

Blind Deconvolution of Seismic Signals in Time and Frequency Domain: a Study on 2004 Sumatra Earthquake

Kedarnath Senapati
Institute of Mathematics and
Applications,
Bhubaneswar, 751003, Orissa, India
kedar_d2@yahoo.co.in

Aurobinda Routray
Department of Electrical Engineering
Indian Institute of Technology,
Kharagpur, 721302
West Bengal, India.
Aurobinda.Routray@iitkgp.ac.in

Santosh Dhubia
Department of Geology and
Geophysics,
Indian Institute of Technology,
Kharagpur, 721302,
West Bengal, India,
santoshdhubia@yahoo.co.in

William Kumar Mohanty
Department of Geology and Geophysics
Indian Institute of Technology,
Kharagpur, 721302, West Bengal, India
wkmohanty@yahoo.co.in

Abstract

Blind deconvolution of Green's Function (GF) as well as the Source Time Function (STF) is a fundamental problem in seismic signal processing. This paper presents a comparative study of two blind deconvolution methods for determining the Green's function as well as the source time function from the recorded seismic data. We have tested these algorithms on the December 26, 2004 Sumatra earthquake ($M_w=9.3$). An improvised Least Mean Square (LMS) algorithm has been used to improve the multi-pulse method proposed earlier by Cooley et al. [1]. We have also evaluated the performance of this improvised multi-pulse method against the Frequency-Domain Blind Deconvolution (FDBD) method. The FDBD method is based on Mutual Information Rate (MIR) proposed by Larue et al. [3]. These methods have also been tested for some of the records of the earthquake at different stations worldwide available on IRIS-dataset. We have found the azimuthal dependency of the source time function on these records. We have also compared both the methods using synthetic as well as real data. The STFs recovered on different stations using multi-pulse model and FDBD method show similarity in shape and duration as well. However the proposed improvisation on the multipulse method is computationally efficient.

Keywords: *Green's function, Source Time Function, Blind Deconvolution, Mutual Information Rate, Multi-pulse method, Least Mean Square Algorithm.*

1. Introduction

The signatures of a seismogram are the combined effect of the source, the propagation path of the seismic wave, the response of the recording instrument and the ambient noise present at the recording site. It is important to know the source properties to determine the kind of rupture and released energy in the event. Mathematically, seismogram is considered to be the convolution of the Source Time Function (STF) and Green's Function (GF). The STF is an important physical quantity characterizing the rupture process of an earthquake source. If an earthquake is considered as a rupture of rocks, the STF of the earthquake is defined as the time history of the relative displacement of the rupture surface. The GF can be defined as the recording at the receiver when STF of the earthquake is a unit impulse.

Numerous studies have been shown that the STF for any point is closely related to the earthquake source process [4, 5]. Estimation of both STF and GF directly from the observed seismic data without any prior information about them is referred to as blind deconvolution. The computation of exact theoretical GF is quite difficult as the earth model becomes more and more sophisticated and complicated. Hence, there is a need of blind deconvolution algorithms that can perform estimation of STF and GF from the recorded seismograms.

There are several methods currently available in literature for the estimation of STF and GF. Some of the methods are linear inversion [9], source estimation from teleseismic data [7] and broadband regional seismograms [10]. Further, the evaluation of Green's functions for elastic layered media has been proposed on the basis of exact discretization of the wave field emitted by the source [13]. In the synthetic ground motion for Large/Great earthquakes, small earthquakes from the source region are used as the Empirical Green's Functions (EGFs). It can be achieved in two ways i.e. one by random summation of EGFs and another by stochastic method [12, 14]. There is a need to record the smaller magnitude earthquakes from the source regions to estimate the empirical Green's functions. Bostock [11] used the spectral properties of the signals to separate the STF from the scattered teleseismic data. Sabra *et al.* [15] proposed a method for the estimation of Green's function between two stations, with the help of ambient noise. Sèbe *et al.* [8] used *coda* waves for the estimation of STF, and the deconvolution of three-component teleseismic P waves is investigated using the autocorrelation of the P to SV scattered waves by Dasgupta *et al.* [6]. However, these methods have not discussed about the time complexity and execution time for estimation of STF and Green's function, which turns out to be a very important consideration for real time seismic signal analysis.

The rest of the paper is organized into sections. Section 2 discusses the preliminary literature review of seismic deconvolution techniques both in time and frequency domain. In Section 3, the theory for multi-pulse model [1] and frequency domain blind deconvolution method [3] to estimate GF has been discussed. We propose a fast method to find the STF from recorded seismic data. The method uses the multi-pulse algorithm and a modified Least Mean Square (LMS) method for estimating the STF. This has been discussed in Section 4. We have also evaluated the performance of this improvised multi-pulse method against the frequency-domain blind deconvolution method. Validation of both the methods has been carried out using synthetic data generated by convolutive models. Then both the methods have been applied on the recording of 26 December 2004 Sumatra earthquake. It has been discussed in Section 5. Section 6 is the conclusion of the work.

2. Seismic Blind Deconvolution

To isolate the source time function from path and recording instrument effects, several methods have been developed to date and are successfully in use, each with its own merits and shortcomings. These methods may be divided into two categories as waveform modeling which includes different approaches, e.g., moment tensor inversion [16]-[18] and waveform inversion [19, 20] and deconvolution methods which involve both time domain [21, 22] and frequency domain [23]-[27]. All these methods require information about the propagation path which may or may not always be available with required accuracy. This can be achieved by using recording of a smaller earthquake close to large earthquake recorded at the same station. If the focal depth and focal mechanism of two events are identical, the earth response to each recording will be the same in both cases. If the small event has an impulse-like source time function then it approximates earth's Green's function, known as Empirical Green's Function (EGF) [28], which includes attenuation, propagation, instrument and radiation pattern effects with corresponding seismic moment. Deconvolving this EGF record from corresponding larger event record in either time domain or frequency domain approximates the larger event source time function normalized by smaller event seismic moment which is known as Relative Source Time Function (RSTF). Azimuthal variations in the RSTF provide directivity patterns which allow the finiteness and rupture process of large event to be studied. Many researchers have applied this method to study source complexity of the larger earthquakes [23]-[27], [29]-[31]. This method has the limitation of frequency band, and is valid for the frequencies below *corner* frequency of smaller event, and also has the difficulty of obtaining smaller event records with similar location and identical focal mechanism recorded at different stations as the larger event. Thus, there is a need of methods which can be applied to the earthquake recordings without having *a priori* knowledge about path or source characteristics.

In this study, we address the problem of single-input single-output (SISO) seismic blind deconvolution. The observation is the convolution between the unknown source (STF) and the unknown path function (GF) and is corrupted by additive noise. If g is the GF and d is the STF then the observed seismogram can be written as follows

$$y(t) = g * d(t) + u(t) \quad (1)$$

where $u(t)$ is the ambient noise present in the recorded data. Blind deconvolution refers to the problem of determining the GF and STF only from the observed seismogram.

3. The Blind Deconvolution methods

This section frames the theory for the multi-pulse model and frequency domain blind deconvolution method to estimate GF.

