

Extended Seismic Source Characterisation using Linear Programming Inversion in a Dual Formulation



B. Caldeira, J.F. Borges & M. Bezzeghoud

Geophysical Centre of Évora/ Physics Department, ECT, University of Évora, Portugal

V. Bushenkov

CIMA, Department of Mathematics, University of Evora, Portugal

G. Smirnov

Centre of Physics, Dep. Mathematics and Applications, University of Minho, Portugal

SUMMARY

A linear programming (LP) inversion method in a dual formulation was applied to reconstruct the kinematics of finite seismic ruptures. In a general setting, this approach can yield results from several data sets: strong ground motion, teleseismic waveforms or/and geodesic data (static deformation). The dual formulation involves the transformation of a normal solution space into an equivalent but reduced space: the dual space. The practical result of this transformation is a simpler inversion problem that is therefore faster to resolve, more stable and more robust. The developed algorithm includes a forward problem that calculates Green's functions using a finite differences method with a 3D structure model. To evaluate the performance of this algorithm, we applied it to the reconstitution of a realistic slip distribution model from a data set synthesised using this model, i.e., the solution of the forward problem. Several other standard inversion approaches were applied to the same synthetic data for comparison.

Keywords: Seismic source reconstitution, numerical techniques, joint seismic and geodetic data inversion.

1. INTRODUCTION

The reconstitution of seismic rupture processes from records of their effects at the Earth's surface requires the use of discrete inverse theory methodological paths. In this context, the modelling of a seismic rupture involves determining the values of the finite-dimensional parameters of a model that numerically reproduces the effects of the earthquake on the Earth's surface. The most popular data sets used to describe these effects are elastic motion recorded at a near-field distances from the source (strong ground motions), elastic motion recorded at a far-field distances (teleseismic waveforms), and inelastic deformations recorded by geodetic techniques such as Global Positioning System (GPS) and Interferometric Synthetic Aperture Radar (InSAR). The detail and accuracy of the computed characteristic parameters for large earthquakes depend on a combination of two factors: the methods used and the input data. The kinematic model of a finite seismic source consists of a spatiotemporal distribution of slip vectors, i.e., a vector field $u(t,r)$ on a fault plane divided into a grid of **small elements**, or subfaults. Each slip represents the motion of a corresponding subfault. All subfaults are considered to have the same size, geometry and orientation (rectangular and oriented along *strike* and *dip* directions). The slip vector associated with each subfault is characterised by the following parameters: the start time, direction (rake angle), magnitude and the evolution with time (source time function) or the rise time of a defined scalar time function.

The most popular approach for determining the slip distribution models are the inversion of the seismic near-source strong ground motion waveform (e.g., Asano and Iwata 2009; Suzuki et al., 2009; Hartzel et al., 2007) and the joint inversion of near-source and teleseismic waveforms (e.g., Delouis et al., 2009; Yagi, 2004; Mozziconacci et al., 2009). Near-source data have the advantage of allowing the rupture kinematics to be reconstructed in more detail than when teleseismic waveforms are used exclusively. However, the use of this kind of data can pose some problems : accelerometer coverage unavailable or poor for some relevant seismic zones; inexistence of accurate of Earth structures models and the exigency of a very high computational power for waveform modelling.

According to the Discrete Inverse Theory, the modelling of any physical system by inversion involves three phases: the formulation of the forward problem, its parameterisation and the determination of the inverse problem. The forward problem corresponds to the application of the laws of physics to compute the data (elastic displacement, velocity, acceleration or deformation) using a given model. In the context of this study, the main issue in the forward problem is the calculation of the Green's functions, which are the approximated solution of the second-order elastodynamic equation at any point in an elastic medium when a perturbation (produced by a source with known mechanism) is applied at another point. Of the several approaches for solving this problem, the simplest is that set forth by Bouchon (1980), using a layered earth structure (a 1D velocity model) to obtain reasonable Green's functions for low frequencies (<1 Hz). When considering the propagation of seismic radiation through a complex 3D anisotropic earth structure, algorithms based on finite differences (e.g., Olsen and Archuleta, 1996; Larsen & Schultz, 1995), finite elements (e.g., Bao et al., 1998), or spectral elements (Komatitsch & Villote, 1998) should be used.

