Bulletin of the Seismological Society of America, Vol. 86, No. 4, pp. 1187–1193, August 1996

On the Use of Simulated Annealing Method and Cross-Validation Theory for Deconvolution of Seismograms

by F. Courboulex, J. Virieux, and D. Gibert

Abstract In order to retrieve the apparent source time function (ASTF) seen at a given station, one must take into account propagation effects, site, and instrumental influences. Removing these effects can be performed by a deconvolution of the mainshock seismogram by a seismogram of a smaller event recorded at the same station. This smaller event must occur nearby the mainshock, and the associated seismogram is considered as an empirical Green's function. We propose a deconvolution based on the simulated annealing method, and we compare it with the often-used spectral division technique.

We show on both synthetic and real signals that the simulated annealing deconvolution (SAD) provides stable and positive ASTF, whereas results from spectral division are very sensitive to an *ad hoc* parameter called water level.

Finally, the application of cross-validation analysis between the three components of the seismogram in addition to the SAD allows us to estimate errors on the ASTF.

Introduction

A seismogram recorded at a particular station contains the effects of the rupture process, propagation inside the Earth, and the site and instrumental response. Removing the path, site, and instrumental effects enables one to recover the apparent source time function (ASTF) seen at each station. This ASTF reflects the temporal evolution of the rupture process, i.e., the source time function (STF), seen by this station. When considering a point source, STF and ASTF are the same. Differences appear considering spatiotemporal propagation of the rupture on the fault plane.

A seismogram of a smaller event located near the event of interest can be used as an empirical Green's function (EGF) (Hartzell, 1978). If the focal mechanisms of both events are similar, deconvolution of the larger event using the smaller one may be used to recover the ASTF at each station (Mueller, 1985). Moreover, if the stations are well distributed in azimuth around the epicenter, the ASTFs obtained at each station may be used to construct the STF and to study the spatiotemporal rupture processes along the fault plane (Mori and Hartzell, 1990; Courboulex *et al.,* 1996).

A key element in such studies is a stable deconvolution method that provides a well-constrained ASTF. Spectral division is frequently used in most cases because of its computer speed and efficiency (Ammon *et al.,* 1994; Velasco *et al.,* 1994). However, this method is prone to instability due to the division by negligible coefficients in the Fourier transform of the small event.

Applying a "water level" correction smoothes the spectrum of the small event signal and regularizes the unstable deconvolution (Helmberger and Wiggins, 1971). Using this additional parameter, one can obtain an ASTF for each good-quality seismogram, but the interpretation of the signal is still difficult: in particular, the deduced ASTF contains negative values while its shape and amplitude is influenced by the chosen "water level." Estimation of uncertainties of the ASTF is an open problem.

We propose an alternative deconvolution method based on the simulated annealing method that might be called the simulated annealing deconvolution (SAD) method.

We first describe the SAD technique and compare it with the spectral division technique using synthetic data. Then, we discuss uncertainty estimations based on cross-validation techniques and finally present an application on real three-component data for P and S waves.

Simulated Annealing Deconvolution

Our goal is to estimate ASTF time series and, more precisely, the amplitude of each point of the ASTF. The basic steps of a global search method as applied to deconvolution are described in Figure 1: After discretization of the model space, an initial ASTF is guessed; convolution is performed between the ASTF and the empirical Green function; then the synthetic result is compared with the observed signal using a given misfit function. If the result is unsatisfactory, another iteration is started by repeating the direct modeling. The choice of a new model depends on the selected algorithm. Among various methods, we have used the simulated