3.1. Multi-pulse Method

Green's function or path characteristic of seismic record is estimated using multi-pulse method [1] and modified for the application in earthquake seismology. This is based on the common AutoRegressive (AR) model used in applications like speech processing and seismic deconvolution [34]. The model can be described as

$$s(n) = \sum_{i=1}^P a_i s(n-i) + x(n) + u(n), \quad 1 \leq n \leq N. \quad (2)$$

Here P is the order of the model, $x(n)$ is the excitation or input, and $s(n)$ is the output or the recorded signal, respectively. $u(n)$ is the noise component which is assumed to be white. The multi-pulse method approximates this model as

$$\hat{s}(n) = \sum_{i=1}^P \alpha_i s(n-i) + \hat{x}(n) \quad (3)$$

where $\hat{s}(n)$ is the approximation of $s(n)$, $\alpha_i (1 \leq i \leq P)$ are the estimates of a_i and $\hat{x}(n)$ is an estimate of $x(n)$. An error sequence $e(n)$ is defined as the difference between actual signal and the predicted signal given by Eq. 3

$$e(n) = s(n) - \hat{s}(n), \quad 1 \leq n \leq N \quad (4)$$

$$\text{or } e(n) = s(n) - \sum_{i=1}^P \alpha_i s(n-i) - \hat{x}(n). \quad (5)$$

In the multi-pulse method, it is assumed that the excitation is not a periodic pulse, and input is made up of I pulses such that i^{th} impulse has amplitude of b_i with position p_i ,

$$\hat{x}(n) = \sum_{i=1}^I b_i \delta(n - p_i) \quad (6)$$

where, the discrete delta function is defined as:

$$\delta(k) = \begin{cases} 1 & k = 0 \\ 0 & \text{otherwise} \end{cases}$$

The error sequence $e(n)$, total error E and Residual Error R_E can then be written as

$$e(n) = s(n) - \sum_{i=1}^P \alpha_i s(n-i) - \sum_{i=1}^I b_i \delta(n - p_i) \quad (7)$$

$$E = \sum_{k=1}^N e^2(k) \quad (8)$$

$$R_E = \frac{\sum_{k=1}^N e^2(k)}{\sum_{k=1}^N s^2(k)}. \quad (9)$$

The pulse locations and amplitudes that will minimize the total error defined by Eq. 8 can be determined with the help of following algorithm:

Step 1

Given the seismic record $s(n)$, obtain filter coefficient α_i from:

$$\mathbf{H}\boldsymbol{\alpha} = \mathbf{h} \quad (10)$$

where the elements of the matrix \mathbf{H} and \mathbf{h} are given by

$$H_{ij} = \sum_{n=0}^{N-1} s(n-i)s(n-j) \quad (11)$$

and,
$$\mathbf{h} = [h(1), \dots, h(P)]^T \quad (12)$$

with,
$$h(i) = \sum_{n=0}^{N-1} s(n)s(n-i). \quad (13)$$

Singular Value Decomposition (SVD) technique [41] is used to find out filter coefficients by solving Eq. 10. The prediction error for current filter coefficient is given by:

$$e(n) = s(n) - \sum_{i=1}^P \alpha_i s(n-i). \quad (14)$$

Model given by Eq. 3 and Eq. 6 reflects that if the filter coefficients are approximately true filter coefficients, then pulse positions will be at largest errors. Furthermore, for noise free recording and with true filter coefficients, the error amplitude matches exactly that of the input pulse. Thus, the pulse amplitude and its position is obtained by $b_i = e(n_i)$ for the I largest errors. Surface wave part which almost covers 90% of the seismogram is used widely for estimation of STF. However, for finding the Green's function the body wave part should also be considered. Therefore we divided the whole seismogram into two parts: one before significant arrival of surface wave and another after it. And 10% weight is given for the pulses in first part and 90% to the rest of the seismogram.

Step 2

The new values of α_i (Eq. 10) are computed using summations which do not include these pulse positions at I largest errors.

Step 3

The prediction error (step 1) is then recalculated using the new filter coefficients obtained in step 2. If the estimated pulse positions are unchanged, then the iteration is said to have converged; if new positions are indicated, return to step 2.

So far the proof of convergence of this method has not been envisaged in the literature. Therefore the number of iterations in this case has been fixed by trial and error to avoid infinite loop situations.

3.2. Frequency Domain Blind Deconvolution (FDBD)

The recorded seismogram $d(t)$ can be modeled as the convolution between the unknown source $r(t)$ and the unknown filter u and is corrupted by additive noise $n(t)$. In discrete time domain, the recorded seismogram $d(t)$ is given by

$$d(t) = (u * r)(t) + n(t) = \sum_{i=-\infty}^{+\infty} u(i)r(t-i) + n(t). \quad (15)$$

One of the many applications of this model is reflectivity characterization in seismology, where a short duration acoustic wave is transmitted into the ground. The reflected energy

resulting from the series impedance in the earth is measured. Therefore, each reflection can be expressed as the convolution of delta function with the emitted wavelet. Then, in the above model (15), the source $r(t)$ is the vertical reflectivity sequence of the earth, while u represents the unknown emitted wavelet and the coupling effect between the source and the earth.

In the blind deconvolution context, only $d(t)$ is accessible to the algorithm, whereas $r(t), n(t), u$ are unknown. The task here is to estimate an inverse filter g for providing the deconvolution output $y(t) = g * d(t)$, the most similar to $r(t)$. Another way to model in (15) is to assume that source samples $r(t)$ are independent and identically distributed (i.i.d) and non-Gaussian. Thus, the solution is unique up to delay and magnitude indeterminacies. The method aims at estimating a deconvolution filter g which provides an i.i.d. output signal $y(t) = g * d(t)$. In fact, we adjust the inverse filter g for maximizing an i.i.d. criterion of the output $y(t)$. And for that purpose, the deconvolution filter is estimated by minimizing the Mutual Information Rate (MIR) used as an i.i.d. criterion.

The mutual information [16] of a random vector $\mathbf{z} = [z_1, z_2, z_3, \dots, z_n]$ is defined by

$$I(\mathbf{z}) = \sum_{i=1}^n H(z_i) - H(z_1, z_2, \dots, z_n) \quad (16)$$

where $H(z_i)$ is the Shannon marginal entropy of z_i , and is defined by $H(z_i) = \int_{\square} p_{z_i}(u) \log p_{z_i}(u) du$ and the second term of Eq. 16, $H(z_1, z_2, \dots, z_n)$, the Shannon joint entropy $H(\mathbf{z}) = \int_{\square^n} p_{\mathbf{z}}(\mathbf{u}) \log p_{\mathbf{z}}(\mathbf{u}) d\mathbf{u}$. The entropy rate [34] of a stochastic process $Z = \{Z_t\}$ is defined as

$$\mathcal{H}(Z) = \lim_{T \rightarrow \infty} \frac{1}{T} H(Z_1, \dots, Z_T). \quad (17)$$

Then, the MIR of a stationary process Z is defined by

$$I(Z) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T H(Z_t) - \mathcal{H}(Z) \quad (18)$$

where $H(Z_t)$ denotes the marginal entropy of the process Z_t and $\mathcal{H}(Z)$ denotes the entropy rate of Z defined by Eq. 17. If we take stationary assumption, then marginal entropy of the process Z_t will not be time dependent. And, all the terms of the right-hand-side in Eq. 18 are equal, so we have

$$\lim_{T \rightarrow \infty} \sum_{t=1}^T H(Z_t) = H(Z_\tau), \quad \forall \tau. \quad (19)$$