The first efforts to reconstruct the spatial and temporal rupture processes of finite seismic sources using the inversion of seismic waveforms were presented in theoretical works (Gilbert, 1975; Hartzell et al. 1978) and applied to the 1979 Imperial Valley earthquake (Hartzell and Heaton, 1983). These early works described the rupture model as a succession of slips on sections of a rectangular fault plane. The initiation time of each section (subfault) is controlled by a rupture front that spread over the fault plane with constant velocity in all directions from the hypocentre. The evolution of each slip is given by a scalar function with a fixed shape (source time function). This rupture scheme is known as the single time window model. Two aspects are difficult to resolve in the original single time window models. The first is related to the fact that the shape and duration of the source time functions, which are equal for all subfaults, limit the frequency range of the modelled data. The second is the incorrect assumption of constant rupture velocity imposed, which also affects the accuracy of the results of the data modelling. These difficulties were partially overcome by considering the slip in each subfault as a succession of elementary source time functions, which requires that the ruptures of each subfault occur in separate time intervals. This model, the multiple time windows model (Olson and Apsel, 1982; Cohee and Beroza, 1994), ensures a more realistic simulation of the ruptures at the expense of calculation time.

In the present framework, the finite-source models (e.g., Asano and Iwata, 2009; Mozziconacci et al., 2009; Robinson and Cheung, 2010; Delouis et al., 2009) are similar to those used in previous works. The major differences are the increase in the scale of the computation and the use of new optimisation techniques.

2. FORWARD PROBLEM

The description of the elastic displacement produced at the Earth's surface as the result of applied body forces, or slip discontinuities, in a semi-infinite elastic medium is the basis for the formal development of the methods for studying seismic sources. The discretisation of the integral form of the representation theorem (Aki and Richards, 1980) through proper parameterisation of the source allows for the computation of the synthetic seismograms. The fault plane is discretised into a set of N subfaults defined by a grid covering the entire surface. Each subfault is disposed along a square orthogonal referential, xOy . The time of rupture is also discretised into Nt intervals of time. Each subfault l ($l=1, N$) constitutes a point source that at certain time step k ($k=1, Nt$) initiates slips (breaks) in direction m according to a source time function $\dot{S}_{k,l,m}(t)$. The slip vector is defined by the magnitude of the two orthogonal components m : one in the *strike* direction ($m=1$) and the other in the *dip* direction ($m=2$).

The rupture described by this model is a sequence of slips characterised by a) position, b) initial time, c) amplitude, d) direction, and e) source time function. The adopted finite source model allows each subfault to be reactivated and break again in a different stage of the rupture after the first break. The complete parameterisation of this model requires the definition of the geometry of the fault plane, the

hypocentre position, the size of each subfault, and the time-step in which the rupture was discretised. According this model, the i^{th} component of displacement at station j , $u_i^j(t)$, is calculated by the following equation (where the asterisk, *, denotes a convolution):

$$u_i^j = \sum_{k=1}^{N_t} \sum_{l=1}^N \sum_{m=1}^2 (\dot{S}_{k,l,m}(t) * G_{i,j,k,l,m}(t)) x_{k,l,m} \quad (1)$$

where i, j, k, l, m represents, respectively, the direction of the displacement at observation point (1=North-South, 2=East-West, 3=Vertical), the observation point, the time step where the time of rupture was discretised, the subfaults and the components of the slip vector (1= strike direction; 2= dip direction). $G_{i,j,k,l,m}(t)$ is the Green's function that represents the temporal evolution of the component i of the displacement at the observation position j due to a unitary slip in the direction m produced at source l at time t , and $x_{k,l,m}$ represents the slip. The Green's functions were computed in terms of wave propagation in 3-D media using the E3D finite-difference code (Larsen & Schultz 1995).

The system of linear equations (1) that defines the computation of the synthetic seismograms (forward problem) can be translated to matrix language as

$$\mathbf{u} = \mathbf{A}\mathbf{x} \quad (2)$$

where \mathbf{u} is the vector that contain all seismograms, \mathbf{x} is the vector of the slips of all subfaults in whole time steps (slip distribution model), and \mathbf{A} is the matrix of Green's functions.

3. INVERSE PROBLEM

Following the procedure described by Das and Kostrov (1990), we denote $\mathbf{u} - \mathbf{A}\mathbf{x}$ by \mathbf{r} . We will minimize the absolute misfit

$$f = \sum_s |r_s|.$$

If we represent $\mathbf{r} = \mathbf{r}^+ - \mathbf{r}^-$, where $r_s^+ \geq 0$, $r_s^- \geq 0$, the function f will be linear

$$f = \sum_s (r_s^+ - r_s^-)$$

and the slip determination problem can be formalized as the linear programming (LP) problem

$$\begin{aligned} \sum_s (r_s^+ - r_s^-) &\rightarrow \min, \\ \mathbf{A}\mathbf{x} + \mathbf{r}^+ - \mathbf{r}^- &= \mathbf{u}, \\ \mathbf{c}^T \mathbf{x} &= M_0 \\ x_i \geq 0, r_s^+ \geq 0, r_s^- \geq 0 \end{aligned} \quad (3)$$

where, as in Das and Kostrov (1990), $\mathbf{c}^T \mathbf{x} = M_0$ is the requirement that total seismic moment equals a known value, and the vector \mathbf{c} represents the medium rigidity at the corresponding subfault multiplied by the subfault area times time step $\Delta\tau$. The weak causality-like constraint is also included, namely, we assume that the slip rate is zero at any subfault and time step which would produce a signal before the first arrival at any station from the hypocentral subfault.

