9th Biennial International Conference
& Exposition on Petroleum Geophysics

"HYDERABAD 2012"

P-328

An Implementation of Double Difference Approach to Improve the Imaging of the Earth's Interior

Akash Adwani[1], Deepanshu Melana[1] , Jean Virieux[2]*

Summary

We have developed a double-difference (DD) tomography code to improve the imaging of a heterogeneous medium containing seismogenic objects. It makes use of both absolute and relative arrival times and can be used to tackle active as well as passive tomography problems. This method is designed into independent major modulus which exchanges results through files. A projection of the inversion grid into the forward grid is the first step. Ray-tracing is performed using eikonal solvers and rays are stored in this second step. The third one is related to travel- time estimation and partial derivative evaluation. The fourth step is the preconditioning of this ill-conditioned system and the fifth and the final step performs the inversion and updates the model parameters. The user can perform each step by tuning the best set of external parameters without redoing the entire workflow. We have tested this method on a synthetic dataset and found that it produces improved velocity models and accurate event locations. We also test this method on 2D bore hole data collected by INERIS, France. The velocity model coming from double difference is more resolved and has sharper velocity contrasts than the delay travel time tomography

Keywords: *Seismic Tomography, Double Difference Technique.*

Introduction

First-arrival travel time tomography can be used to image the earth's interior at various scales, from near-surface to global scale using active and passive sources. For seismic imaging, determination of the near-surface velocity structure is a key step when trying to image deeper structures. After the work of Aki and Lee (1976),(who used P-wave delay time readings from Test Ban monitoring arrays like LASA and NORSAR to delineate the seismic velocity structure directly under these arrays) earthquake location and tomography have become widely used means to deduce active and passive structures of the Earth interior from the available seismological data. In recent years, Passive Tomography has gained popularity in Geophysical exploration due to its numerous advantages over conventional techniques. However, resolution of obtained 3D velocity models can be greatly affected by the location uncertainty of routinely determined hypocenters which is typically many times larger than the source dimension of the events itself, thus putting limits on the study of the fine structure of seismicity.

Pavlis (1986) showed that the accuracy of event hypocenters depends on the several factors, including the network geometry and arrival time accuracies. Due to the presence of noise, the arrival times picked either manually or automatically generally have errors. However, implementation of waveform cross-correlation (WCC) and event clustering techniques improves arrival time estimates or determine high-precision relative arrival times, thus providing improvements in location accuracy for earthquakes and explosions (VanDecar and Crosson, 1990). These studies are based on the assumption that two earthquakes produce similar waveforms at a common station if their source mechanisms are virtually identical, and waveform cross-correlation can then be used to determine precise relative arrival times. These and other studies have shown that seismogenic region knowledge may be deeply changed by the accurate relocation of earthquakes.

[1] *Indian School of Mines ,Dhanbad,India*
[2] *LGIT, Universite´ Joseph Fourier, Grenoble, France.*
D-205 Sapphire Hostel, Indian School of Mines, Dhanbad-826004, Jharkhand.
Email:akash.ismu@gmail.com

Waldhauser and Ellsworth (2000) demonstrated a method to determine the relative event locations by directly using relative arrival times. However it is only valid for closely spaced events and not for far apart events. As a result, such far apart event locations may be biased due to velocity heterogeneity.

We will be using a method that overcomes this disadvantage. It is based on the code tomoDD of Zhang and Thurber (2003), and makes use of both absolute and relative arrival time data. The method determines a three-dimensional (3D) velocity model jointly with the absolute event locations while using extended information. This approach has the advantage of including relative arrival times with their quality values along with absolute arrival times. Also, the present code is written in such a way that it takes into account all the cases, i.e., (1) Different Sources and same Receivers, (2) Different receivers and same sources and (3) Different sources and Different receivers (also known as Quad difference), provided that the cross correlation between them is possible.

This program has also been tested on Active tomographic example and it is observed that the velocity model obtained with DD tomography is superior to that from standard tomography. With standard tomography, event locations are somewhat scattered due to imprecise picks and correlated errors, but in DD tomography, the use of the differential arrival times removes most of these errors, which in turn removes some "fuzziness" from the velocity model.