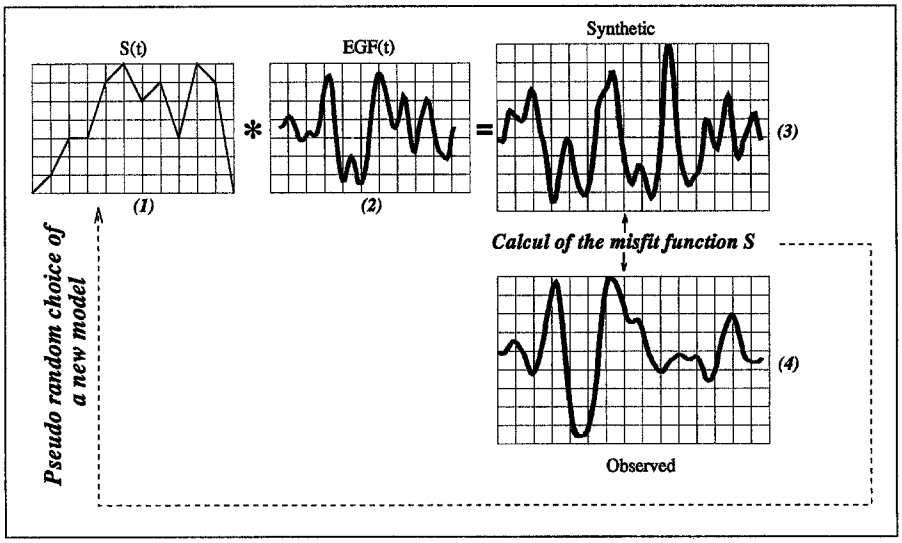

Figure 1. Principle of deconvolution of two signals by a global search method: (1) the selected apparent source time function at a given iteration, (2) the empirical Green's function, (3) the result of convolution of signals (1) and (2), and (4) the seismogram we want to obtain.

annealing method that readily allows to impose positivity constraint and provides a solution quite rapidly.

The simulated annealing algorithm was proposed by Kirkpatrick *et al.* (1983) and Cerny (1985) and relies on the work of Metropolis *et al.* (1953). Like physical annealing, which produces crystalline assemblies with very low internal energy, simulated annealing aims to produce models (i.e., parameter configurations) with low misfit function (i.e., energy). This is achieved by performing some "guided random walk" in the model space. The guidance, which boosts the convergence toward the solution, is driven by the misfit function. During iterations, the random walk allows models with a misfit function not always lower than the preceding ones, permitting the acceptance of "worse" models that may be on the path to escape from local minima of the misfit function. The balance between the guided and random behavior of the walk is controlled by a parameter called the "temperature." The procedure starts at a sufficiently high temperature for the walk to be almost purely random. Then, while walking in the model space, the temperature is lowered until the walk is so "guided" that it becomes similar to that of a pure gradient descent.

In the present study, we have implemented the algorithm by embedding a heat bath looping [i.e., a walk at constant temperature, see Creutz (1980)] into an adaptive cooling schedule designed by Huang *et al.* (1986). A detailed description of this procedure is given by Gibert and Virieux (1991).

For real data, the highest frequency we can consider for the ASTF is the corner frequency of the small event used as an EGF. The SAD sampling frequency is then set to be equal to two times this value. Discretization of amplitude mainly depends on the final precision result one wants to reach. Choosing a coarse discretization provides a quick result, whereas a denser sampling defines a more refined result at the expense of computer time.

Synthetic Illustration

We first test the SAD method on a numerical example already studied by Zollo *et al.* (1995) and compare the results with those of two other methods: spectral division and simplex deconvolution.

We consider two earthquakes that occurred at almost the same location but with a different STF. As we use in this example a point source model, STF and ASTF have the same meaning. The bigger earthquake record is obtained by convolution of a complex STF with a synthetic Green's function (Fig. 2, right side). The smaller event record, used as an empirical Green's function, is obtained by convolution of a simple triangular STF with a Green's function slightly different to that of the main event (Fig. 2, left side). The difference between the two Green's functions has been introduced to consider possible small differences in location between the two events. To both signals obtained by convolution, we add zero-mean Gaussian noise, whose amplitude is some fraction of the amplitude of the first pulse of the EGF. This provides more realistic seismograms (Fig. 3). Using these seismograms without any further filtering, we perform deconvolution of the main event seismogram with the EGF using both spectral division and SAD technique.

Figure 4 shows results obtained by spectral division with different values of the water level. Clearly, the shape and amplitude of the STF are influenced by the choice of water level value. Furthermore, if no water level is imposed, the deduced STF is very unstable, and if the water level is