For LP problem (3) can be formulated the dual LP problem, as follows:

$$\begin{aligned} \mathbf{u}^T \mathbf{z} + M_0 z_0 &\rightarrow \max, \\ \mathbf{A}^T \mathbf{z} + \mathbf{c}^T z_0 &\leq 0, \\ -1 \leq z_p \leq 1, p &= 1, \dots, n_u. \end{aligned} \quad (4)$$

where z_0 and the coordinates z_p of the vector \mathbf{z} , $p = 1, \dots, n_u$, are the corresponding dual variables, n_u

is the dimension of \mathbf{u} .

It is interesting to note that in the dual formulation (4), all seismograms data are localized in the functional. The problem restrictions are unaffected by seismogram vector \mathbf{u} and total seismic moment M_0 . This makes it possible to find a feasible solution for the dual LP problem (4) in advance. When new seismograms are given, the resolution of LP problem (4) starts from a good initial basis.

Another advantage of dual formulation (4) is that the variables z_p are limited to the interval from -1 to $+1$. They can not take large values which increases the computational stability of the algorithm.

When the restriction $\mathbf{c}^T \mathbf{x} = M_0$ is omitted in (3), the corresponding dual problem takes especially simple structure which is illustrated geometrically in Figure 1

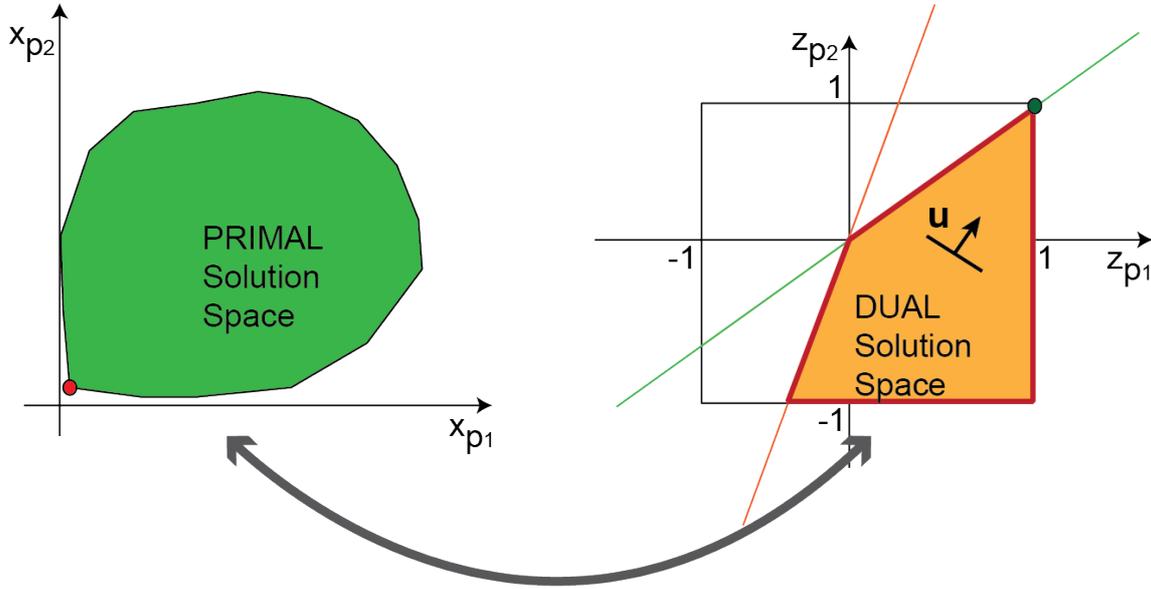


Figure 1. Geometric structure of the primal and dual linear programming formulations

4. CASE STUDY SCENARIO

To evaluate the proposed algorithm, we applied it to a synthetic seismic rupture scenario similar to real sources. This type of evaluation is extremely important because it is the only reliable way to analyse method performance when the expected results are known (Beresnev, 2003). The synthetic waveforms were calculated from (1) for a defined rupture model (Fig. 2) based on a set of 13 seismic stations distributed around the source in the geometry represented in Figure 3.

