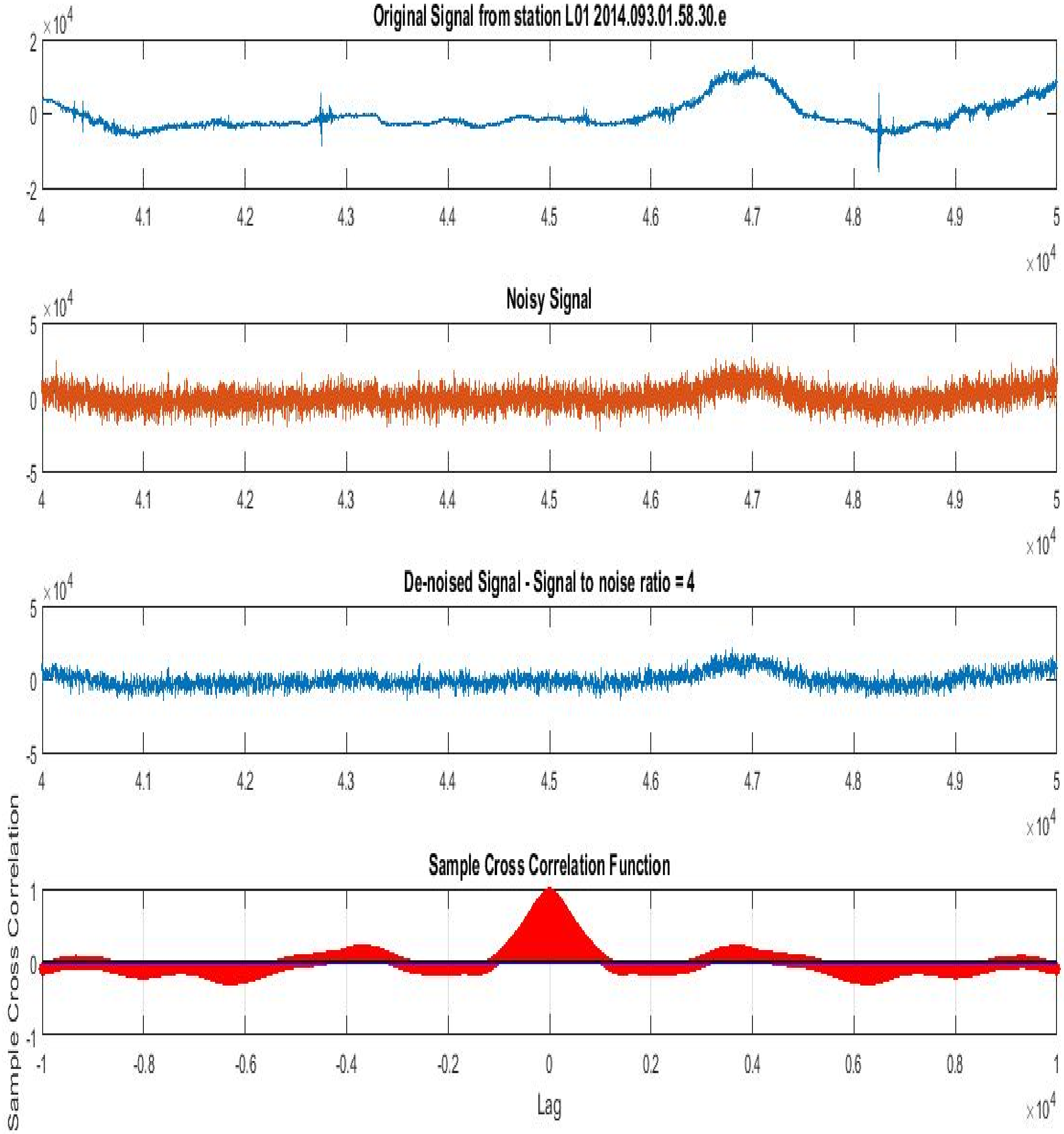
The similar computation follows for all the rest of tested stations.