Double Difference Tomography Method

The body wave arrival time T from an event i to a seismic station k is expressed using ray theory as a path integral

$$T^i = \tau^i + \int_i^k u\, ds \tag{2.1}$$

Where τ^i is the origin time of event ^{i} , u is the slowness field and ds is an element of path length. The source coordinates (x_1 , x_2 , x_3), origin times, ray paths, and the slowness field are the unknowns. The relationship between the arrival time and the event location is highly nonlinear, so a truncated Taylor series expansion is generally used to linearize equation (2.1). This linearly relates the misfit between the observed and predicted arrival times r_k^i to the desired perturbations to the hypocenter and velocity structure parameters.

$$r_k^i = \sum_{l=1}^{3} \frac{\partial T_k^i}{\partial x_l^i} \Delta x_l^i + \Delta\tau^i + \int_i^k \delta u\, ds \tag{2.2}$$

Subtracting a similar equation for event j observed at station ^{i} from equation (2.2), we have

$$dr_{kl}^{ij} = r_k^i - r_l^j = \sum_{m=1}^{3} \frac{\partial T_k^i}{\partial x_m^i} \Delta x_m^i + \Delta\tau^i + \int_i^k \delta u\, ds \tag{2.3}$$

Where dr_{kl}^{ij} is the so called double difference (Waldhauser and Ellsworth, 2000). This term is the difference between observed and calculated differential arrival times for the two events, and can also be written as

$$dr_{kl}^{ij} = r_k^i - r_l^j = \left(T_k^i - T_l^j\right)^{obs} - \left(T_k^i - T_l^j\right)^{cal} \tag{2.4}$$

The observed differential arrival times $\left(T_k^i - T_l^j\right)^{obs}$ can be calculated from both waveform cross-correlation techniques for similar waveforms and absolute catalog arrival times. Equation (2.4) is known as the DD algorithm (Zhang and Thurber, 2003).

The double-difference equation system may therefore be written

$$G\,\delta m = r \tag{2.5}$$

Where G is the matrix of the partial derivatives of the theoretical travel time differences with respect to the model parameters, δ m is the model perturbation vector, r is the travel time difference residual vector.

The system expressed by equation (2.5) being generally ill- conditioned, Waldhauser and Ellsworth (2000) and Zhang and Thurber (2003) have chosen to solve the system

$$W \begin{pmatrix} G \\ \lambda I \end{pmatrix} \delta m = W \begin{pmatrix} r \\ 0 \end{pmatrix} \tag{2.6}$$

Where W contains a priori quality weights. G is the matrix

containing the partial derivatives of the travel time differences (Waldhauser and Ellsworth, 2000) or the partial derivatives of the travel times and travel time differences (Zhang and Thurber, 2003) with respect to the model parameters; λ is a damping factor, I is the identity matrix, and r is the vector of residuals in travel time differences (Waldhauser and Ellsworth, 2000) or travel times and travel time differences (Zhang and Thurber, 2003). Waldhauser and Ellsworth (2000) add the constraint (for an N event cluster)

$$\sum_{i=1}^{N} \delta m_i = 0 \quad (2.7)$$

Earthquake location and tomography are, however, simultaneously nonlinear and ill-posed. Determining the hypocenter location and the seismic velocities is an optimization problem whose solution has been discussed by Tarantola and Valette (1982) and Tarantola (1987). The tomographic problem should actually be posed

$$g(m) = d \quad (2.8)$$

Where, g represents the functional used to solve the direct problem and d is the travel time data.

Montellier et. al [2003] showed that minimizing equation (2.7) is equivalent to solving iteratively the system

$$\begin{pmatrix} C_d^{-1/2} G \\ C_m^{-1/2} \end{pmatrix} \delta m_k = \begin{pmatrix} C_d^{-1/2} (d - g(m_k)) \\ C_m^{-1/2} (m_0 - m_k) \end{pmatrix} \quad (2.9)$$

Where C_d , C_m , m_0 are the data covariance matrix, the a priori model covariance matrix and the a priori model vector, respectively.