Hence, adding stationary assumptions, the MIR (18) is simply given by

$$I(Z) = H(Z_\tau) - \mathcal{H}(Z). \quad (20)$$

Where, τ is chosen arbitrarily. In the blind deconvolution context, we can obtain a simpler cost function [2], [32], [35], noticing that the entropy rate of the deconvolution output $y(t) = (g * d)(t)$ is

$$\mathcal{H}(Y) = \mathcal{H}(D) + \frac{1}{2\pi} \int_0^{2\pi} \log \left| \sum_{t=-\infty}^{\infty} g(t) \exp(-jt\theta) \right| d\theta. \quad (21)$$

With Eq. 21, the MIR of the deconvolution output becomes

$$I(Y) = H(y(\tau)) - \mathcal{H}(D) - \frac{1}{2\pi} \int_0^{2\pi} \log \left| \sum_{t=-\infty}^{\infty} g(t) \exp(-jt\theta) \right| d\theta. \quad (22)$$

Then, since the entropy rate $\mathcal{H}(D)$ is independent of the inverse filter g , instead of (22), one can consider the simplified criterion $\tilde{I}(Y)$

$$\tilde{I}(Y) = H(y(\tau)) - \frac{1}{2\pi} \int_0^{2\pi} \log \left| \sum_{t=-\infty}^{\infty} g(t) \exp(-jt\theta) \right| d\theta \quad (23)$$

which, like Eq. 22, is minimum when the process Y is i.i.d. Taleb *et al.* [2], estimate the inverse filter g in the time domain, by minimizing (23) with respect to the impulse response $g(t)$:

$$g^{MAML}(t) = \arg \min_{g(t)} \tilde{I}(Y). \quad (24)$$

The minimum search of (23) is made by a gradient technique with respect to the coefficients of the impulse response of the inverse filter g . This method is called Moving Average Maximum Likelihood (MAML). The cost function to be minimized according to a gradient iterative procedure, with respect to the inverse filter frequency response $\mathbf{G} = [G(0), \dots, G(T-1)]$, is derived by Larue *et al.* [3] and is given by

$$J(\mathbf{G}) = H(y(\tau)) - \frac{1}{T} \sum_{\nu=0}^{T-1} \log |G(\nu)| + \lambda_1 \sum_{\nu=0}^{T-1} |G(\nu) - G(\nu+1)|^2 + \lambda_2 \sum_{\nu=0}^{T-1} |G(\nu)|^p. \quad (25)$$

Where λ_1 and λ_2 denote two hyperparameters which balance the last two regularization terms of Eq. (25) and $G(\nu) = \sum_{t=1}^T g(t) \exp\left(-\frac{j2\pi t\nu}{T}\right)$, $\forall \nu = 0, \dots, T-1$.

Thus, the gradient of the frequency-domain blind deconvolution criterion (25), with respect to $G(\nu)$ is derived by

$$\begin{aligned} \nabla \hat{J}(\mathbf{G}) = & -\frac{1}{2T^2} \Psi_Y(\nu) D^*(\nu) - \frac{1}{2T} \frac{1}{G^*(\nu)} \\ & + \lambda_1 (2G(\nu) - G(\nu+1) - G(\nu-1)) + \lambda_2 \frac{p}{2} \frac{|G(\nu)|^p}{G^*(\nu)}. \end{aligned} \quad (26)$$

Where, $\Psi_Y(\theta)$ is the Fourier transform of the score function [2] $\psi_Y(u) = -\frac{d}{du} \log p_Y(u)$ of the process Y . After computing the gradient, we can minimize the criterion (25) with a gradient descent procedure. Therefore, the FDBD algorithm can be summarized as follows

- 1) Initialization of the inverse filter and of the deconvolution output;
- 2) Estimation of the score function;
- 3) Computation of the gradient estimate (26);
- 4) Updating of $G(v) \leftarrow G(v) - u \nabla \hat{J}(\mathbf{G})$;
- 5) Computation of the deconvolution output $y(t)$;
- 6) Normalization step.

Step 2 to 6 is repeated till convergence is achieved. The normalization Step is required for taking care of scale indeterminacy in $G(v)$.

In present study, the score function $\psi_Y(x)$ is estimated using Gram-Charlier expansion [33], [42] and is given by

$$\begin{aligned} \psi_X(x) = & -x - (E[x^2] - 1)H_1(x) + \frac{1}{2}E[x^3]H_2(x) \\ & + \frac{1}{6}(E[x^4] - 6E[x^2] + 3)H_3(x) + \dots \end{aligned} \quad (27)$$

where $H_k(x)$ is the standard k^{th} Chebyshev-Hermite polynomial. This kind of expansion has already been used by Comon [36] and Gaeta et al. [40] in the linear case and by Taleb and Jutten [38] and Yang et al. [42] in the nonlinear source separation case. The main advantages of this approach are its simplicity and its low computational cost.

4. Estimation of Source Time Function

Estimation of GF is discussed in the previous section using modified multi-pulse method and frequency domain blind deconvolution and once the GF is extracted from earthquake recordings, STF can be estimated using a modified LMS method proposed here. This method considers the seismograph $u_C(t)$ as the convolution of STF $B_C(t)$ and Green's function $G_C(t)$ as follows

$$u_C(t) = \int_{-\infty}^t G_C(t-t')B_C(t')dt'. \quad (28)$$

The STF can be estimated by deconvolving the seismic record with the estimated Green's function. Deconvolution can be carried out directly in frequency domain. However, due to the presence of finite poles in the Green's function, it may lead to inconsistent results. In addition to it the STF should be casual and positive in time domain.

In this paper a computationally effective modified Least Mean Square (LMS) method has been proposed to estimate the STF iteratively. The algorithm has been constrained to positivity and causality.

4.1. Modified LMS Method for Deconvolution

The Least Mean Square (LMS) algorithm, introduced by Widrow and Hoff in 1959 is an adaptive algorithm, which uses a gradient-based method of steepest decent. LMS algorithm uses the estimates of the gradient vector from the available data. LMS incorporates an iterative procedure that makes successive corrections to the weight vector in the direction of the negative of the gradient vector which eventually leads to the minimum mean square error. Compared to other algorithms LMS algorithm is relatively simple; it does require neither the correlation function calculation nor the matrix inversions.

To facilitate the seismic deconvolution, a new scheme based on a modified LMS method is proposed with constraint to positivity and causality as follows:

Step 1

Obtain an initial estimate of the discrete source time function B . The estimate of seismic record \hat{y} is calculated as

$$\hat{y}(n) = \sum_{i=1}^N G(i)B(n-i) \quad (29)$$

where, $G(i)$ is discrete time Green's function calculated by the above method.

Step 2

Calculate error in estimated and observed record by

$$e(n) = y(n) - \hat{y}(n) \quad (30)$$

The error vector is $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$, where,

$\mathbf{e} = [e_1 \ e_2 \ \dots \ e_N]^T$, $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_N]^T$, $\hat{\mathbf{y}} = [\hat{y}_1 \ \hat{y}_2 \ \dots \ \hat{y}_N]^T$ and, for $n = 1, 2, \dots, N$.