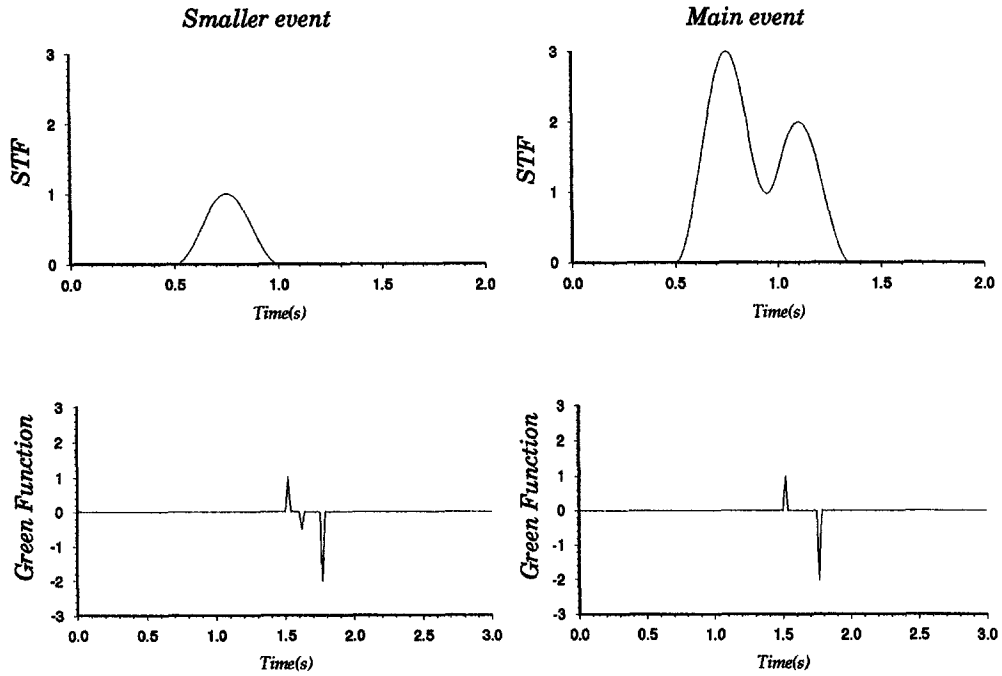

Figure 2. STF (top) and synthetic Green's function (bottom) used to create synthetic seismograms of the mainshock (right side) and the EGF (left side).

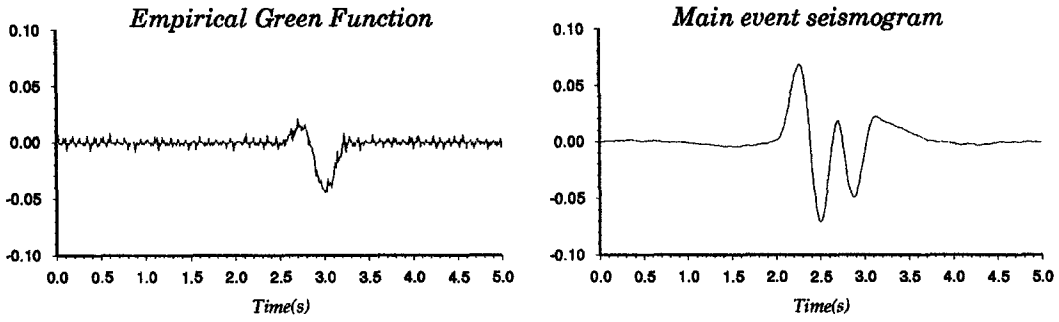

Figure 3. Synthetic seismograms of the smaller earthquake used as an EGF (left) and the main event (right). A white noise has been added to both signals.

too high, a large number of spectral values are deleted before division, and the signal is highly modified. Retaining a water level value of around 10 to 20 dB, we observe that the result obtained is not always positive, and so measuring the amplitude of the STF becomes difficult.

Figure 5 presents the result obtained using the SAD technique with 30 iterations, a discretization equal to 30 points in amplitude and 100 points in time. The global shape and the amplitude of the result is close to that obtained by spectral division (with a reasonable choice of water level) with the advantage of being very stable and always positive. The result seems not to be affected by noise, except for the very small spikes of the STF seen in Figure 5. Our results are in good agreement with the results of Zollo *et al.* (1995).

Uncertainties by Cross Validation on Real Three-Component Seismograms

The cross-validation method (Stone, 1974; Peterson and Davey, 1991) allows us to estimate the noise variance of the data. For this purpose, we consider the three components of the signal independently. The SAD method is first performed on the vertical component of the signal, obtaining the best-fitted ASTF. Next, the misfit function is calculated with this solution for the other two horizontal signals (north–south and east–west components). This is then repeated for the two other components thus obtaining, for a given signal, three values of the misfit function. The lowest value corresponds to the best model for this component while the two other