The fault plane was divided into a grid of 36, $2 \text{ km} \times 2 \text{ km}$, subfaults. The rupture starts at the initial time at the 12-km-deep hypocentral node and travels in all directions with a variable velocity. The slip of each node is specified by the initial time, two components of the slip vector and a triangular Source Time Function (STF) with a rise time τ . The rupture time is discretised using a temporal gridding of 0.3 s. The defined source model assumes that subfaults slip more than once at different stages of the rupture. The Green's functions were computed using a finite-difference spatial and temporal grid spacing scheme of 0.5 km and 0.03 s respectively. The velocity model is a $100 \text{ km} \times 100 \text{ km} \times 70 \text{ km}$ fragment of the 3D velocity model of SW Iberia (Grandin et al. 2007). Based on the velocity model and the spatial grid, the Green's functions are significant to a maximum frequency of 1.3 Hz. Thus, the computed Green's functions were filtered with a low-pass Butterworth filter with a cut-off frequency of 1.3 Hz to avoid the numerical noise.

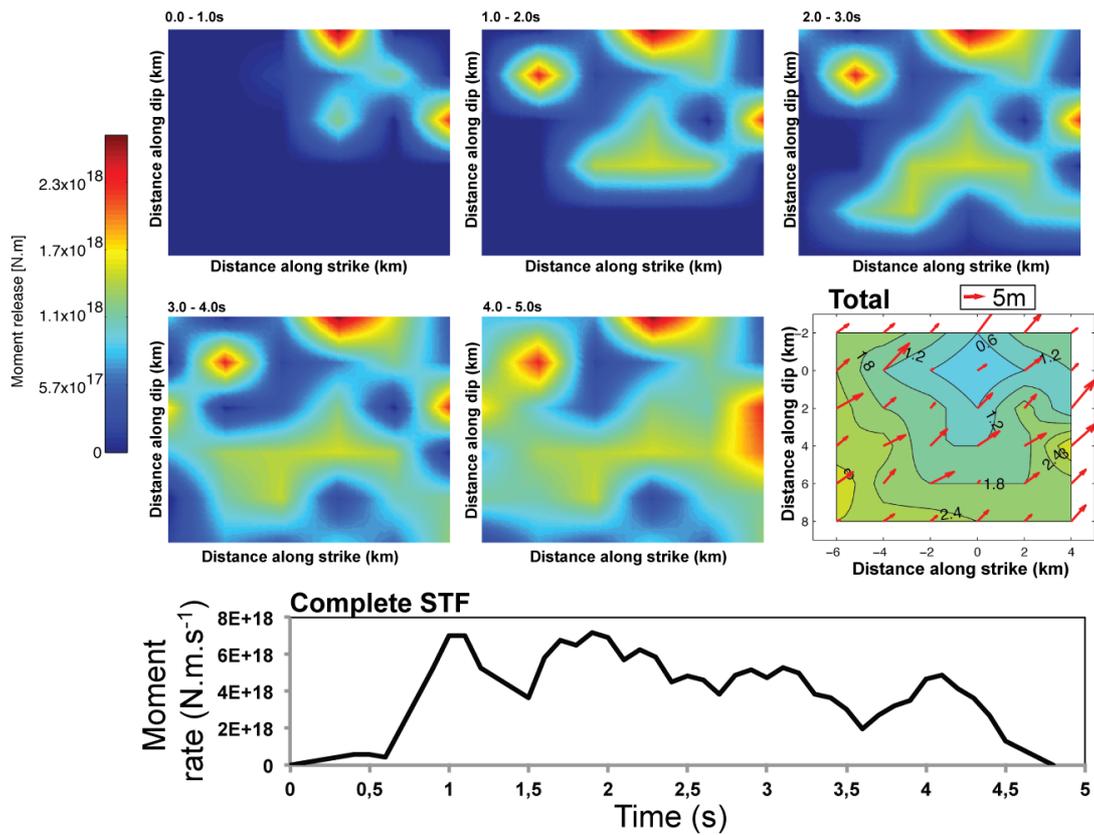


Figure 2. Rupture model defined. The first five individual panels (0-1s to 4-5s) show the distribution (in 1s time windows) of the seismic moment release. The sixth panel (denoted by Total) shows the final slip distribution (red arrows) and the coloured contours show rupture time in 0.6-sec contours. The bottom panel denoted “Complete STF” represents the rate of moment release.