2. Station: L01_2014.092.23.22.47.e



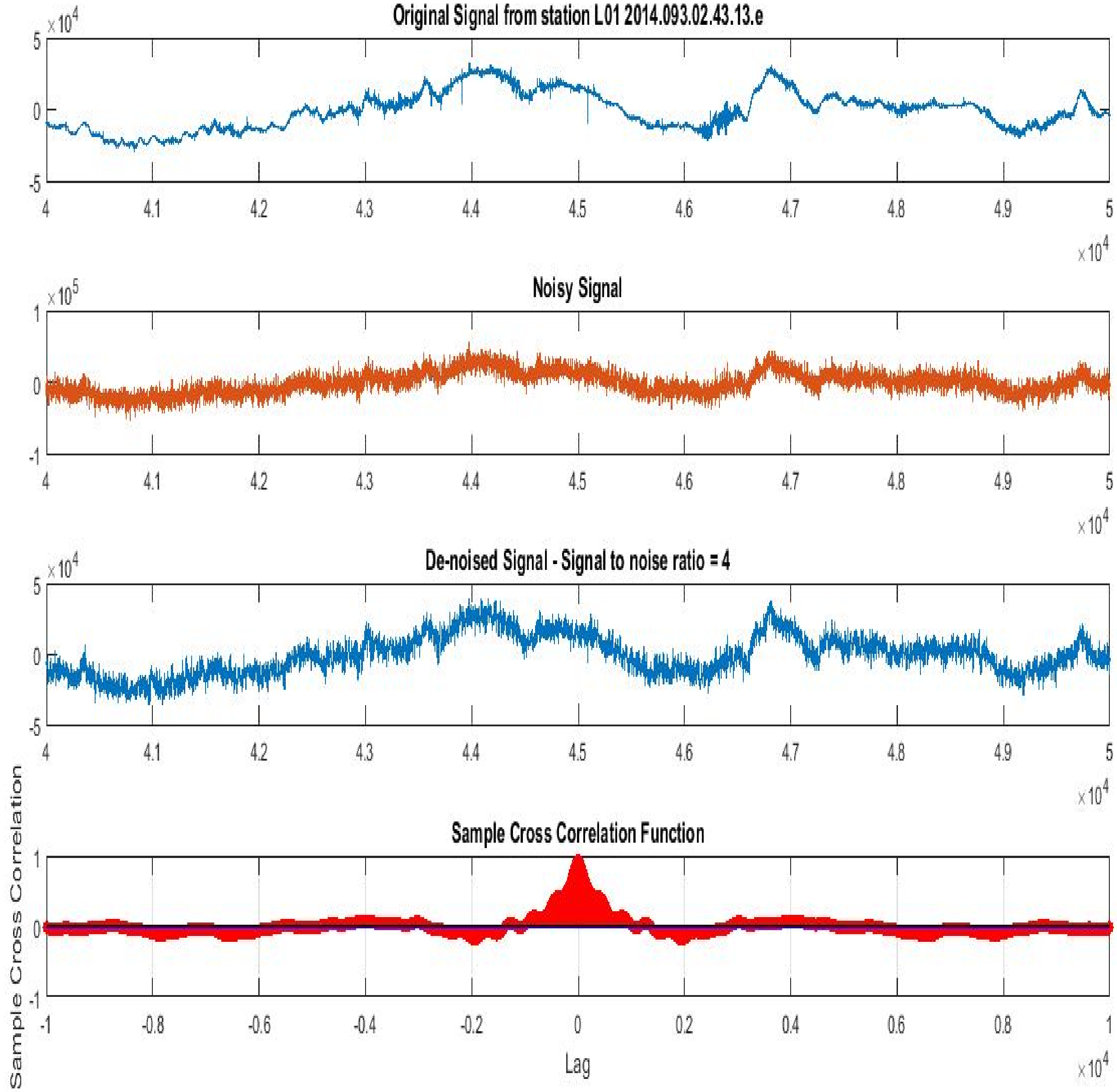
Lag = −1

3. Station: 'L01_2014.093.01.58.30.e



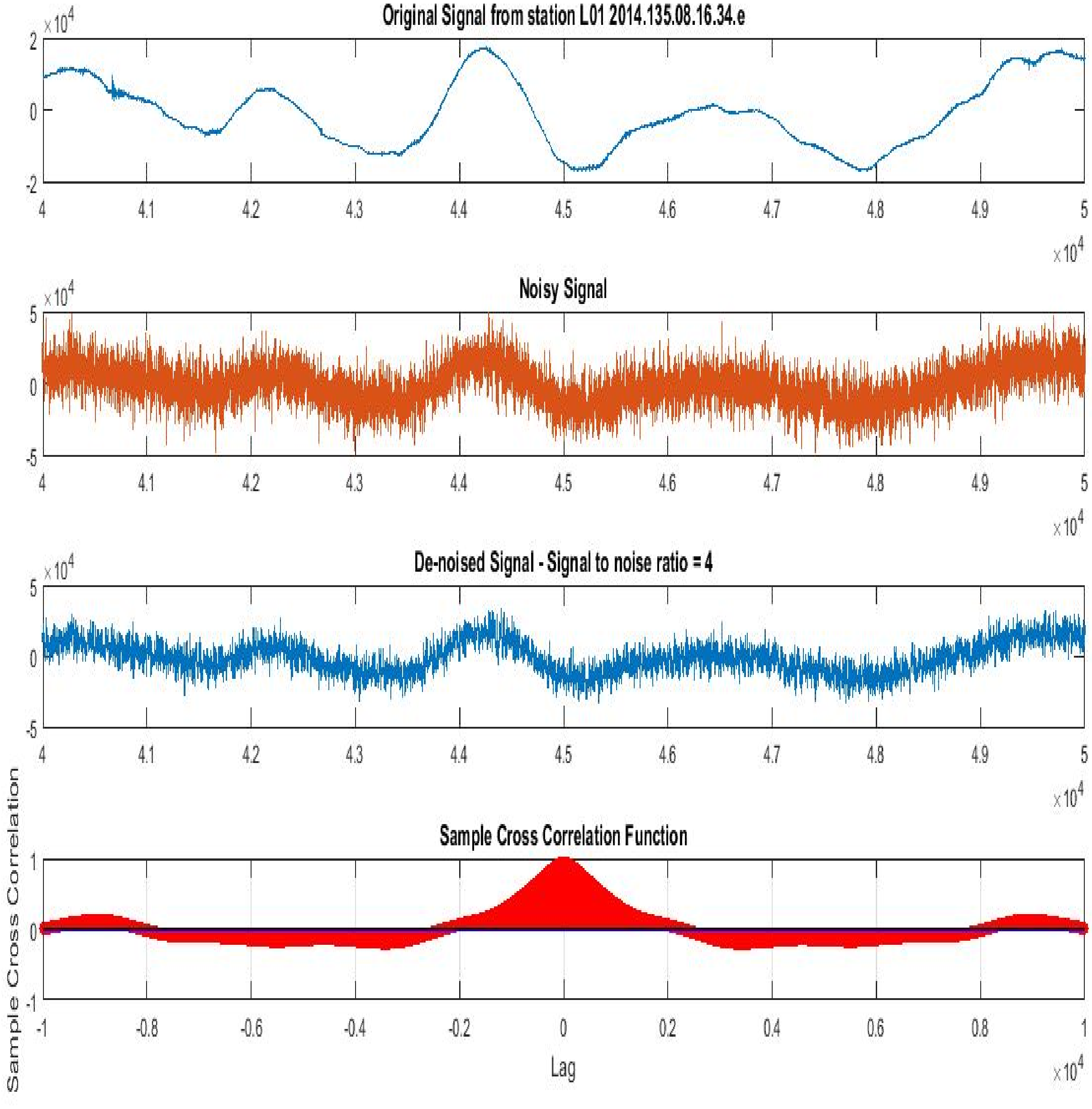
Lag = −1

4. Station: L01_2014.093.02.43.13.e



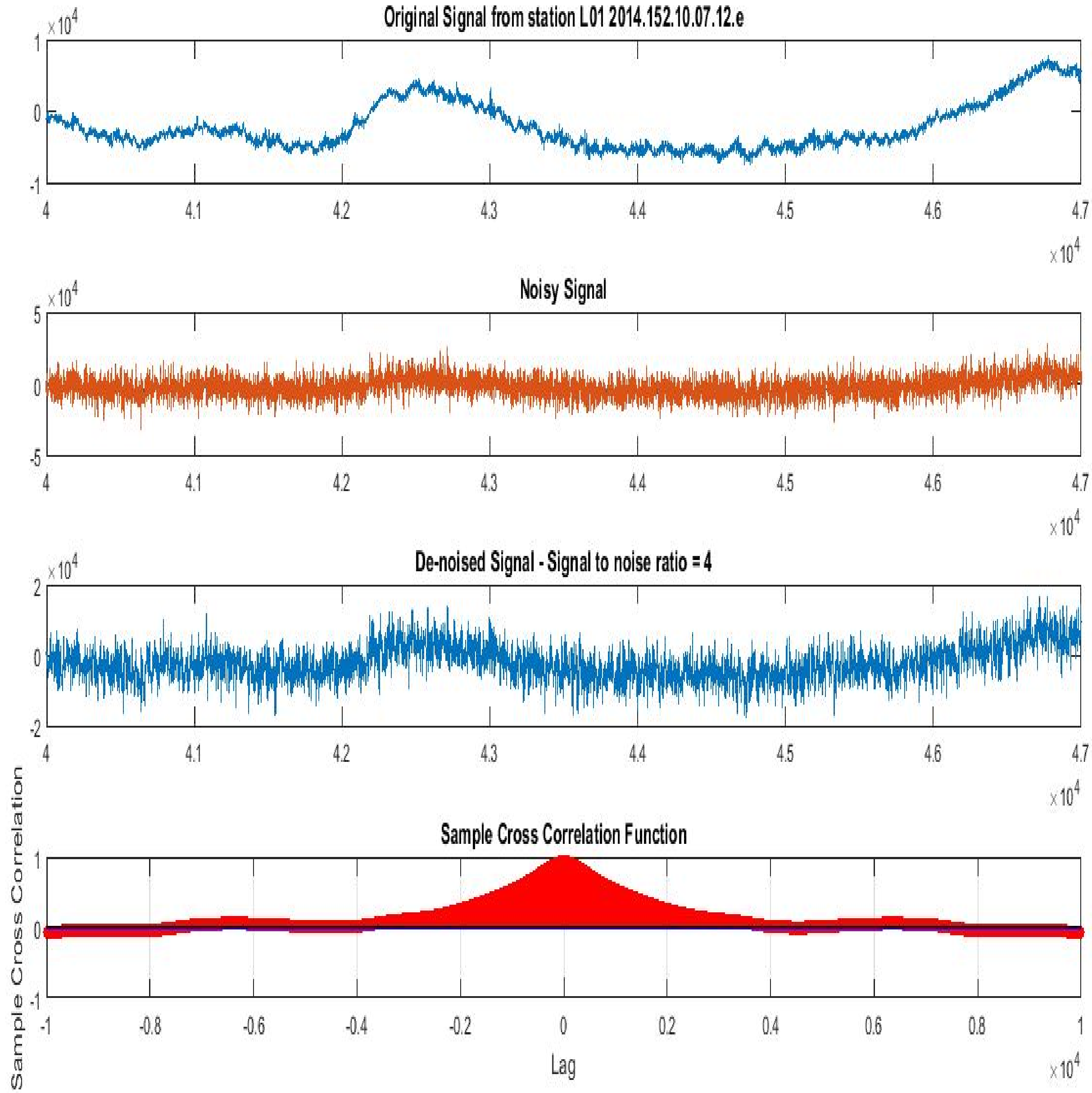
Lag = -1

5. Station: L01_2014.135.08.16.34.e



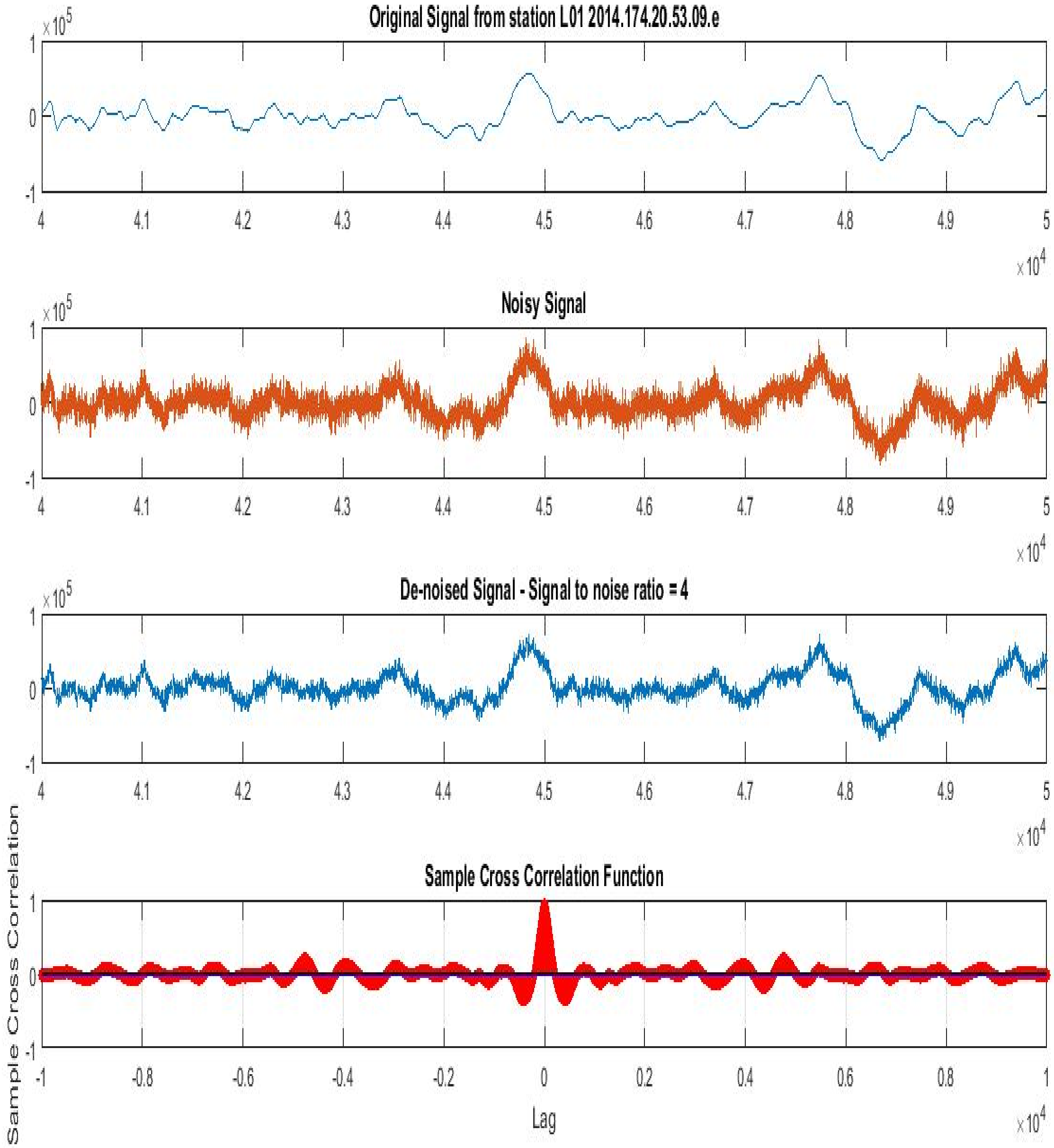
Lag = −1

6. Station: L01_2014.152.10.07.12.e



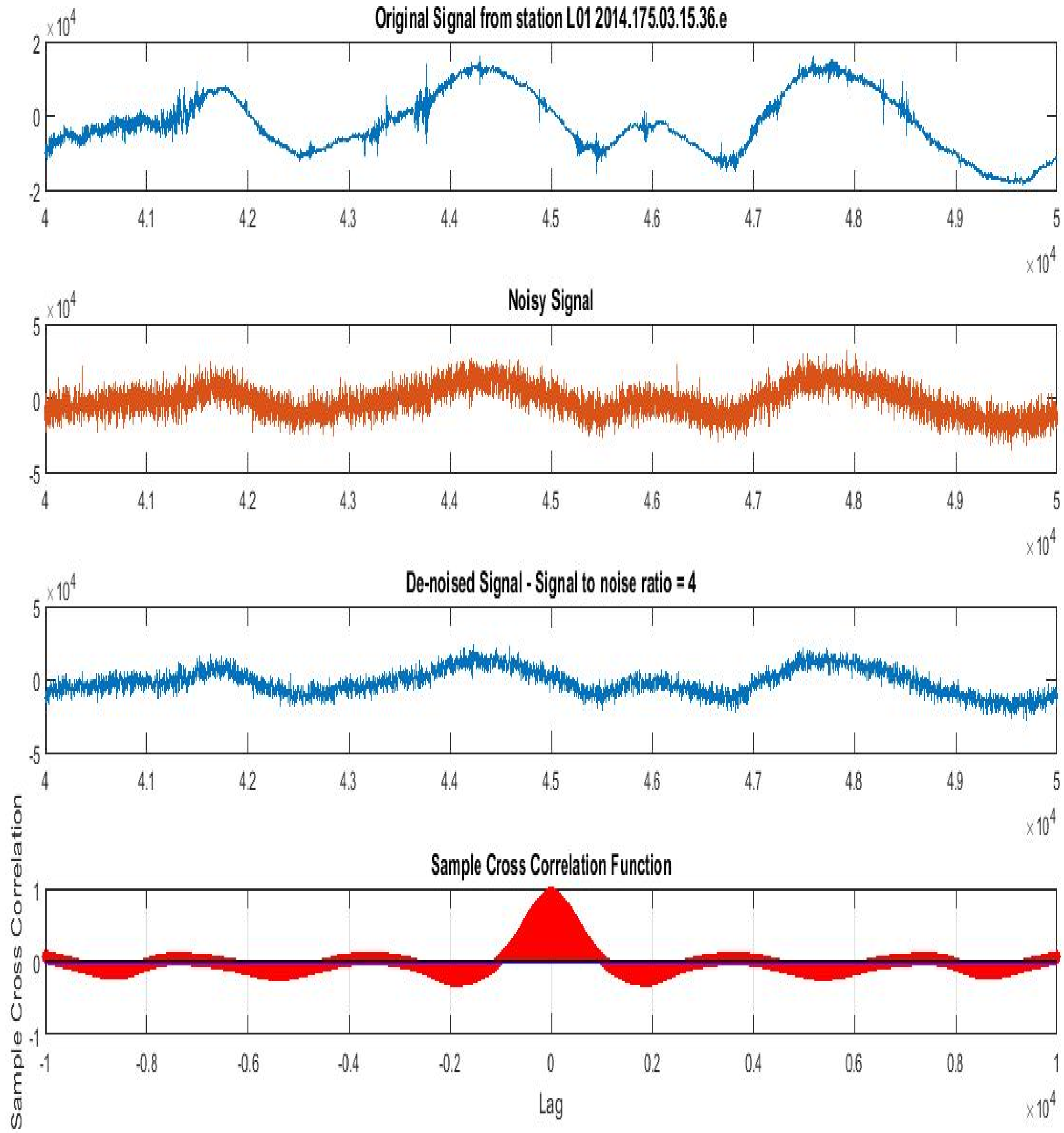
Lag = −1

7. Station: L01_2014.174.20.53.09.e



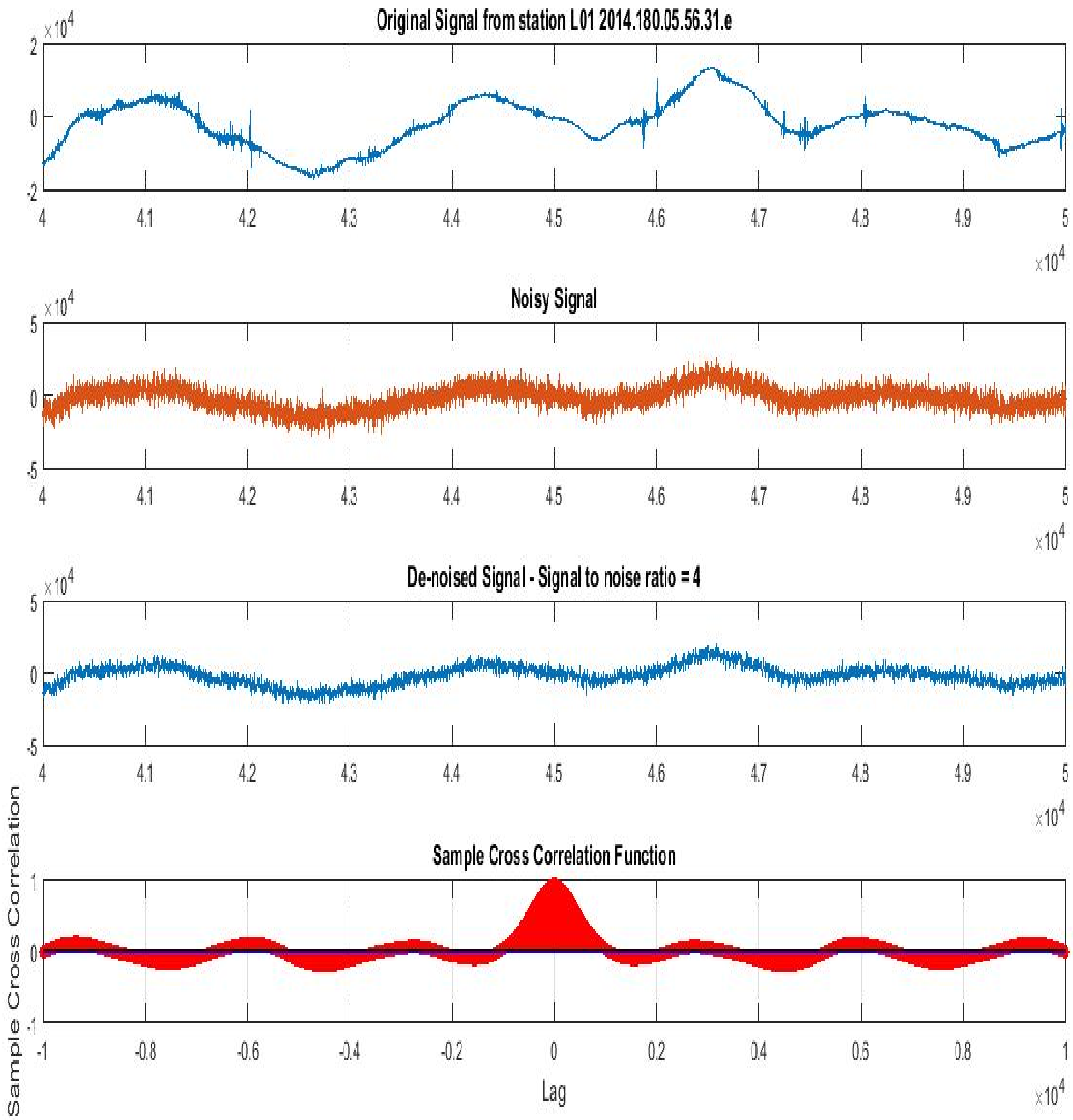