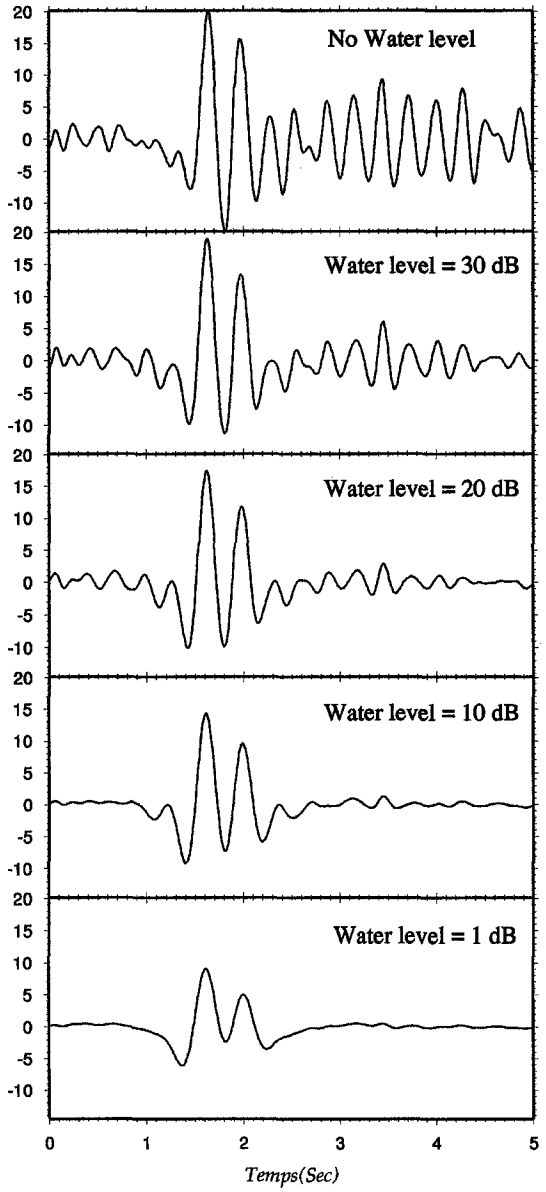

Figure 4. Deconvolution of the mainshock by the EGF obtained by the spectral division technique. Results are shown for different water level values.

ones estimate the independent quality of the fit for the two other components. Those two misfit values are called cross values. From all of these values, we obtain a matrix of nine values with three minimal values along the diagonal and six other cross values. These quantities are related to the global noise variance σ^2 of our data:

$$\sigma^2 = \frac{|\bar{V}_c|}{N}, \tag{1}$$

where N is the number of points of the observed signal and V_c are the cross values. Then we perform a global SAD on the three components together; the misfit function is chosen as

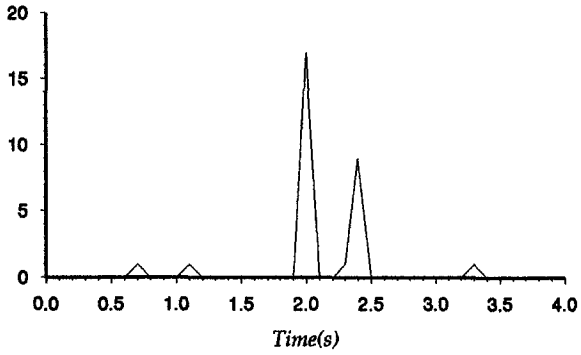

Figure 5. Deconvolution of the mainshock by the EGF obtained by the simulated annealing method.

$$C = \sum_{i=1}^{I} [(S_V(i) - O_V(i)]^2 + [S_{NS}(i) - O_{NS}(i)]^2 + [S_{EW}(i) - O_{EW}(i)]^2,$$

where I is the number of points in time, $O(i)$ represents the observed signal, and $S(i)$ is the synthetic signal for each component. During the global procedure, instead of decreasing the temperature to zero, we stop it at a critical value T_c equal to the noise variance calculated from the cross values (equation 1).

At that critical temperature, we perform a large number of iterations and retain the complete set of models. The distribution of these models is used to bind the final solution. The selected solution is the average of those models, and the estimation of uncertainties is equal to the standard deviation.

Data Application

The aim of deconvolution by empirical Green's functions on real data is to recover the ASTF at a given station.