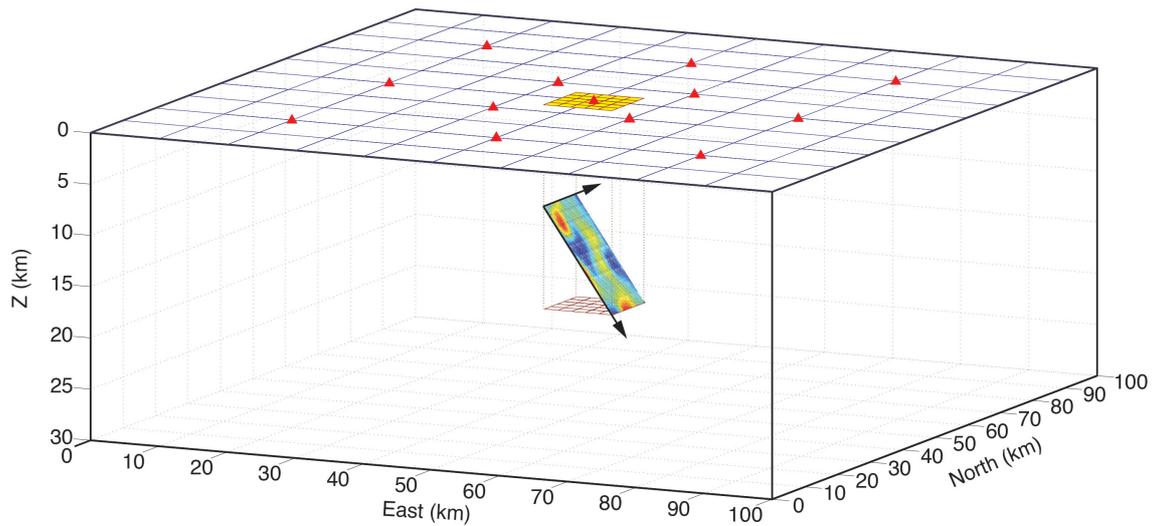


Figure 3. Geometry of the tested situation. The 13 red triangles on the surface represent the seismic observation points; at the interior of the volume is represented the fault plane.

5. RESULTS

Inversions were performed on the same data set using three different algorithms: the dual linear programming formulation presented herein and two standard algorithms, the least-squares method of Lawson and Hanson (1974) using the formulation of Hartzell and Heaton (1983) and the primal linear programming formulation developed by Das and Kostrov (1990). The same Green's functions and source parameterisations were used for all the procedures. Figure 4 presents the reconstructed rupture model calculated using the dual linear programming algorithm developed herein. The likenesses between this reconstructed model and the synthetic origin model (Figs. 4 and 2) are clear in both the spatial distribution of the slip and its spatial occupation over time, characterised by a non-uniform rupture front. The “total” slip distribution of the synthetic and reconstructed models indicates the spatial likeness; the evolution of the rupture, displayed in the sequence of snapshots and in the STF of the synthetic and reconstructed models, indicates a suitable temporal reproduction.

For the reconstitution by the primal linear programming formulation, convergence to a similar solution was obtained but required a processing time approximately 100 times longer than the reconstruction using the dual formulation (approximately 12 min with the dual and approximately 14 h with the primal formulation) and with a great number of warnings.