Our purpose is to determine the absolute locations and the velocity structure. Also, the model derivatives in equation (2.3) will cancel in the region where the ray paths substantially overlap. For this reason, we include the absolute arrival times in the inversion to resolve the velocity structure outside the source and the station region. By doing this, we can jointly determine the velocity structure and the relative event locations as well as the absolute event locations accurately.

Examples:

1. Synthetic Example

1.1) Formulation of the synthetic example:

We considered an example with 20 equally spaced earthquakes in X- and Y- direction each. All the sources are considered at the same level i.e., at 5000 meters . There are 36 stations above the events at 15000 meter level, with all the events observed at each station. As a result, there are 14,400 travel times in total for all the event and station pairs. Distribution of sources and receivers is shown in (figure 1). As it is shown, the sources and stations are distributed from 15000 meters to 35000 meters in x and y direction. Hence, the total volume is 20*20*10 km3. Synthetic P model is designed by adding a central negative velocity anomaly (figure 2) Theoretical travel times are computed in these models and are used as observed times.

1.2) Data

We have used 39,600 P- wave differential arrival times from WCC (Waveform cross-correlation), in addition to 14,400 absolute catalog P-wave arrival picks. WCC data was chosen with the help of Quality control on the distance between sources and distance between receivers.

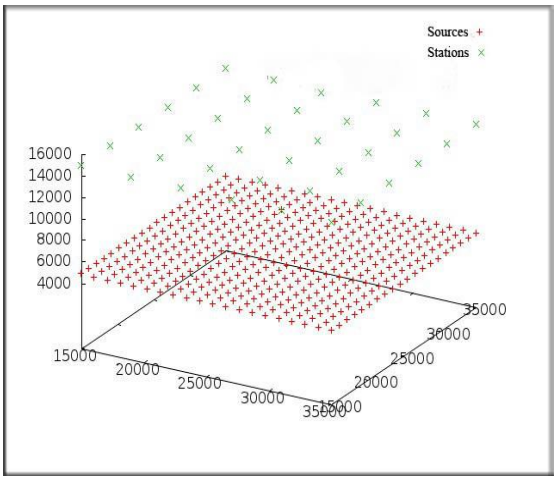

Figure 1: Distribution of quakes and the receivers

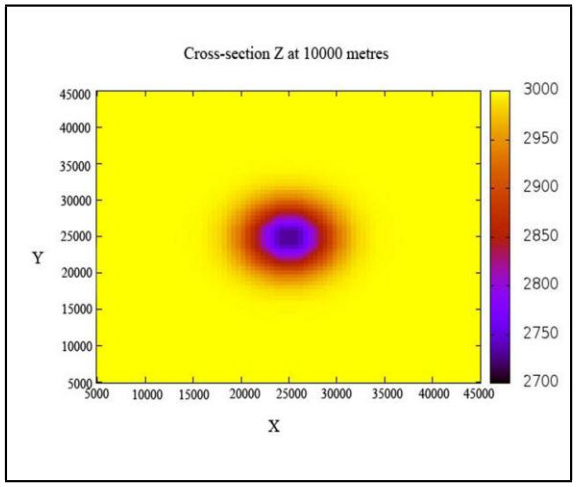

Figure 2: Synthetic model plot showing a boule in the center

1.3) Model Parameterization

We use a homogeneous velocity model with constant velocity of 3000 m/s as the starting model for DD tomography. Grid dimensions of forward cubical grid are chosen in such a way that it is finer enough to perform the ray tracing properly without providing any convergence problem of the rays due to bending but at the same time it is coarser enough to save the computational time, whereas the grid stepping of inversion grid is kept coarser to get the better resolution (Latorre et.al, 2004).

1.4) Inversion strategy

Relative weighing was done for the two datasets, i.e., Absolute arrival picks and differential data pick. For the first 7 iterations, the Absolute arrival picks were given more weight (weight 1) than differential data (weight 0.01) to reconstruct the large scale variations, while for the next 8 iterations the differential data are weighted more (weight 1) than the absolute (weight 0.01).