Consider a line source with a unilateral rupture propagation. If θ is the angle between the station and the rupture directivity, V_R is the rupture velocity, and c is the velocity of the studied wave, then the apparent duration ΔT of the ASTF at this station can be written as

$$\Delta T = T_R\left(1 - \frac{V_R \cos\theta}{c}\right), \tag{2}$$

where T_R is the STF duration. This equation shows that the directivity effect depends on the location of the station with respect to the source directivity and also on the ratio between V_R and c. The directivity effect is more important on S waves than on P waves. The accuracy of ASTF determination on several stations and for different types of waves when possible is key in order to estimate the directivity of the seismic rupture process.

An example of the importance of deconvolution is shown here on real three-component data from the region of

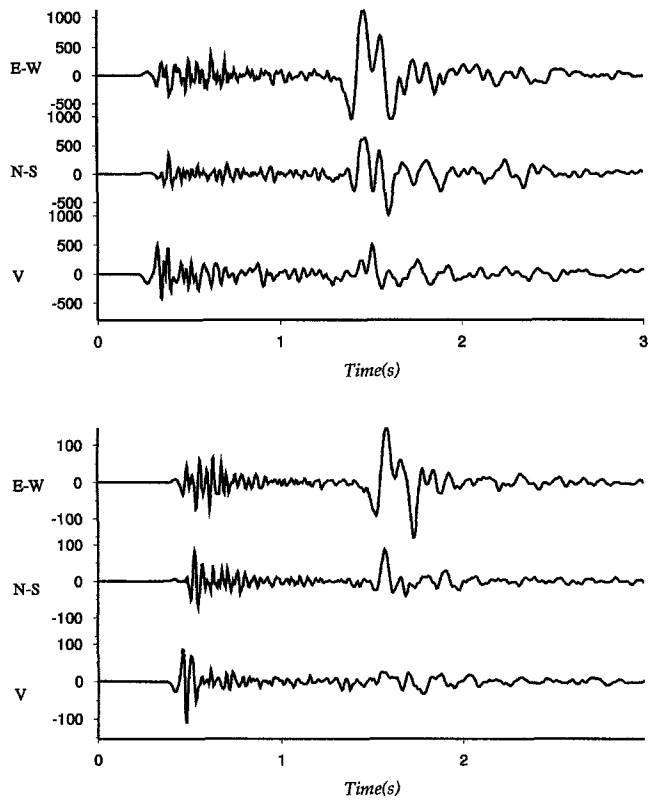

Figure 6. Seismograms of the bigger earthquake (top) and the smaller one (bottom) used as an empirical Green's function (data of Campi Flegrei, Italy).

Campi Flegrei (Italy) (De Natale *et al.*, 1987) for which careful analysis of seismic source is necessary in order to understand the seismic crisis that takes place in this volcanic caldera.

We select two earthquakes located at the same place, within location error bounds, with the same focal mechanism and recorded by the same station (Fig. 6). The smaller event ($Md = 0.9$) is considered as an empirical Green's function for the larger one ($Md = 1.9$). The first deconvolution applied on P and S waves is obtained by a spectral division (Fig. 7). The selected "water level" is set to 20 dB for each of the six signals.

These results highlight three different problems, the first of which is the zero level. Since spectral deconvolution provides both positive and negative values, and because the negative values obtained cannot be neglected, determination of the final duration of the principal pulse is problematic. The second problem is that there are clear differences between the ASTF obtained with the three components of the signal. Finally, what confidence should be accorded to the large number of second pulses obtained by deconvolution?

The results obtained by the SAD method are shown in Figure 8 for P waves and in Figure 9 for S waves. Each figure presents (a) the ASTF obtained (bold line) with its uncertainties (dashed lines) and (b) the observed signal and the synthetic signal obtained by convolution of the mean solution with the EGF.