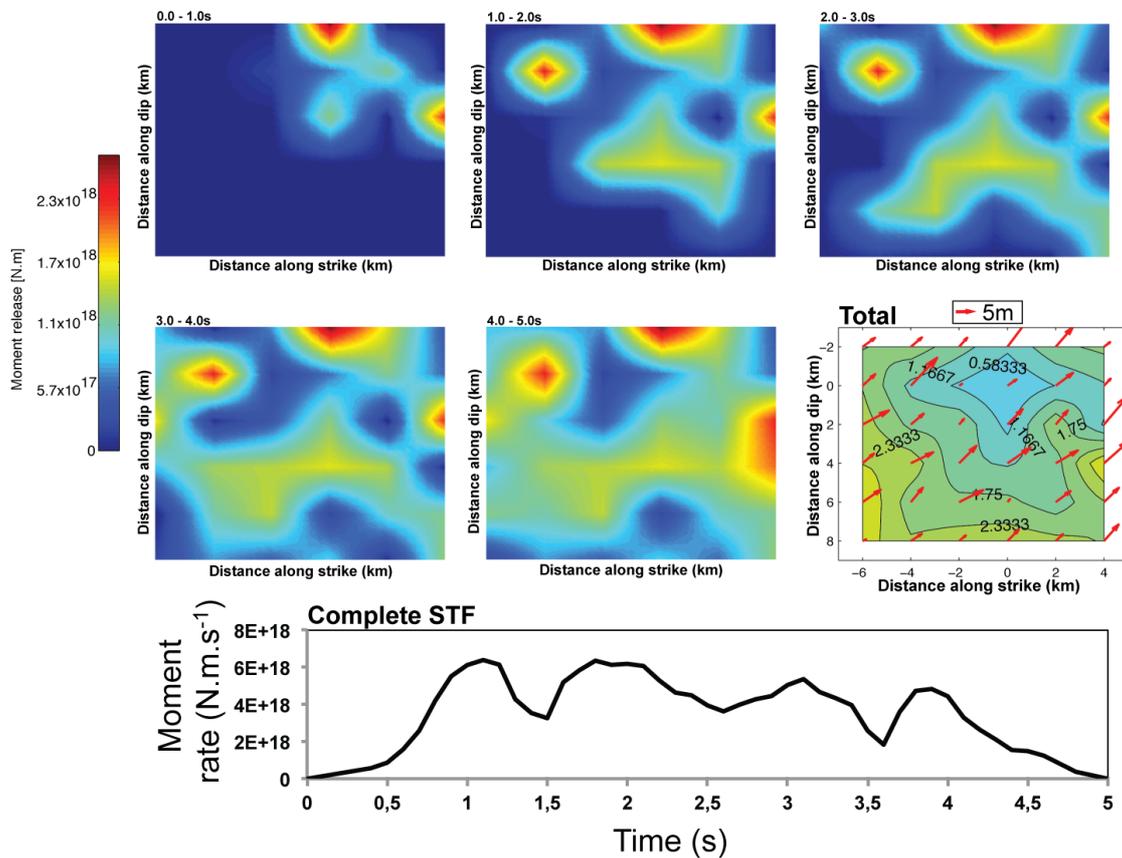


Figure 4. Rupture model reconstructed through the Dual version of Linear Programming algorithm developed. The first five individual panels (0-1s to 4-5s) show the distribution (in 1s time windows) of the seismic moment release. The sixth panel (denoted by Total) shows the final slip distribution (red arrows) and the coloured contours show rupture time in 0.6-sec contours. The bottom panel denoted “Complete STF” represents the rate of moment release.

As shown in Figure 5, the model reconstructed using the formulation of Hartzell and Heaton (1983), via the NNLS algorithm, differs significantly from the synthetic origin model (Fig. 2) in both total slip spatial distribution and slip temporal distribution.

Comparing the observed and model-predicted waveforms is the only way to validate the

reconstructions in many practical applications. In the trials performed in this work, a surprisingly good fit was obtained in all cases, even when the model differs from the real, reconstituted model. This result proves that comparing seismograms is not a reliable indication of the quality of a solution.

5. DISCUSSION

The reconstruction of the rupture kinematics for a large earthquake from a set of recorded effects on the Earth's surface remains a non-completely solved problem. Primal linear programming (LP) techniques (Das Kostrov, 1990; Hartzel and Liu, 1995) are an appropriate tool to resolve it to relatively small scale of solution domains. Primal LP techniques utilise an inversion method that explores the full solution space. The ease of incorporating constraints to improve solution convergence is an advantage of this type of approach over other global iterative methods. However, when the problem is parameterised to involve a large number of equations, its solution through the primal LP inversion techniques can become computationally expensive and difficult. In such cases, for example, reconstructing the rupture kinematics of large earthquakes, the dual LP inversion method reported herein is preferable because it reduces the dimension of the variable space and inputs the observed data (u) into the objective function, thus increasing the stability of the computation process. We demonstrate that the dual formulation has clear advantages in terms of both convergence and computing time when compared with the primal formulation used by previous authors.