Step 3

Update STF using step size μ by:

$$\mathbf{B} = \mathbf{B} + \mu \times e(n) \times \mathbf{b} \quad (31)$$

where \mathbf{b} is the vector of length equal to length of \mathbf{B} , and obtained by taking window of length L from the right of the vector obtained by folded Green's function about its first element. This window will shift towards right by one element after every iteration, and a check is done in each iteration for positivity of the STF \mathbf{B} and negative values are clamped to zero. Fig. 1 shows graphical illustration of this step for the particular iteration number $L+k$, and Fig. 2 shows how \mathbf{b} is chosen for the vertical velocity record of Sumatra earthquake for different iteration number.

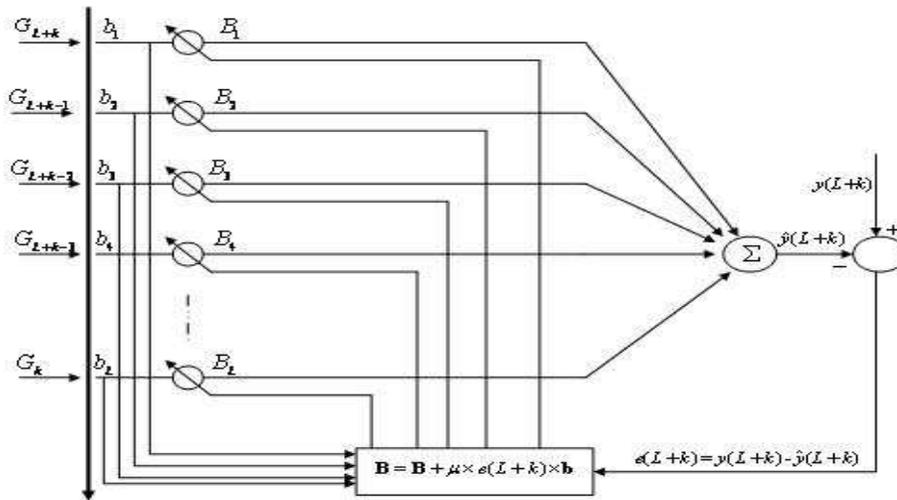


Figure 1. Graphical illustration of the step 3 for iteration number $L+k$, where L is the length of STF B .

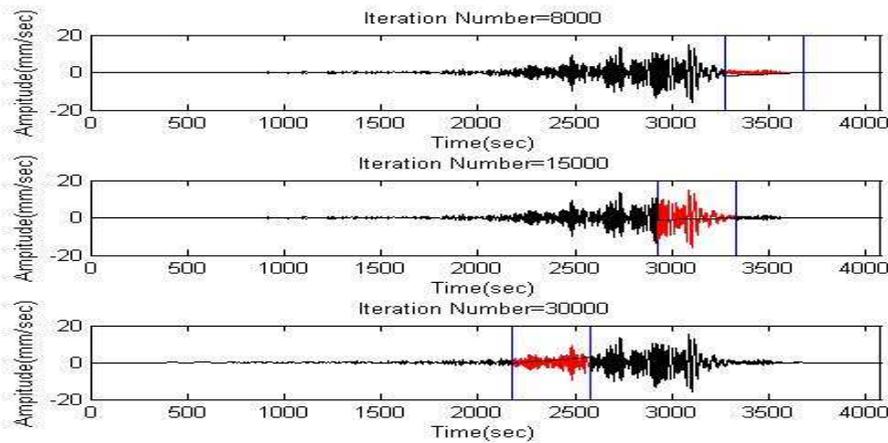


Figure 2. Graphical illustration of the step 3 with different number of iterations for 26th December 2004 Sumatra Earthquake vertical component recorded at IIT Kharagpur, and folded about its first element.

Step 4

Stop if total error E is less than acceptable value or it does not change, otherwise repeat step 1 to step 3. In each iteration step 1 to steps 4 are repeated. The convergence of algorithm depends on the maximum amplitude of the record and step size μ .

Parameters selection for the deconvolution:

- a) Length of STF, for simple earthquakes can be chosen up to maximum 100 seconds. For major earthquakes (like 2004 Sumatra earthquake) it should be chosen carefully

on the basis of earthquake-size. However, for initial estimates it can be chosen to be around 600 seconds.

- b) To satisfy the convolution theorem, on which the present deconvolution scheme stands we have to reduce the length of the estimated Green's function with no loss of information. One way is to reduce this length of Green's function by removing the number of zeros prior to the first arrival. If number of zeros prior to the first arrival is less than the chosen length of the STF then the length of the Green's function should reduced by the length of the STF.

The above algorithm has been found to have second order time complexity Time for single iteration depends on the step size μ . As expected it has been found that a low value of μ increases the time of convergence while keeping the error within bounds, whereas a higher value of μ leads to faster convergence, occasionally leading to instability. Therefore, μ has to be carefully chosen for optimal convergence of the adaptive algorithm. Fig. 3 shows the application of modified LMS method for deconvolution of signal 2 with convolutive output of signal 1 and signal 2, to recover signal 1. The figure also shows the total error in the estimation of signal 1 exhibiting good convergence.

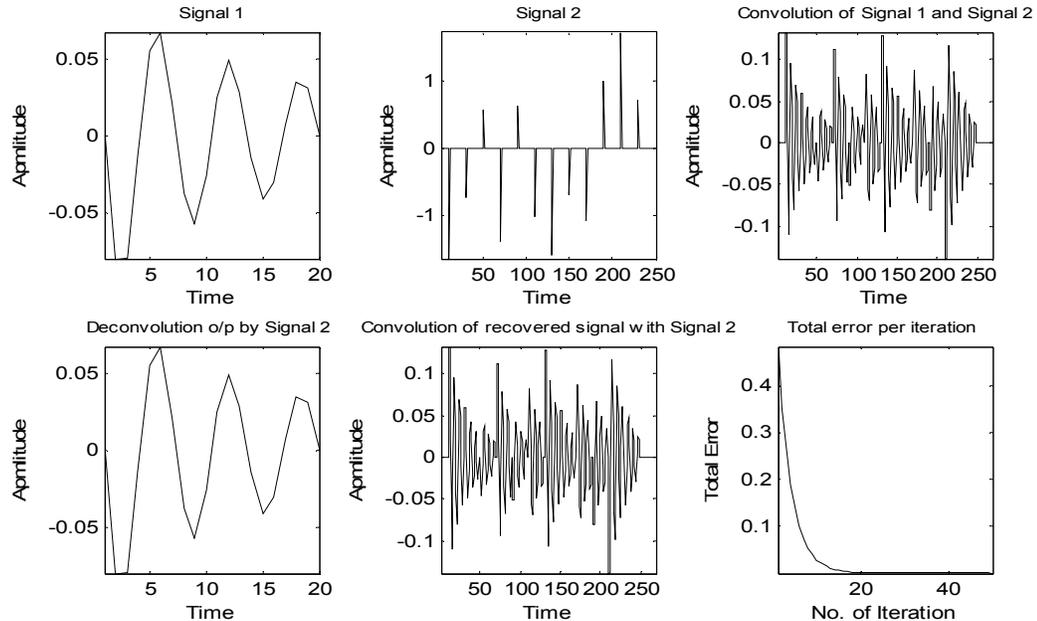


Figure 3. Application of Modified LMS method of deconvolution on synthetic data

5. Experimental Results and Discussion

5.1. Validation using Synthetic data

A synthetic seismogram is generated to validate the algorithms discussed in previous sections. Source wavelet is constructed as a “sinc” pulse and an impulse train is generated to simulate Green's function with Gaussian distribution. The source wavelet and the impulse train are convolved to generate the synthetic seismogram. Generated Source wavelet, Green's function and convolved output are shown in Fig. 4. Multi-pulse and frequency-domain blind

deconvolution method are employed to recover the Green's function. The order of the multi-pulse model is chosen as $P=30$ with number of pulses $1/10^{th}$ of total length of observed data for multi-pulse model. Method of generalized inverse and LMS are used to recover the source wavelet.