Considering the uncertainties estimation, we can see

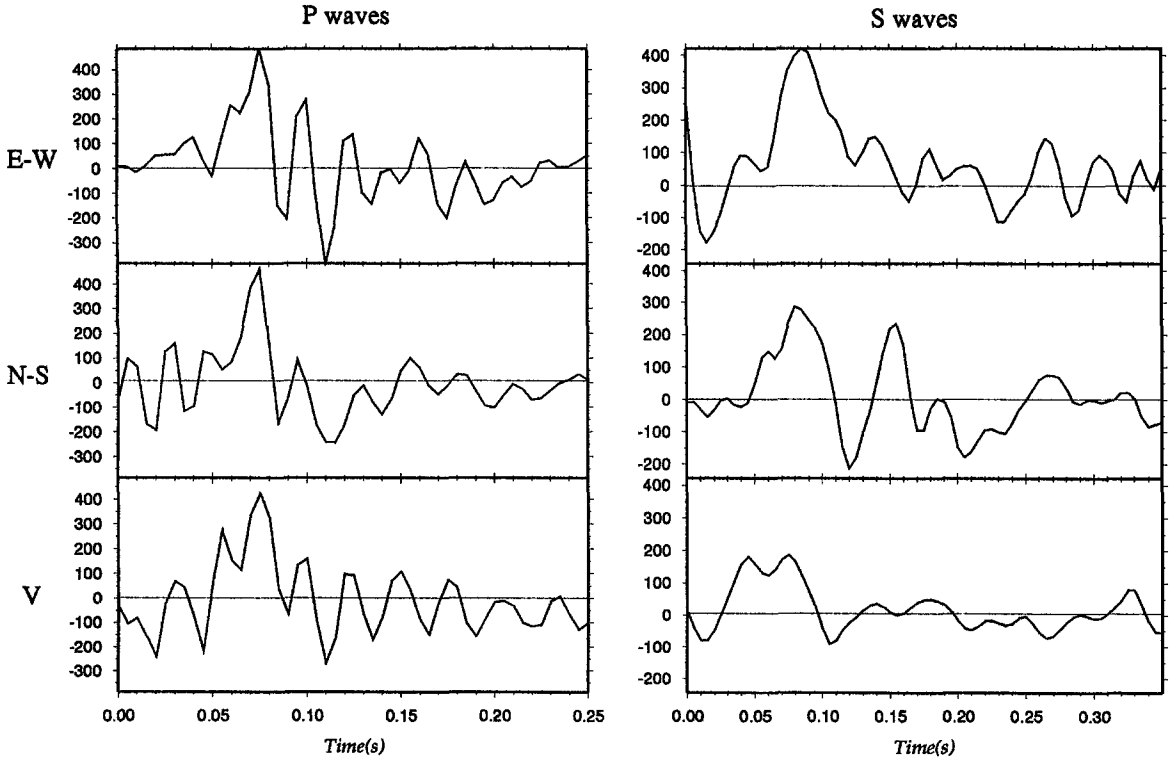

Figure 7. Deconvolution obtained by spectral division of the three-component seismograms of Campi Flegrei for P and S waves.