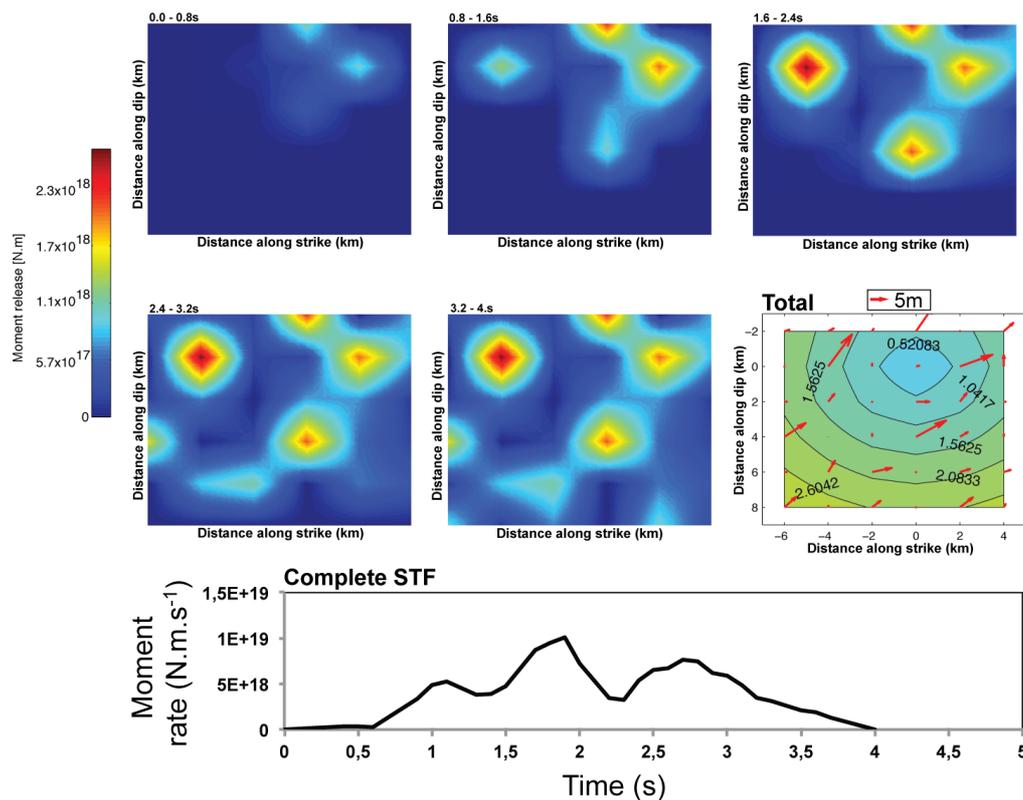


Figure 5. Rupture model reconstructed through the NNLS algorithm. The first five individual panels (0-0.8s to 3.2-4s) show the distribution (in 0.8s time windows) of the seismic moment release. The sixth panel (denoted by Total) shows the final slip distribution (red arrows) and the coloured contours show rupture time in 0.52-s contours. The bottom panel denoted “Complete STF” represents the rate of moment release.

The ability of the slip inversion methods to construct detailed rupture scenarios makes them particularly attractive for studying a seismic source. However, a detailed analysis of the

solutions provided by these methods reveals the requirements for the successful application of these tools. The first is the choice of method; a number of different ways of obtaining inversion scenarios are available in the literature, all of which have similar physical and numerical requirements (e.g., Hartzel and Heaton, 1983, Hernandez et al. 2001; Valle and Bouchon, 2004). However, applying these methods to the same data for the same events produces different results, as we can observe by comparison of the works of Wald and Heaton (1994) and Cohee and Beroza (1994) for the 1992 Landers earthquake.

Because the complex nature of this problem does not allow for an a priori knowledge of the best solution, the use of real data is not proper to compare methods or investigate parameterisation schemes by analysing solutions (Beresnev, 2003). This investigation can only be performed using synthetic data calculated from defined rupture models. Therefore, we tested the stability and robustness of the algorithm using synthetic waveforms computed from a defined slip distribution model, similar to real sources.

The results reveal the good likeness between the reconstructed model using both primal and dual LP inversion methods. Both versions converge to the same solution but with very different computing costs. Using the same simplex inversion routine, the dual method converges after approximately 12 min compared with 14 h for the primal method. The model reconstructed by the NNLS algorithm differs significantly from the synthetic origin model.

The method presented herein can be generalised to jointly utilise other data types (geodetic and teleseismic waveforms). The results obtained in this study encourage us to apply the proposed algorithm to real seismic and geodetic data, which is the next step.

ACKNOWLEDGEMENTS

This work was developed with the support of the Fundação para a Ciência e a Tecnologia (FCT) of the Ministério da Educação e Ciência through the projects SISMOT/LISMOT (PTDC/CTE-GIN/82704/2006) and ATESTA (PTDC/CTE-GIX/099540/2008).

REFERENCES

- Aki, K., and P. G. Richards, (1980). *Quantitative Seismology: Theory and Methods* (Vol. I and II), Freeman and Company, San Francisco.
- Asano, K., and T. Iwata. (2009). Source Rupture Process of the 2004 Chuetsu, Mid-Niigata Prefecture, Japan, Earthquake Inferred from Waveform Inversion with Dense Strong-Motion Data. *Bull. Seism. Soc. Am.* 99:1, 123-140, doi:10.1785/0120080257.
- Bao, H., J.Bielak, O.Ghattas, L.F.Kallivokas, D.R.O'Hallaron, J.R.Shewchuk, and J.Xu. (1998). Large-scale simulation of elastic wave propagation in heterogeneous media on parallel computers. *Comput. Methods Appl. Mech. Engrg.* 152, 85-102.
- Beresnev, I. A. (2003). Uncertainties in Finite-Fault Slip Inversions: To What Extent to Believe? (A Critical Review). *Bull. Seismol. Soc. Am.*, 93:6, 2445-2458, doi:10.1785/0120020225.
- Bouchon, M., (1981). A simple method to calculate Green's functions for elastic layered media, *Bull. Seismol. Soc. Am.*, 71, 959-971.
- Cohee, B. and Beroza, G. (1994). Slip Distribution of the 1992 Landers Earthquake and its Implications for Earthquake Source Mechanics. *Bull. Seim. Soc. Am.* 84, 692-712.
- Das, S., and B. V. Kostrov. (1990). Inversion for seismic slip rate history and distribution with stabilizing constraints: application to the 1986 Andreanof Islands earthquake. *J. Geophys. Res.* 95, 6899-6913.
- Delouis, B., M. Pardo, D. Legrand, and T. Monfret. (2009). The Mw 7.7 Tocopilla Earthquake of 14 November 2007 at the Southern Edge of the Northern Chile Seismic Gap: Rupture in the Deep Part of the Coupled Plate Interface, *Bull. Seism. Soc. Am.* 99(1), 87-94, doi:10.1785/0120080192.
- Gilbert, F. (1973). Derivation of source parameters from low-frequency spectra, *Phil. Trans. R. Soc. Lond. Ser. A.* 274, 369-371.
- Grandin, R., Borges, J.F., Bezzeghoud, M., Caldeira, B., and Carrilho, F. (2007). Simulations of strong ground motion in SW Iberia for the 1969 February 28 (MS = 8.0) and the 1755 November 1 (M ~ 8.5) earthquakes – I. Velocity model. *Geophys. J. Int.* 171, 1144-1161, doi:10.1111/j.1365-246X.2007.03570.x