Lag = −1

8. Station: L01_2014.175.03.15.36.e



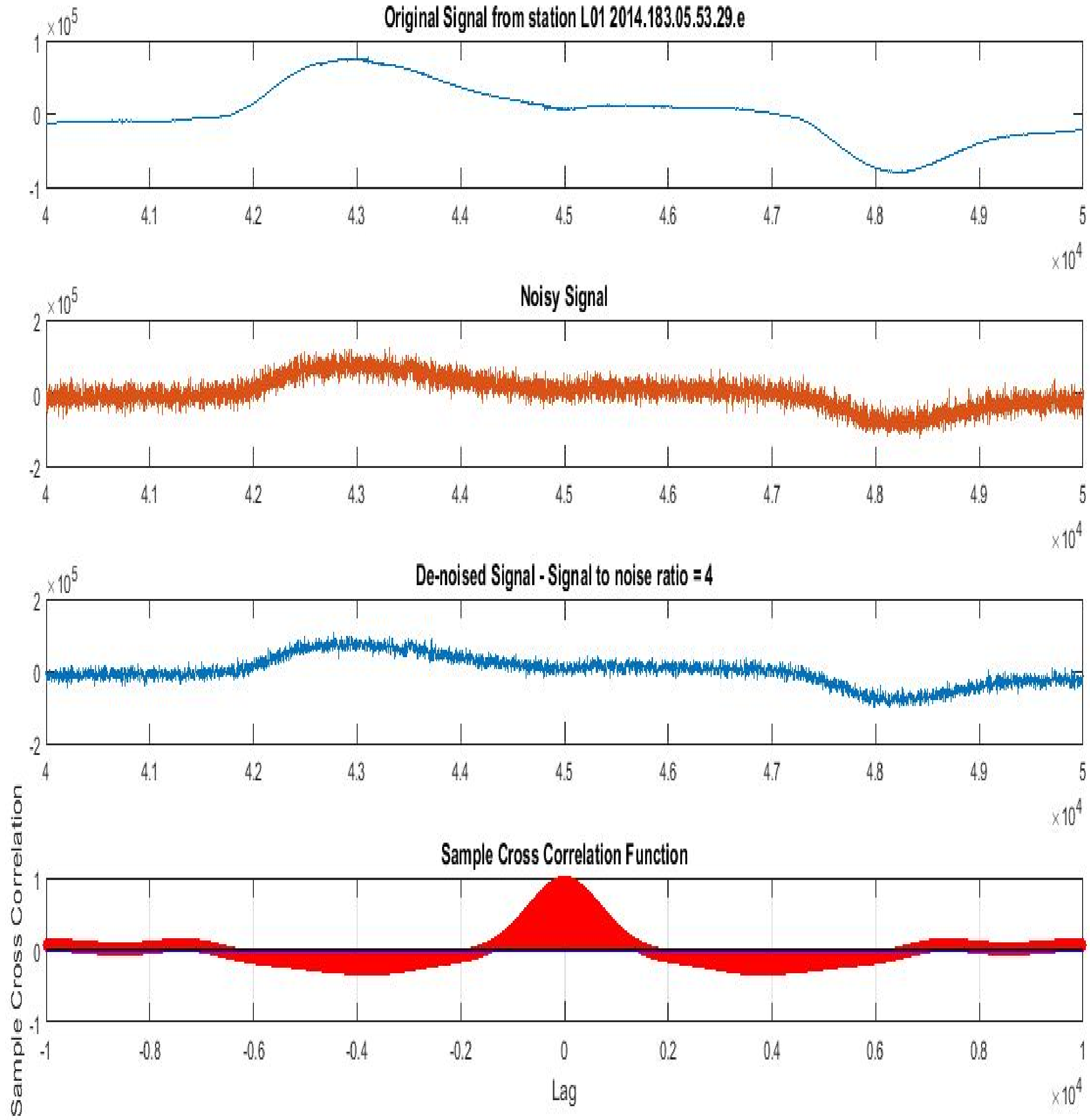
Lag = −1

9. Station L01_2014.180.05.56.31.e



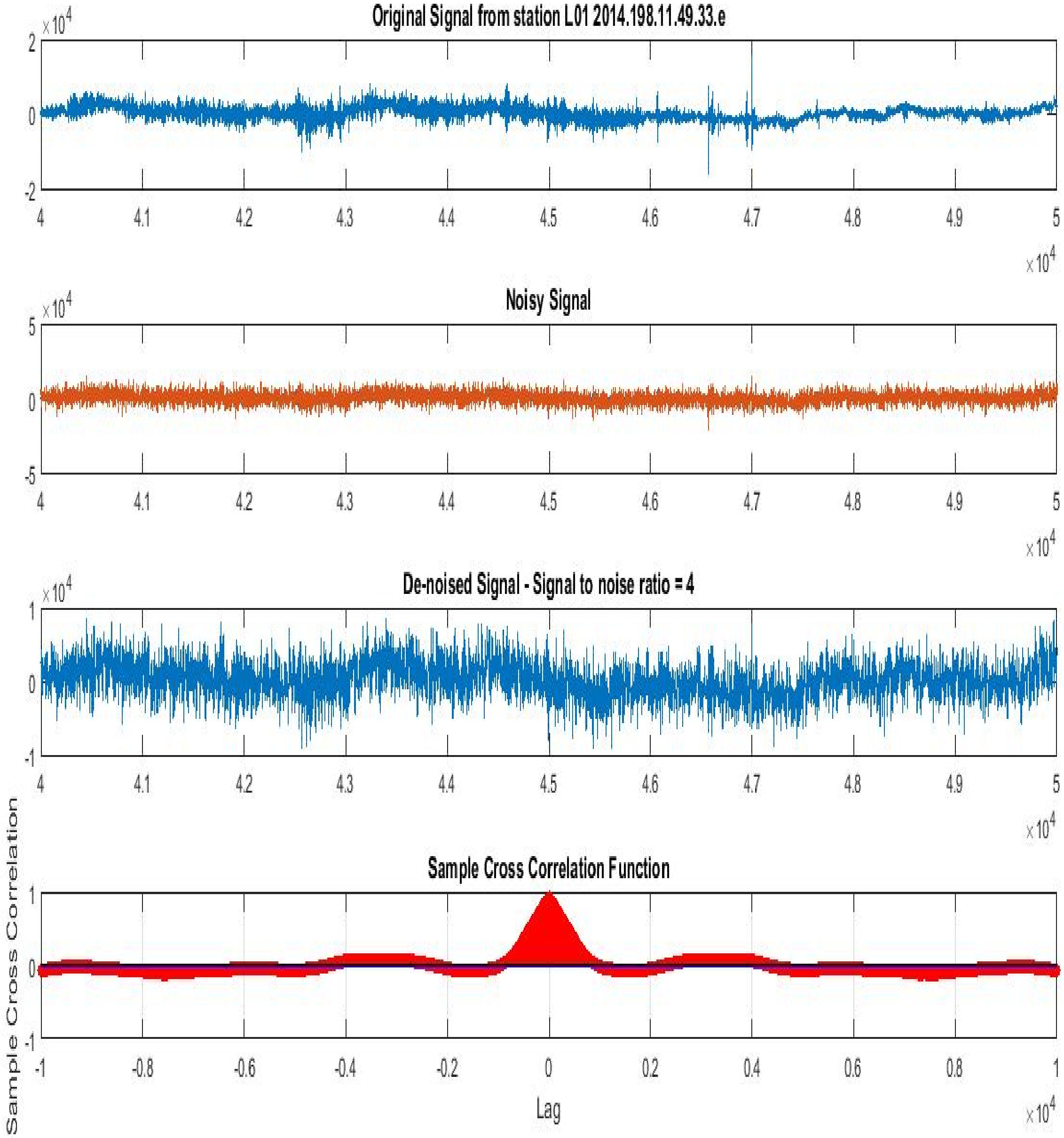
Lag = −1

10. Station: L01_2014.183.05.53.29.e



Lag = −1

11. Station: L01_2014.198.11.49.33.e



Lag = –1

Chapter 5

Results and Discussions

The Digital filter I used as a denoising method is able to retrieve the original signal upto the signal to noise ratio per sample equal to $4\ dB$.