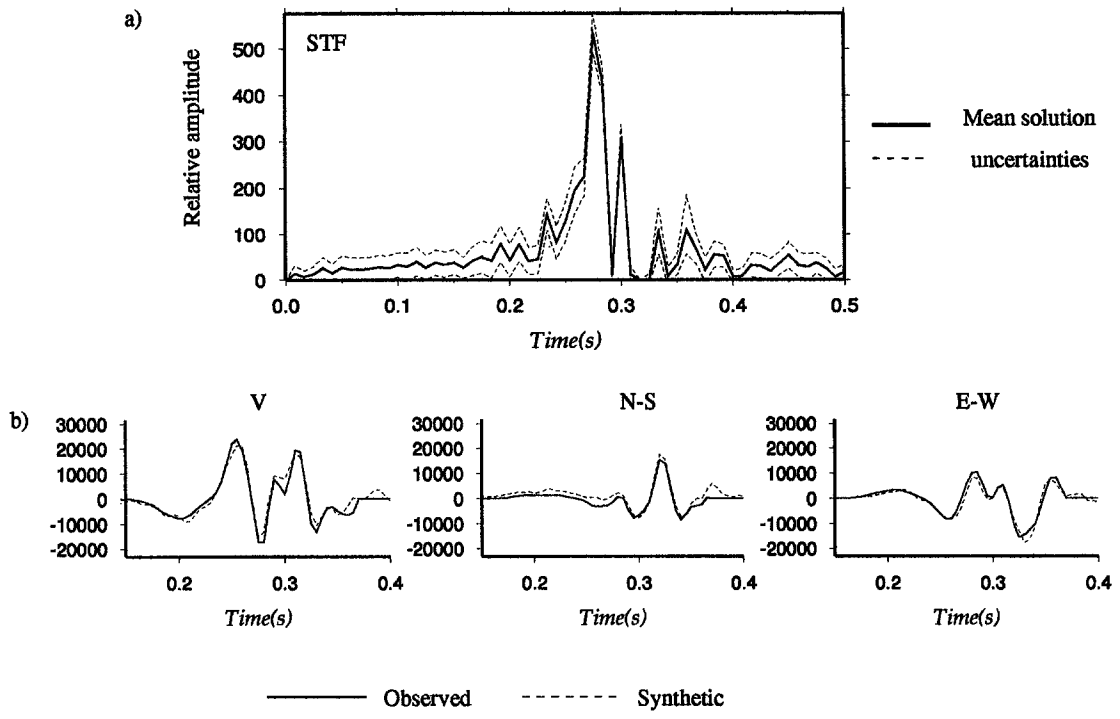

Figure 8. (a) The mean ASTF found by the SAD method is drawn (bold line) with its uncertainties (dashed lines) for *P* waves. (b) Observed and synthetic seismograms obtained by convolution of the mean ASTF with the three components of the EGF.

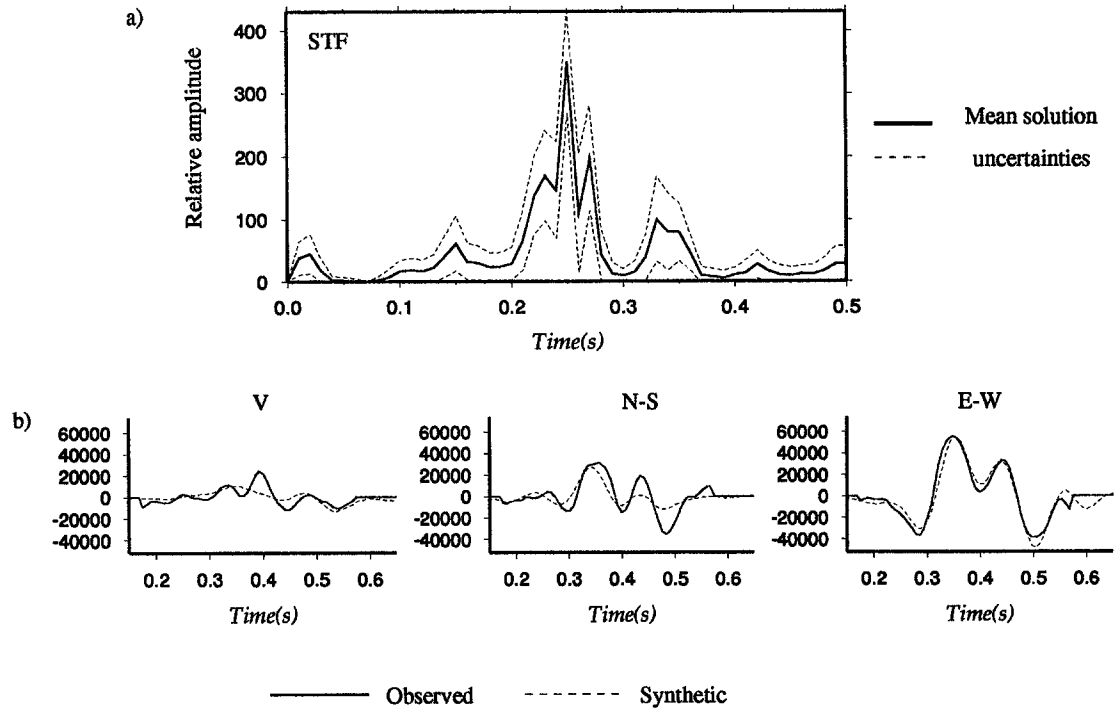

Figure 9. (a) The mean ASTF found by the SAD method is drawn (bold line) with its uncertainties (dashed lines) for *S* waves. (b) Observed and synthetic seismograms obtained by convolution of the mean ASTF with the three components of the EGF.

that the ASTF is predominantly represented by a main pulse duration of 0.08 sec for the *P*-wave ASTF and 0.1 sec for *S*-wave ASTF. A second pulse can also be determined on the two ASTFs, but the amplitude is very small. Looking at these results, we can deduce that the rupture propagates away from the station because the ASTF main pulse is larger using *S* waves rather than *P* waves.

An accurate determination of ASTF is a difficult task. The main problem is to be sure that all the propagation effects have been considered in the Green's function. Using a small event as an empirical Green's function removes most of these problems. Nevertheless, a perfect empirical Green's function is impossible to find, as small differences in the event locations always exist. Then, the ASTF obtained must always be looked at carefully. This notwithstanding, the SAD method greatly reduces numerical errors of the obtained ASTF and allows the investigation of source directivity of moderate earthquakes.