The damping term is required for fitting of delay times to obtain physically plausible velocity variations and earthquake distributions. This term removes high-frequency variations, which otherwise would not be constrained during the inversion procedure (latorre et. al 2004).For different damping values, in a range from 0.01 to 10, we have inverted simultaneously for both velocity and hypocenter parameters. Test results shows that the damping value of around 3 give the best compromise between well-retrieved velocity models and earthquake locations and a rather low data misfit.

1.5) Results

By inverting the selected data set, final 3-D P velocity structure and earthquake locations are obtained after 15 iterations. Figure 3 shows the relocation of the events after the Double difference tomography. The weighted rms residuals for cross- correlation data and catalog data decrease from 19.17 to 1.58 msec and from 78.33 to 1.26 msec, respectively. While, in case of standard tomography it decreases from 116 to 5.37 msec. The reconstructed velocity model have been shown in Figure 4.It can be seen that the negative anomaly has been recovered and is well resolved except for some artifacts near the anomaly.

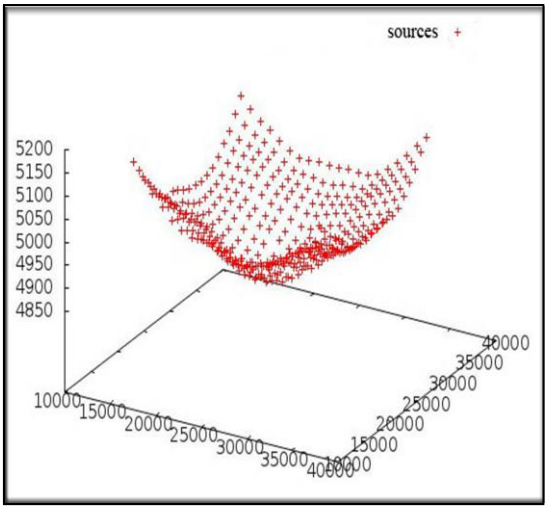

Figure 3. Relocation of Earthquakes through double difference due to the Boule in the middle of the distribution.

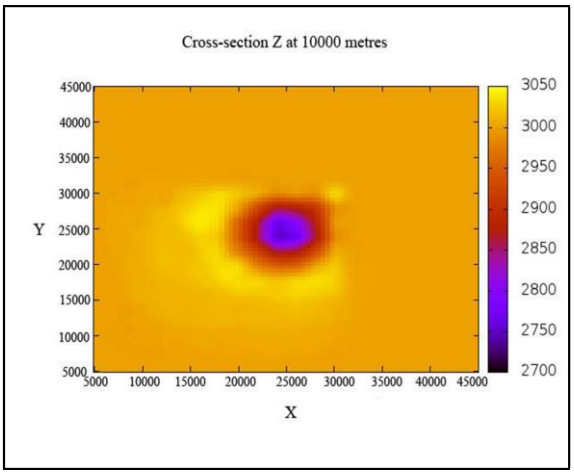

Figure 4. Reconstructed velocity model (Z- cross section at depth 10000 metres)

2. Application on 2D Bore hole Data collected by INERIS:

2.1) Formulation of the Problem

We applied the Double Difference tomography on the 2D borehole data collected by the INERIS to detect the void, which is due to the result of intensive mining activity in the north eastern parts of France. Two holes of depth 25 meters were dig, Shots and receivers were placed in the separate holes. The holes were inclined to an angle such that the maximum rays passes through the void. There are 48 shots and 47 receivers, with all the events observed at each station. As a result, there are 2256 travel times in total for all the event and station pairs.

2.2) Selection of the Data and Initial Model:

We have used 2,58,387 P- wave differential arrival times from WCC, in addition to 2,256 absolute P-wave arrival picks, which were picked using cross-correlation. WCC data was chosen with the help of Quality control on the distance between sources and distance between receivers. For accepting the cross-correlated data, maximum permissible distance between sources as well as between receivers was taken as 5 meters.