Furthermore decrement in the signal to noise ratio (snr) leads to information loss and $|lag|$ between two signals starts increasing.

I initiated the computations by adding noise in the seismological data taken from different locations, by fixing at a higher value of signal to ratio. Then after the removal of noise using digital filter the “lag” value between the two times series came out to be 0 which means there is a perfect correlation between original signal and denoised signal. In other words the original signal was perfectly retrieved after removal of noise.

The next task was to scrutinize the “lag” values by keep reducing the signal to noise ratio or to check the confidence limit (signal retrieving capacity) of the opted denoising algorithm.

According to the plots mentioned, if we reduce signal to noise ratio below $4\ dB$ then the lag value starts increasing and we won’t be able to recuperate the signal completely and consequently the information will be erroneous.

The nature of noise being added to the signal can also effect the signal differently and we might get a distinct confidence limit.

We are now able to handle different types of noise that can be developed in a signal due to various reasons.

When seismologists collects the seismological data from different location then it’s very important for them to keep the information preserved so that appropriate predictions can be made about earthquakes. But sometimes the data which is collected over a course of time gets corrupted by noise that might have generated due to circumstantial conditions. In order to regain the information, Moving average filter can be used as an effective algorithm for smoothening of signal as it’s easily applicable on large data sets and it can be handled according to the nature of signal.

Chapter 6

Conclusion

These methods of denoising plays a significant role not only in keeping the information stored in seismological data (which helps in quantifying and detecting risks of earthquakes) conserved but also they are useful in various fields such as medical imaging. Information stored in an image which might have corrupted by noise, can be retrieved using the 'digital filter'.