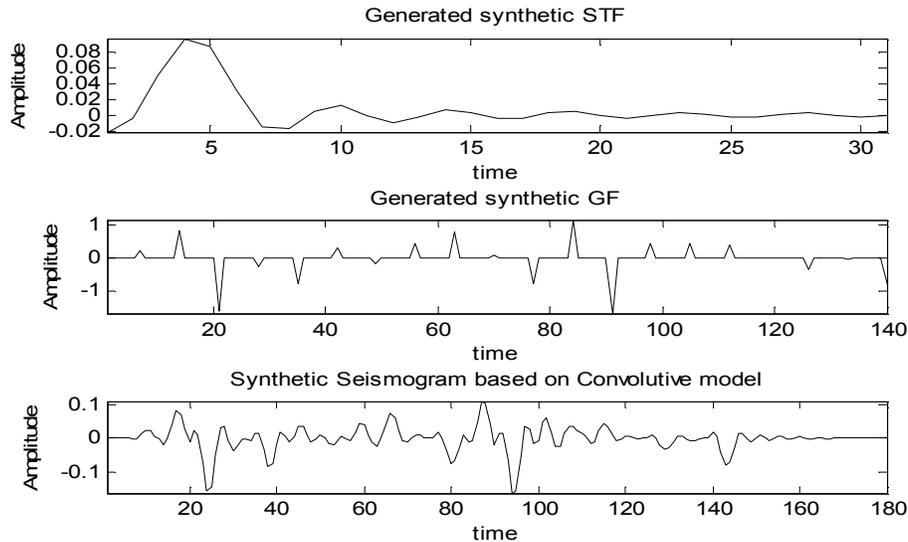


Figure 4. Generated STF, GF and Seismogram based on convolutional model.

5.2. Validation of Modified Multi-pulse Algorithm

Multi-pulse model is employed on the synthetic data with the 30^{th} order model, and as no convergence proof of method is available maximum number of iteration for *step 3* of Section 3 is set to be 10. Fig. 5 shows the recovered normalized GF with comparison to actual one using multi-pulse model. It shows that the position of impulses is estimated more or less correctly with a spread at its center value, while in the case of amplitude it's the normalized amplitude that is estimated correctly.

Source wavelet or STF has been estimated using modified LMS method as shown in Fig. 6 with total error per iteration and the parameters discussed in Section 4, chosen for the application is as follows

- Length of STF: 30.
- Step size μ : 0.01.
- Window is chosen depending on the length of the Green's function so as to satisfy the convolution model.

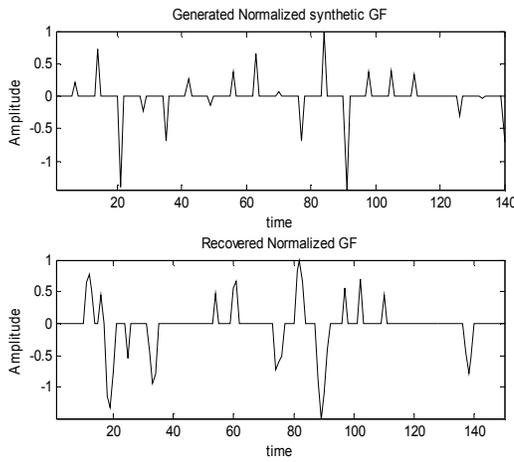


Figure 5. Generated normalized synthetic GF and recovered normalized GF using multi-pulse method.

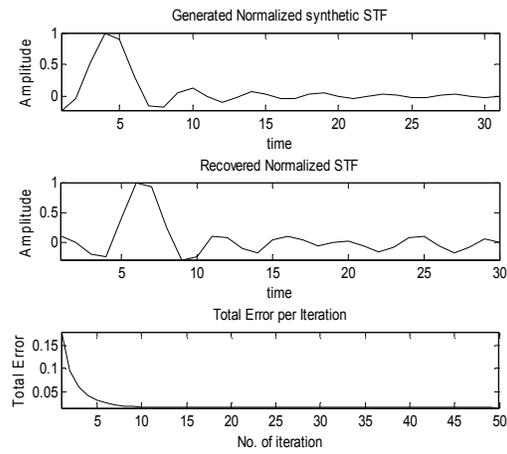


Figure 6. Generated normalized STF, recovered STF and total error per iteration using modified LMS method.

5.3. Validation of Frequency-Domain Blind Deconvolution

Frequency domain blind deconvolution is employed on the above synthetic data following algorithm given in Section 3. Fig. 7 shows the recovered normalized GF with the actual one. It shows that the recovered GF matches well with the actual one but a time shift is noticed during estimation process. Source wavelet or STF is estimated using modified LMS method with the same parameters as in the case of multi-pulse model. Fig. 8 shows estimated normalized STF with actual one and total error per iteration.

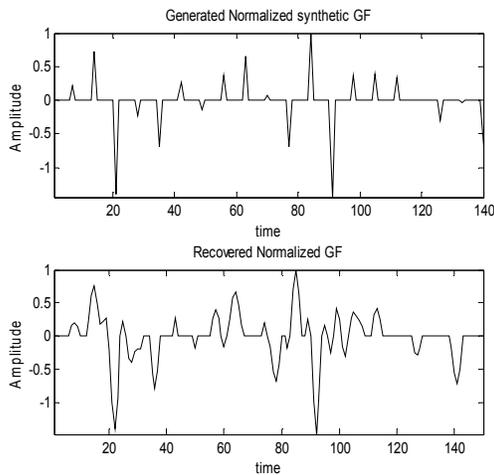


Figure 7. Generated normalized synthetic GF and recovered normalized GF using FDBD method.

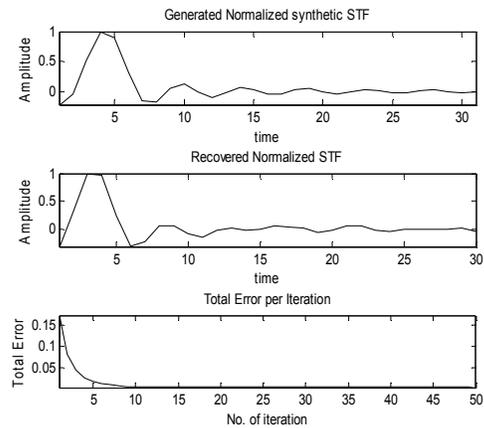


Figure 8. Generated normalized STF, recovered STF and Total error per iteration.