Kissling et al. (1994) demonstrated how the inappropriate choice of the initial model may give rise to artifacts in inversion results. Hence, special care was taken while selecting the best initial model. With the help of the recovered model provided by the INERIS , we were able to choose the best initial model after analyzing several other several models. The selected initial model is a homogeneous model with constant velocity 2100 m/s and it gives the least rms value out of other input models which were tested.

2.3) Model Parameterization

Grid dimensions should be chosen in such a way so that it should be able to delineate, as much as possible, shape and position of heterogeneities. We have tested different grid spacing to find the best compromise between model parameterization, spatial resolution and a reliable representation of the velocity structure. These tests have demonstrated that x-direction grid distance of 0.7 m and a z-direction grid distance of 0.75 m are the smallest spacing achievable without introducing any a priori smoothing for our present data set. Therefore, we have parameterized the 25 × 4 × 37 m3 inversion volume by an inversion grid step of 0.7 × 4 × 0.75 m3 .While the forward grid is considered with finer grid stepping so as to improve the results of ray tracing. Since it's a 2D dataset, we have copied the value of the velocity at y=0 to its all the four nodes, just to avoid any gradient because of that.

2.4) Inversion strategy

Like the previous example, Relative weighing was done for the two datasets, i.e., Absolute arrival picks and differential data pick. For the first 15 iterations, the Absolute arrival picks were given more weight (weight 1) than differential data (weight 0.01), while for the next 15 iterations the differential data are weighted more (weight 1) than the absolute data (weight 0.01).

To know the best choice for the Damping value, we have inverted simultaneously for both velocity and hypocenter parameters for damping values in a range from 0.01 to 100. Test results show that the final root mean square (rms) value is rather high for damping values greater than 10. On the contrary, for damping values smaller than 1, very low rms values are reached after only five iterations but the image of recovered P velocity anomaly is blurred by the artifacts to an extent that it is impossible to relate with the true dataset. Smoothing parameters have been used as well to deal with this problem. Ultimately, damping values around 20 give the well-retrieved velocity models.

2.5) Results

Reconstructed 3-D P velocity structure is obtained after 30 iterations. The weighted rms residuals for cross-correlation data and catalog data decrease from 0.46 to 0.36 msec and from 0.746 to 0.477 msec, respectively. While, in case of standard tomography it decreases from 0.746 to 0.40. The reconstructed velocity model has been shown in Figure. 5.

It can be seen that the reconstructed model is well resolved with a noticeable drop in resolution where the finer gridding is used. There is a gradient with minimum velocity near the free surface which is clay medium and then it increases downwards. It is further observed from the plot that there is a negative anomaly at the middle which is the void that has been recovered. For comparison, Figure 6 shows the reconstructed model obtained by travel time tomography with the same model parameterization and smoothing parameters. The anomaly is recovered in both the cases but it is sharper in case of double difference tomography. Additionally, some areas are not recovered properly by the delay travel time tomography, especially

near the source and the station region. For example, observe both the plots at a depth between -5 to -10 meter, it can be seen that there is a high velocity zone with velocity approx 2170 m/s, which is not recovered in case of delay travel time tomography. Clearly, in this case application of Double difference tomography provides more resolved image than the delay travel time tomography.

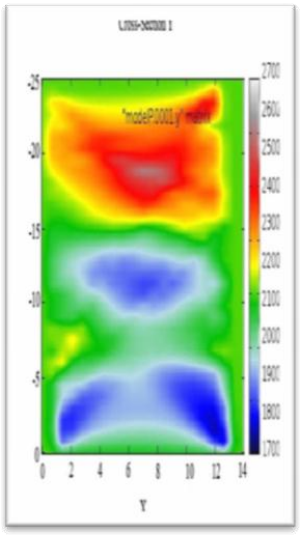

Figure 5: Reconstructed velocity model through Double Difference. A sharp velocity gradient has been noticed with a well recovered void at depth 10-14m.