However, as discussed earlier, there have been proposed many methods of denoising, one of them which can be reserved as a future scope of this project, is "Stacking and Denoising with Discrete Orthonormal S – Transform".

I will present a glimpse of this algorithm.

It works on S – Transform which is similar to that of Fourier transform as it decomposes the data from time domain to frequency domain. The only difference is that it enables better time resolution of high-frequency signal.

At first, every component, x_m, taken from station, i, at location r_i having particle displacement, $x_m(r_i, t)$, is divided into N_K smaller segments then every segment is cross correlated with each component, x_n taken from another station.

Now after stacking, the stacked cross correlation is defined by,

$$C_{mn}(r_i, r_j, t) = \sum_{k=1}^{N_k} \int_{t_k}^{t_{k+1}} x_m(r_i, s)\, x_n(r_j, t+s) ds$$

The derivative of this cross correlation is related to Green's function as,

$$\dot{C}_{mn}(r_i, r_j, t) = b[G_{mn}(r_i, r_j, t) - G_{mn}(r_i, r_j, -t)]$$

Where,

$G_{mn}(r_i, r_j, t)$ is Green's function.

Then the designed algorithm "time frequency filter" is applied to smoothen the signal. The benefit of using this method is that it permits for noise in same pass-band as signal can be precluded from signal so long as it's temporally parted from arrivals.

Bibliography

[1] Automatic detection and denoising of signals in large geophysical datasets *by Gabriel O. Trisca*

[2] Albert Tarantola, 2005, *Inverse problem theory and methods for model parameter estimation*, SIAM Publications, USA

[3] Singular value Decomposition. http://www.wou.edu/~beavers/Talks/Willamette1106

[4] William Menke, 2012, *Geophysical data analysis: Discrete inverse theory*, 3rd edition, Academic press,Elsevier Inc., UK

[5] Invertible matrix theorem.http://mathworld.wolfram.com/InvertibleMatrixTheorem.html

[6] Baig, A. M., M. Campilo, and F . Brenguier (2009), Denoising seismic noise cross correlations, J. Geophys. Res., 114, B08310, doi: 10.1029/2008JB006085.

[7] Least Square Fitting. http://mathworld.wolfram.com/LeastSquaresFitting.html

[8] A. Fettweis, "Wave digital filters: Theory and practice," in *Proceedings of the IEEE*, vol. 74, no. 2, pp. 270-327, Feb. 1986, doi: 10.1109/PROC.1986.13458.

[9] Singular Value Decomposition (SVD) tutorial BE.400 / 7.548
http://web.mit.edu/be.400/www/SVD/Singular_Value_Decomposition.htm

[10] Hadamard, J. (1902) Sur les Problèmes aux Dérivées Partielles et Leur Signification Physique, Princeton University Bulletin, 13, 49-52.

[11] Oppenheim Alan V. and Willsky Alan S., 1997, Signals and Systems, 2nd Edition, Prentice Hall, USA

[12] Efficient denoising algorithms for large experimental datasets and their applications in fourier transform ion cyclotron resonance mass spectrometry.
http://www.pnas.org/content/111/4/1385.full.pdf

[13] SAC Headers. https://ds.iris.edu/files/sac-manual/manual/file_format.html

[14] Michiel de Bondt, Arno van den Essen, Singular Hessians, Journal of Algebra, Volume 282, Issue 1,2004, doi: 10.1016/j.jalgebra.2004.08.026.

[15] Browne K, Qiao S, Wei Y (2009) A Lanczos bidiagonalization algorithm for Hankel matrices. Linear Algebra Appl 430:1531–1543.

[16] Golub GH, Loan CFV (1996) Matrix Computations (Johns Hopkins Univ Press, Baltimore).

[17] Cai JF, Candès EJ, Shen Z (2010) A singular value thresholding algorithm for matrix completion. SIAM J Optimiz 20:1956–1982.

[18] Stockwell, R. G. (2007), A basis for efficient representation of the S-transform, Digit. Signal Process., 17, 371 – 393, doi:10.1016/j.dsp. 2006.04.006.

[19] Schimmel, M., and H. Paulssen (1997), Noise reduction and detection of weak, coherent signals through phase weighted stacks, Geophys. J. Int., 130, 497 – 505, doi:10.1111/j.1365-246X.1997.tb05664.x.

[20] Pooja Jain, Ram Bilas Pachori (2015), An iterative approach for decomposition of multi-component non-stationary signals based on eigenvalue decomposition of the Hankel matrix, Journal of the Franklin Institute, Volume 352, Issue 10, ISSN 0016-0032.

[21] Introduction to Seismic Analysis Code (SAC). http://geophysics.eas.gatech.edu/classes/SAC/

[22] Georgios K. Apostolidis, Leontios J. Hadjileontiadis (2017) Swarm decomposition: A novel signal analysis using swarm intelligence, Signal Processing, Volume 132,ISSN 0165-1684.

[23] E.C. Titchmarsh, "Introduction to the theory of Fourier integrals" , Oxford Univ. Press (1948)

[24] Donoho DL (1995) De-noising by soft-thresholding. IEEE Trans Inf Theory 41:613–621.

[25] Schimmel, M., and J. Gallart (2007), Frequency-dependent phase coherence for noise suppression in seismic array data, J. Geophys. Res., 112, B04303, doi: 10.1029/2006JB004680.