Thus the following points can be inferred by the application on the synthetic data. In multi-pulse model, the estimated GF has impulses with correct locations, while normalized amplitude is estimated correctly. Number of impulses is the important parameter and should be chosen carefully. In FDBD method, the estimated GF matches well with the actual one, but it is time shifted. It estimates correct length of GF, so the deconvolution using modified LMS will be easier as compare to multi-pulse model.

In the present study, December 26, 2004 Sumatra Earthquake ($M_w=9.3$) [37] downloaded from the *IRIS* data center is used to estimate the Green's function for chosen stations. The details of the chosen stations are shown in table. 1. All the seismogram of the event has been filtered using band pass filter with pass band frequency 0.008-0.1 Hz, and deconvolved with the instrument response to get the ground velocity. The horizontal components have been rotated using the back azimuth at the respected stations to transform the north and south components to radial and transverse components respectively. The order of the multi-pulse model is chosen to be $P=500$ in this case.

Table 1. List of the selected stations connected to *IRIS* data center.

S.N.	Station Name	Latitude	Longitude
1.	AAK	42.639	74.494
2.	CTAO	-20.0882	146.2545
3.	MAJO	36.5457	138.2041
4.	GUMO	13.5893	144.8684
5.	BRVK	53.0581	70.2828
6.	YSS	46.9587	142.7604
7.	OBN	55.1146	36.5674
8.	INCN	37.4776	126.6239
9.	ULN	47.8652	107.0528
10.	YAK	62.0308	129.6812
11.	FURI	8.8952	38.6798
12.	GNI	40.148	44.741
13.	LSZ	-15.2779	28.1882
14.	HNR	-9.4387	159.9475
15.	WAKE	19.2834	166.652

5.4. Extraction of STF and GF Using Modified Multi-pulse Model

Modified multi-pulse method is applied to the three components of displacement records retrieved by velocity record to obtain the reflectivity series or Green's function corresponding to the layered earth structure. The STF's are retrieved by deconvolution of the observed waveforms with the Green's function waveforms in time domain with the help of modified LMS method discussed in Section 5. Recovered GF's and STF's are shown in the Fig. 9 and Fig. 10 at different stations.

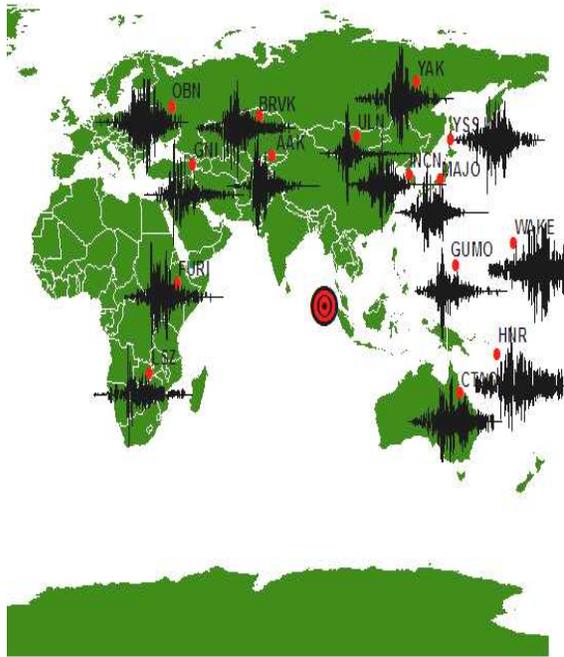


Figure 9. Estimated Green's functions from the vertical components of the respected stations using modified multi-pulse method.

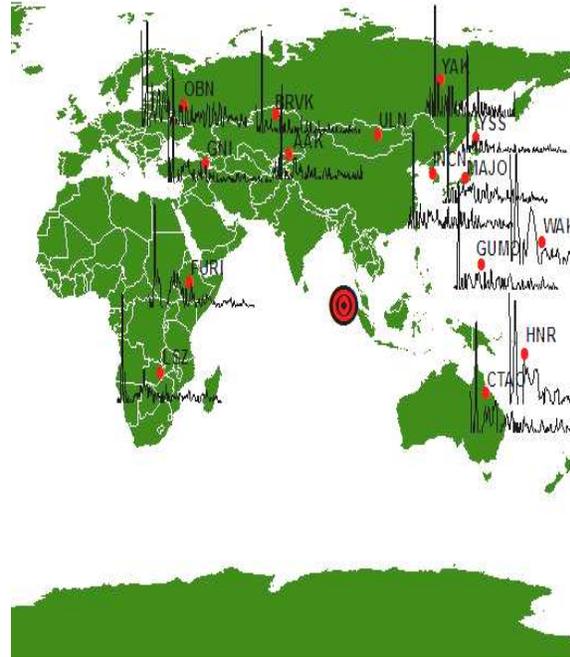


Figure 10. STF of the Sumatra earthquake at the recording stations on the map using modified multi-pulse method.

To get the better picture of rupture process all the recovered STF's are averaged and shown in Fig. 11.

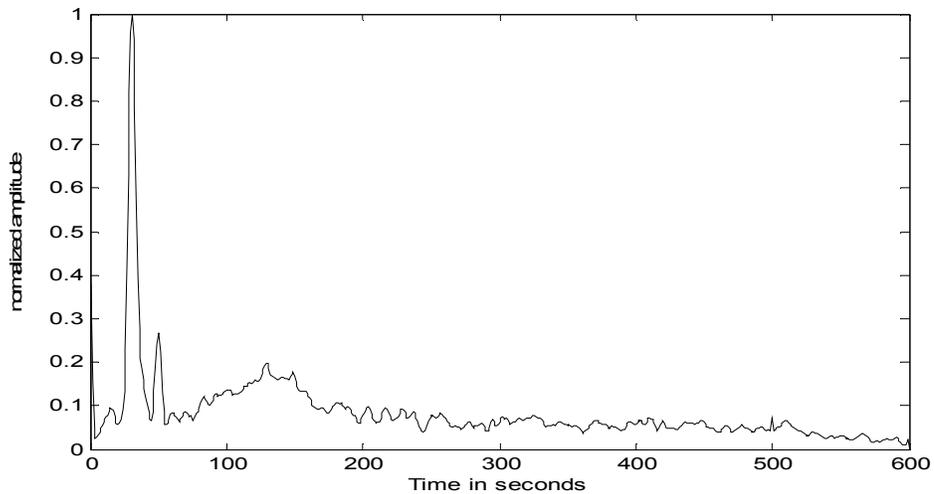


Figure 11. Estimated average STF of the Sumatra earthquake, using modified multi-pulse method.

Parameters used in the above methods are

- Length of STF is taken as 600, in agreement with the time length given in literature.
- Step size μ is taken to 0.01 in LMS method.
- Window chosen from the Green's function of appropriate length to satisfy the convolution model centered at midpoint of the recovered GF.
- Number of impulse is chosen as $1/5^{th}$ of the total length of STF in multi-pulse model.

5.5. Extraction of STF and GF using Frequency Domain Blind Deconvolution

Frequency domain blind deconvolution is employed on the above collected earthquake data following algorithm given in Section 3. Fig. 12 shows the recovered normalized GF at different stations. Recovered STF's are shown in Fig. 13 using modified LMS method with the same parameters as in the case of multi-pulse model. The key point in this method is to initialize filter G discussed in Section 3.