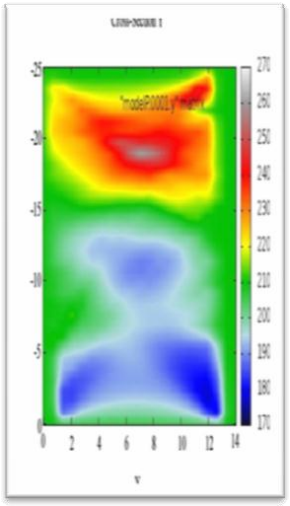

Figure 6:Reconstructed velocity model through delay travel time tomography. A void has been recovered at 10-14m depth but it is not well resolved.

Conclusion

The Double Difference tomography method introduced inside the TOMOTV software is efficient in tackling passive as well as active tomographic problems.It is able to relocate large numbers of earthquakes accurately (in case of passive tomography) as well as reconstructing the local velocity structure finely. However, its accuracy is restricted by the inaccurate cross-correlation picking and the large distances between sources and the receivers. Hence special care has to be taken while using the accurate WCC-derived relative arrival times. Inclusion of absolute arrival times in addition to the relative arrival times gives the absolute event locations. As it was shown through the example Double Difference tomography method is able to sharpen the image of the velocity structure, especially near the source and the station region.

References

Aki, K., and W. H. K. Lee (1976), Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model, J. Geophys. Res., 81, 4381–4399.

Kissling, E., W. L. Ellsworth, D. Eberhart-Phillips, and U. Kradolfer (1994). Initial reference models in local earthquake tomography, J. Geophys. Res. 99, 19,635–19,646.

Latorre, D., J. Virieux, T. Monfret, V. Monteiller, T. Vanorio, J.-L. Got, and H. Lyon-Caen (2004), A new seismic tomography of Aigion area (Gulf of Corinth, Greece) from the 1991 data set, Geophys. J. Int., 159, doi:10.1111/j.1365-246X.2004.02412.x.

Lee, W. H. K., and S. W. Stewart (1981). Principles and applications of microearthquake networks, Adv. Geophys. Suppl. 2, 293 pp.

Monteiller, V., J.-L. Got, J. Virieux, and P. G. Okubo (2005), Double difference tomography of Kilauea volcano, Hawaii, J. Geophys. Res, VOL. 110, B12306, doi:10.1029/ 2004JB003466

Paige, C. C., and M. A. Saunders (1982). LSQR: sparse linear equations and least squares problems, ACM Trans. Math. Softw. 8, no. 2, 195–209.

Pavlis, G. L., and J. R. Booker (1980). The mixed discrete continuous inverse problem: application to the simultaneous determination of earthquake hypocenters and velocity structure, J. Geophys. Res. 85, 4801–4810.

Tarantola, A. (1987), Inverse Problem Theory, 613 pp., Elsevier, New York.

Tarantola, A., and B. Valette (1982), Generalized nonlinear inverse problems solved using the least-squares criterion, Rev. Geophys., 20, 219–232.

Thurber, C. H. (1983). Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, J. Geophys. Res. 88, 8226–8236.

VanDecar, J. C., and R. S. Crosson (1990). Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, Bull. Seism. Soc. Am. 80, 1548–1560.

Waldhauser, F. (2001). hypoDD: a computer program to compute doubledifference hypocenter locations, U.S. Geol. Surv. Open File Rept. 01- 113, 25 pp.

Wolfe, C. J. (2002). On the mathematics of using difference operators to relocate earthquakes, Bull. Seism. Soc. Am. 92, 2879–2892.

Zhang, Haijiang and Thurber, C. H. (2003). Double-Difference Tomography: The Method and Its Application to the Hayward Fault, California, Bull. Seism. Soc. Am. Vol. 93, No. 5, pp. 1875–1889

Acknowledgement

We thank to Mark Wathelet (ISTerre), without his guidance this project would not have materialized. We express our gratitude to Romain Brossier (ISTerre) for providing us data and the numerous discussions we have had. We would also like to extend our gratitude to ISTerre for providing us the computer resources. This work was supported by EGIDE grants.