- Hartzel, S. and P. Liu. (1995). Determination of earthquake source parameters using a hybrid global search algorithm. *Bull. Seism. Soc. Am.* 85, 516-524.
- Hartzell, S., Frazier, G. and Brune, J. (1978). Earthquake Modeling in a Homogeneous Half-Space. *Bull. Sismol. Soc. Am.* 68, 301—316.
- Hartzell, S., P. Liu, C. Mendoza, C. Ji, and K. M. Larson. (2007). Stability and Uncertainty of Finite-Fault Slip Inversions: Application to the 2004 Parkfield, California, Earthquake. *Bull. Seism. Soc. Am.* 97(6), 1911-1934, doi:10.1785/0120070080.
- Hartzell, S.H., and Heaton, T.H. (1983). Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake. *Bull. Seism. Soc. Am.* 73:6, 1153-1184.
- Hernandez, B., Shapiro, N. M., Singh, S., Pacheco, J., Cotton, F., Campillo, M., Iglesias, A., Cruz, V., Gómez, J. and Alcántara, L. (2001). Rupture History of September 30, 1999 Interplate Earthquake of Oaxaca, Mexico (Mw=7.5) from Inversion of Strong-Motion Data. *Geophys. Res. Lett.* 28, 363—366
- Komatitsch, D., and J. P. Vilotte. (1998). The spectral-element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bull. Seism. Soc. Am.* 88:2, 368–392.
- Larsen, S.C. , and Schultz, C.A. (1995). ELAS3D, 2D/3D Elastic Finite-Difference Wave Propagation Code, Lawrence Livermore National Laboratory, UCRLMA-121792, 18 pp.
- Lawson, C. and Hanson, R. (1974). Solving Least Squares Problems, Prentice-Hall (series in automatic computation), New Jersey.
- Mozziconacci, L., B. Delouis, J. Angelier, J. Hu, and B. Huang. (2009). Slip distribution on a thrust fault at a plate boundary: the 2003 Chengkung earthquake, Taiwan. *Geophysical Journal International.* 177(2), 609-623, doi:10.1111/j.1365-246X.2009.04097.x.
- Olsen, K. B., and R. J. Archuleta. (1996). Three-dimensional simulation of earthquakes on the Los Angeles Fault System, *Bull. Seism. Soc. Am.* 86, 575–596.
- Olson, A. and Apsel, R. (1982). Finite Faults and Inverse Theory with Applications to the 1979 Imperial Valley Earthquake. *Bull. Seism. Soc. Am.* 72, 1969-2001.
- Robinson, D. P., and L. T. Cheung. (2010). Source process of the Mw 8.3, 2003 Tokachi-Oki, Japan earthquake and its aftershocks. *Geophysical Journal International.* 181:1, 334-342, doi:10.1111/j.1365-246X.2010.04513.x.
- Suzuki, W., S. Aoi, and H. Sekiguchi. (2009). Rupture Process of the 2008 Northern Iwate Intraslab Earthquake Derived from Strong-Motion Records. *Bull. Seism. Soc. Am.* 99(5), 2825-2835, doi:10.1785/0120080331.
- Wald, D. and Heaton, T. (1994). Spatial and Temporal Distribution of Slip for the 1992 Landers, California, Earthquake. *Bull. Seism. Soc. Am.* 84, 668—691.
- Yagi, Yuji. (2004). Source rupture process of the 2003 Tokachi-oki earthquake determined by joint inversion of teleseismic body wave and strong ground motion data. *Earth Planet and Space.* 56, 311-316.