To get the better picture of rupture process all the recovered STF's are averaged and shown in Fig. 14.

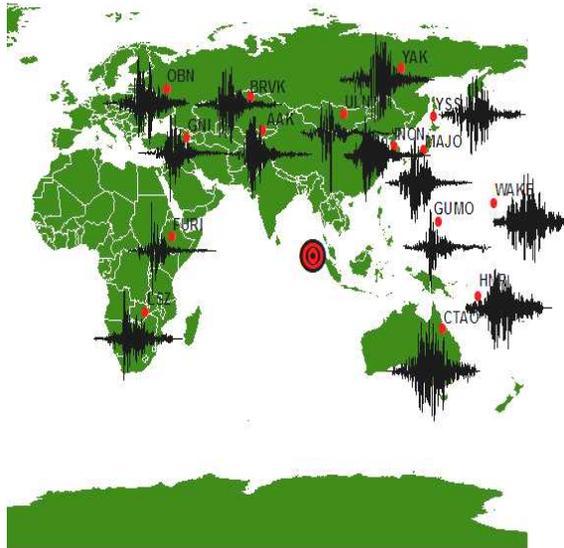


Figure 12. Estimated Green's functions from the vertical components on the respected stations using FDBD method.

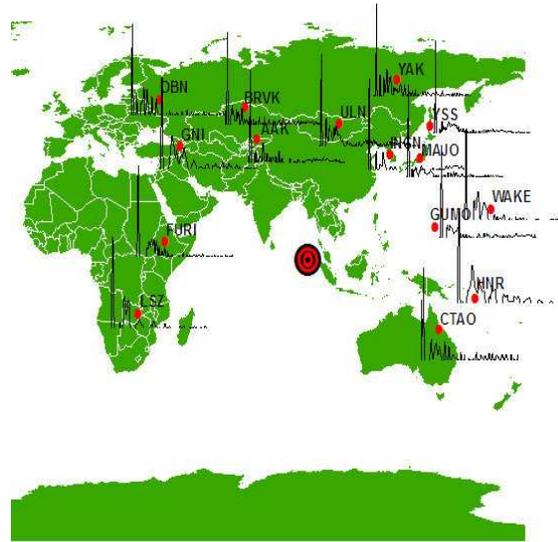


Figure 13. STF of the Sumatra earthquake at the recording stations on the map using FDBD method.

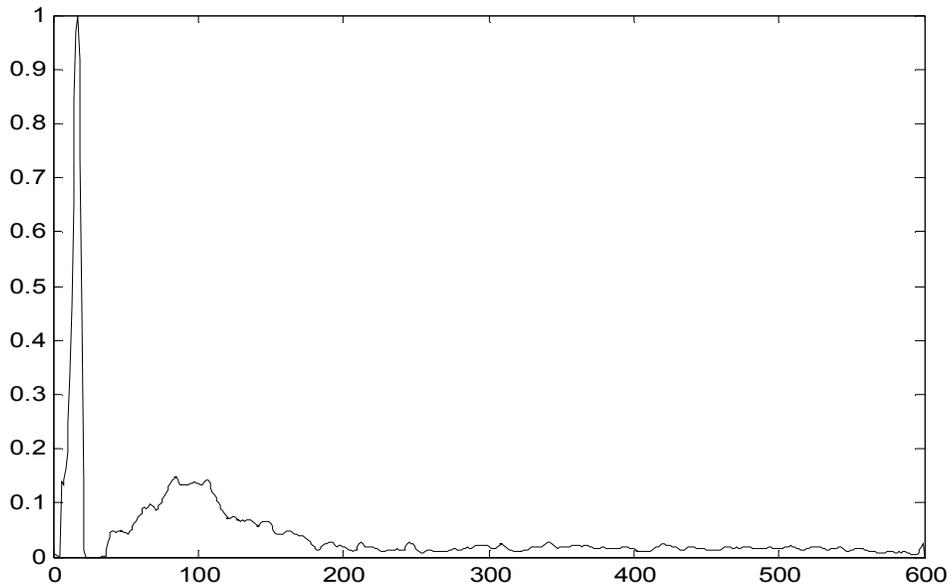


Figure 14. Estimated average STF of the Sumatra earthquake, using FDBD method.

6. Conclusions

This paper discusses the advantages of blind deconvolution as a technique to recover the GF and STF from the observed seismogram. The procedure to obtain GF using blind deconvolution is addressed, considering the convolutive models of seismogram. Some improvement has been introduced in multi-pulse model to take advantage of it in earthquake seismology. LMS method is also modified for use in deconvolution and its validation with synthetic data shows good results. In the validation of blind deconvolution with synthetic data we use the two methods, the multi-pulse model and the frequency domain blind deconvolution, in conjunction with modified LMS method. The result shows that both the methods work well with acceptable total error for convolutive models. After validation, both the methods are used for estimation of source time function from the earthquake records of 26 December 2004 Sumatra earthquake from worldwide stations connected to *IU* network of *IRIS* data center. The STFs recovered on different stations using the multi-pulse model and the FDBD method show similarity in shape and duration as well, although the principles of both the methods are entirely different. That is, the application of both the methods on these data sets shows that the retrieved STF's are in good agreement with surface wave time function discussed in the literature. Blind deconvolution is thus found to be a potential method for recovering GF and STF from the observed seismogram, without any constraint on *a priori* information about the source or station.