Conclusions

Deconvolution of a seismic signal by empirical Green's functions can be an unstable process. The most frequently used method, spectral division, has the advantage of being very fast, but no positivity constraint can be performed, and no error estimations for the overall result of the three components of the signal can be made.

An alternative approach based on the simulated annealing technique provides a stable and positive source time function by simultaneous deconvolution of the three components at a given station, even if the data are noisy. The effect of the noncoherent noise between components on the ASTF is strongly reduced. Moreover, an analysis by the cross-validation method allows error estimations to be made. Thus, the ASTFs may be used more confidently in order to retrieve kinematic source parameters such as rupture directivity or slip distribution on the fault plane.

The spectral division approach remains a very attractive method because of its efficiency when one wants to produce routine analysis of data sets, but for deeper analysis, especially on moderate events, we think that the SAD method is worth the effort to apply it.

Acknowledgments

We are grateful to Aldo Zollo for his help and for providing us synthetic data. Thank you to Greg McIntosh for his careful reading. We thank the partial support of the Institut National des Sciences de l'Univers through the program PNRN 1994. Publication 4 Géosciences Azur.

References

Ammon, C., T. Lay, A. Velasco, and J. Vidale (1994). Routine estimation of earthquake source complexity: the 18 October 1992 Colombian earthquake: *Bull. Seism. Soc. Am.* **84,** 1266–1271.

Cerny, V. (1985). A thermodynamical approach to the travelling salesman problem: *J. Optimisation Theory Appl.* **45,** 41–51.

Courboulex, F., J. Virieux, A. Deschamps, D. Gilbert, and A. Zollo (1996). Source investigation of small events using empirical Green functions and simulated annealing: *Geophys. J. Int.,* **125,** in press.

Creutz, M. (1980). Monte Carlo study of quantized SU(2) gauge theory, *Phys. Rev.* **21,** 2308–2315.

De Natale, G., G. Ianaccone, M. Martini, and A. Zollo (1987). Seismic sources and attenuation properties at the Campi Flegrei volcanic area, *Pageoph* **125-6,** 883–917.

Gibert, D. and J. Virieux (1991). Electromagnetic imaging and simulated annealing, *J. Geophys. Res.* **96,** 8057–8067.

Hartzell, S. (1978). Earthquake aftershocks as Green's functions, *Geophys. Res. Lett.* **5,** 99.

Helmberger, D. and R. Wiggins (1971). Upper mantle structure of the midwestern United States, *J. Geophys. Res.* **76,** 3229–3245.

Huang, M., F. Romeo, and A. Sangiovanni-Vincentelli (1986). An efficient general cooling schedule for simulated annealing, Santa Clara, *Proc. IEEE Int. Conf. Computer-Aided Design,* 381–384.

Kirkpatrick, S., J. Gellat, and M. Vecchi (1983). Optimization by simulated annealing, *Science* **220,** 671–680.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller (1953). Equation of state calculation by fast computing machines, *J. Chem. Phys.* **21,** 1087–1092.

Mori, J. and S. Hartzell (1990). Source inversion of the 1988 Upland, California, earthquake: determination of a fault plane for a small event, *Bull. Seism. Soc. Am.* **80,** 507–518.

Mueller, C. (1985). Source pulse enhancement by deconvolution of an empirical Green's function, *Geophys. Res. Lett.* **12,** 33–36.

Peterson, J. and A. Davey (1991). Crossvalidation method for crosswell seismic tomography, *Geophysics* **56,** 385–389.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions, *J. Roy. Stat. Soc.* **2,** 111–147.

Velasco, A., C. Ammon, and T. Lay (1994). Empirical Green function deconvolution of broadband surface waves: rupture directivity of the 1992 Landers, California (Mw = 7.3), earthquake, *Bull. Seism. Soc. Am.* **84,** 735–750.

Zollo, A., P. Capuano, and S. Singh (1995). Use of small earthquake record to determine source function of a larger earthquake: an alternative method and an application, *Bull. Seism. Soc. Am.* **85,** 1249–1256.

Institut de Géodynamique
Université de Nice-Sophia Antipolis
250 Rue A. Einstein
06560 Valbonne, France
(F.C., J.V.)

Géosciences
Rennes I
Avenue Général Leclerc
35042 Rennes Cédex, France
(D.G.)

Manuscript received 13 September 1995.