References

- [1.] M. Cooke, H. J. Trussell, I. J. Won, "Seismic deconvolution by multipulse methods" IEEE Trans. Acous. Speech, Signal Process. 38, 1990, pp. 156-160.
- A. Taleb, J. Solé, Casals, and C. Jutten, "Quasilinear blind inversion of Wiener systems", IEEE Trans. Signal Process., vol. 49, no. 5, May 2000, pp. 917-924,
- [2.] Larue, Jérôme I. Mars, and C. Jutten, "Frequency-Domain Blind Deconvolution Based on Mutual Information Rate" IEEE Trans. Signal Process, vol. 54, no. 5, May 2006, pp.1771-1781.
- [3.] R. Madariaga, Earthquake source theory: A review, In Kanamori H and Boschi E (eds), Earthquake: Observation, Theory and Interpolation, Soc Ital di Fisica Bologna, 1983, Italy, 1-14.
- [4.] Y. T. Chen, X. F. Chen and L. Knoppof, Spontaneous growth and autonomous contraction of two-dimensional earthquake fault, In: Wesson R I (ed) Mechanics of Earthquake Faulting. Tectonophysics, 144, 5-7, 1985.
- [5.] S. Dasgupta, R. L. Nowack, "Deconvolution of Three-Component Teleseismic P Waves Using the Autocorrelation of the P to SV Scattered Waves", Bull. Seism. Soc. Am., 96, 2006, pp. 1827-1835.
- [6.] D. V. Helmberger, D. M. Hadley, "Seismic source functions and attenuation from local and teleseismic observations of the NTS events Jorum and Handley", Bull. Seism. Soc. Am. 71, 1981, pp. 51-67.
- Sèbe, P. Y. Bard and J. Guilbert J, "Extracting time-domain Green's function estimates from ambient seismic noise", Geophys. Res. Lett. 32, L03310 Doi 10.1029/2004GL021862, 2005.
- B. W. Stump, L. R. Johnson, "The determination of source properties by the linear inversion of seismograms", Bull. Seism. Soc. Am. 67, 1977, pp. 1488-1502.
- [7.] L. S. Zhao, D. V. Helmberger, "Source estimation from broadband regional seismograms". Bull. Seism. Soc. Am., 84, 1994, pp. 92-104.
- [8.] M. G. Bostock, "Green's functions, source signatures, and the normalization of teleseismic wave fields", J. Geophys. Res., 109:B03303, doi 10.1029/2003JB002783, 2004.
- [9.] D. M. Boore, "Stochastic simulation of high-frequency ground motion based on seismological models of radiated spectra", Bull. Seism. Soc. Am., 73, 1983, pp. 1865-1884.
- [10.] M. Bouchon, "A simple method to calculate Green's functions for elastic layered media", Bull. Seism. Soc. Am. 71, 1981, pp. 959-941.
- [11.] T. C. Hanks, R. K. McGuire, "The Character of high-frequency strong ground motion", Bull. Seism. Soc. Am., 71, 1981, pp. 2071-2095.
- [12.] K. G. Sabra, P. Gerstoft, P. Roux, W. A. Kuperman, M. C. Fehler, "Extracting time-domain Green's function estimates from ambient seismic noise", Geophys. Res. Lett., 32:L03310, doi 10.1029/2004GL021862, 2005.
- A. M. Dziewonski and J. H. Woodhouse, "An experiment in systematic study of global seismicity: Centroid-moment tensor solutions for 201 moderate and large earthquakes of 1981", J. Geophys. Res, 88, 1983, pp. 3247--3271.
- B. W. Stump, L. R. Johnson, "Near-field source characterization of contained nuclear explosions in tuff", Bull. Seism. Soc. Am., 74, Feb. 1984, pp. 1-26.
- [13.] J. Ritsema and T. La, "Rapid source mechanism of determination of large ($M_w > 5$) earthquakes in the western United States", Geophys. Res. Lett., 20, 1993, pp. 1611-1614.
- [14.] S. A. Sipkin, "Estimation of earthquake source parameters by the inversion of waveform data, synthetic seismogram", Phys. Earth Planet. Inter., 30, 1982, pp. 242-259.
- [15.] S. A. Sipkin, R. E. Needham, "Moment-tensor solutions estimated using optimal filter theory: global seismicity, 1990", Phys. Earth Planet Int. 70, 1992, pp. 16-21.
- [16.] S. A. Sipkin, and L. A. Lerner-Lam, "Pulse-shape distortion introduced by broadband deconvolution", Bull. Seism. Soc. Am., 82, 1992, pp. 238-258.
- [17.] S. H. Hartzell and T. H. Heaton, "Teleseismic time functions for large shallow Subduction zone earthquakes", Bull. Seism. Soc. Am, 75, 1985, pp. 965-1004.
- [18.] Muelle, "Source pulse enhancement by deconvolution of an empirical Green's function", Geophys. Res. Lett. 12, 1985, pp. 33-36.
- A. A. Velesco, C. J. Ammon, T. Lay, "Empirical Green's function deconvolution of broadband surface waves: rupture directivity of the 1992 Landers California ($M_w=7.3$) earthquake", Bull. Seism. Soc. Am, 84, 1994a, pp. 735-750.
- [19.] A. Velesco, C. J. Ammon, T. Lay, "Recent Large earthquakes near Cape Mendocino and in the Gorda plate: Broadband source time functions, fault orientation and rupture complexity", J Geophys Res. 99, 1994b, pp. 711-728.
- [20.] J. Ammon., A. A. Velesco, T. Lay, "Rapid estimation of rupture directivity: Application to the 1992 Landers ($M_w = 7.3$) and Cape Mendocino ($M_w = 7.3$) California earthquakes", Geophys Res. Lett. 20, 1993, pp. 97-100.

- [21.]J. Ammon, T. Lay, A. A. Velesco, E. J. Vidale, "Routine estimation of earthquake source complexity: The 18th October 1992 Columbia earthquake". *Bull. Seism. Soc. Am*, 84, 1994, pp. 1264-1271.
- [22.]S. H. Hartzell, "Earthquake aftershocks as Green's function", *Geophys Res Lett*, 5, 1978, pp. 1-4.
- [23.]Y. T. Chen, J. Y. Zhou and J. C. Ni J, "Inversion of near-source broadband accelerograms for source time function", *Tectonophysics*, 197, 1991, pp. 89-98.
- [24.]S. Dreger, "Empirical Green's function study of the January 17, 1994 Northridge, California earthquake", *Geophys. Res. Lett.*, 21, 1994, pp. 2633-236.
- [25.]L. S. Xu and Y. T. Chen, "Source time function of Gonghe, China earthquake, retrieved from long-period digital waveform data using Empirical Green's function technique", *Acta Seismologica sinica*, 9(2), 1996, pp. 209-222.
- [26.]T. Pham, "Contrast functions for blind separation and deconvolution sources", In *Proc. Int. Conf. Independent Components Analysis (ICA)*, San Diego, CA, Dec. 2001, pp. 37-42.
- [27.]T. Pham, P. Garat and C. Jutten, "Separation of a mixture of independent sources through a maximum likelihood approach", In *Proc. Eur. Signal Processing Conf. (EUSIPCO)*, vol. 2, Brussels, Sep. 1992, pp. 771-774.
- [28.]Cover, T., and J. Thomas, *Elements of Information Theory*, ser. Wiley Series in Telecommunications. New York: Wiley, 1991.
- [29.]Zhang, L., and Cichocki, A., Multichannel blind deconvolution of nonminimum-phase systems using filter decomposition. *IEEE Trans. Signal Process.*, vol. 52, no. 5, 1430-1441, May 2004.
- [30.]Comon, P., Independent component analysis, a new concept? *Signal Process.*, vol. 36, no. 3, pp. 287-314, Apr. 1994.
- [31.]T. La, "The Great Sumatra-Andaman Earthquake of 26 December 2004", *Science* 308, 2005, pp. 1127-1133.
- A. Taleb, and C. Jutten, "Source separation in post nonlinear mixtures", *IEEE Trans. Signal Process.*, 47, no. 10, Oct. 1999, pp. 2807-2820.
- [32.]M. Gaeta, and J. L. Lacoume, "Source separation without a priori knowledge: The maximum likelihood solution", In: Torres, Masgrau and Lagunas, eds., *Proc. EUSIPCO Conf.*, Barcelona, Elsevier, Amsterdam, 1990, pp. 621-624.
- [33.]Lawson, C., R. Hanson, *Solving Least Squares Problems*. Prentice-Hall, Englewood Cliffs, NJ, 1974.
- [34.]N. Charkani, and Y. Deville, "Optimization of the asymptotic performance of time-domain convolutive source separation algorithms", In *Proc. ESANN*, Bruges, Belgium, Apr. 1997, pp. 273-278, Apr.
- [35.]H. H. Yang, S. Amari and A. Cichocki, "Information back-propagation for blind separation of sources from nonlinear mixture", In *Proc. ICNN*, Houston, TX, 1996.