Reverse modelling for seismic event characterization – A new tool for passive seismology

D. Gajewski and E. Tessmer

email: gajewski@dkrz.de **keywords:** seismic event localization, reverse modelling, passive seismology

ABSTRACT

The localization of seismic events is of utmost importance in seismology. Current techniques rely on the fact that the recorded event is detectable at most of the stations of a seismic array. Weak events, not visible in the individual seismogram of the array, are missed out. We present a new approach, where no picking of events in the seismograms of the recording array is required. The observed wavefield of the array is reversed in time and then considered as the boundary value for the reverse modelling. Assuming the correct velocity model, the reversely modelled wavefield focusses on the hypocenter of the seismic event. The origin time of the event is given by the time where maximum focussing is observed. The spatial extent of the focus resembles the resolution power of the recorded wavefield and the acquisition. This automatically provides the uncertainty in the localization with respect to the bandwidth of the recorded data. The new method is particularly useful for the upcoming large passive arrays since no picking is required. It has great potential for localizing very weak events, not detectable in the individual seismogram, since the reverse modelling stacks the energy of all recorded traces and therefore enhances the signal-to-noise ratio. The method is demonstrated by 2-D and 3-D numerical case studies which demonstrate the potential of the new technique. Events with a S/N ratio smaller than 1 where the events cannot be identified in the individual seismogram of the array are localized very well by the new technique.

INTRODUCTION

The problem of earthquake location is one of the most basic problems in seismology. Although numerous applications exist worldwide, the inherent non-linearity prevents earthquake location and tomography from being a standardized routine tool. The earthquake location problem is stated as follows (Pujol, 2004): Given a set of arrival times and a velocity model, determine the origin time and the coordinates of the hypocenter of the event. This definition inherently assumes that the arrivals of an event are visible on a certain number of the recorded seismograms of the observing array. This also means, that the arrival has to be identified in the seismogram prior to the actual localization of the event. This not only requires the correct identification of the onset of the arrivals but also the proper correlation of the individual phases among the different stations of the array.

Source location methods can be categorized into absolute location methods and relative location methods. For the first type, the determination of the excitation time and hypocenter of a seismic source is traditionally performed by minimizing the difference between the observed and predicted arrival times of some seismic phases. For recent advances in seismic event location using such approaches see, e.g., Thurber and Rabinowitz (Thurber and Rabinowitz, 2000). The second class of methods considers relative location within a cluster of events using traveltime differences between pairs of events or stations, see, e.g., Waldhauser and Ellsworth (Waldhauser and Ellsworth, 2000). These methods allow improvement of the relative location between seismic sources but are exposed to the same difficulties mentioned above.

Annual WIT report 2004

In the above techniques, it is assumed that the event is visible on at least a few stations of the recording array. In this paper we follow a different approach. Similar to Kao and Shan (Kao and Shan, 2004) we will not assume that the event is visible on the individual seismogram of the array and we therefore do not pick arrival times. The information present in all recordings of the whole array is exploited. Since the image of the seismic source is observed from different observation angles by the different receivers of the array, we can exploit this redundancy in the data. In applied seismics, the redundancy or the fold of the recorded reflection data is the key issue for stacking with its potential to increase S/N ratio. For passive seismic arrays the fold for an observed seismic event is determined by the numbers of the receivers in the array. Kao and Shan (Kao and Shan, 2004) developed the Source-Scanning-Algorithm where they exploit the stacking advantage by computing brightness functions for trial locations and origin times.

In our approach a reverse modelling technique is used to propagate the emitted seismic energy back to its origin. The wave propagation for this situation is basically a one way process. We consider the recorded seismograms, which are reversed in time, as the boundary values of our modelling. These boundary values are downward continued, i.e., propagated backwards through the assumed model. If the model is correct the backward propagated energy will focus at the source location and the corresponding zero time. The spatial extent of the focus depends on the bandwidth of the recorded data and reflects the localization potential of the event recorded by the used acquisition geometry, i.e., the localization uncertainty is automatically obtained.

After the introduction we present the general outline of the methodology and describe its numerical realization. Two-dimensional and three-dimensional numerical case studies for simple and complex models demonstrate the feasibility of the method and show its potential for future development. Issues related to model complexity, poor S/N ratio, velocity errors and sparse receiver arrays are considered in the case studies. In the discussion section we suggest applications of the method and formulate extensions not yet put into practice.

METHOD

Outline of the Methodology

The backward propagation of the recorded wavefield is so far entirely based on acoustic numerical seismic modelling. In the present work the Fourier method, a pseudo-spectral method Kosloff and Baysal (1982), is applied. However, finite-differences algorithms or elastic modelling could also be used as well. The main idea is that, in contrast to seismic forward modelling initiated by a (highly) localized source, seismograms reversed in time are used as initial conditions at the receiver locations. For the sake of simplicity the algorithm is presented here for a 2-D space.

The modelling scheme is used to propagate the wavefield, which is fed into the numerical model at the receiver stations, backwards in time. This can be done since the wave equation has no inherent preferred direction in time. It is expected that after propagating the wavefield backwards in time, all energy will be focussed at the source location, which will lead to large amplitudes at this position. These large amplitudes can be easily detected by scanning the image at every time step. A continuation of the process of backward propagation beyond maximum focussing will spread the wavefield away from the source.

The propagation in time is performed iteratively by a time-stepping scheme step by step with small time increments Δt . At each time step the entire grid representing the subsurface pressure field is scanned for its maximum value and the spatial position thereof is stored. The location of the maximum pressure amplitude over all time is the source position and its time is the excitation time of the source we are searching for. The excitation time is obtained from the absolute time of the first sample fed into the back propagation and the number of time steps needed to focus the event.

Assuming the recorded wavefield stems from a point source and the subsurface velocity model is known, ideally the wavefield should focus perfectly. The spatial extent of the focus reflects the resolution power of the experiment. This depends on the bandwidth of the recorded seismograms and the acquisition geometry. Thus, the localization uncertainty is automatically obtained with this method. However, the receiver locations are not perfectly sampled around the source position. Therefore, much information for reconstructing the total wavefield at the source is missing. Consequently, we cannot expect that the wavefield perfectly collapses at the source location. Nevertheless, the examples presented below show that this

method works surprisingly well even for erroneous velocities and low fold arrays.

Modelling

Various types of wave equations (WE) can be used for the propagation algorithm. There are the variabledensity WE and the constant-density WE, which in pseudo-spectral algorithms is computationally half as expensive as the former. These two WEs naturally produce reflections at interfaces. If reflections need to be avoided, which is desirable in our case, a one-way WE Baysal et al. (1983) which shows no reflections at interfaces at all, or a so called impedance matching WE Baysal et al. (1984) with vanishing reflection coefficient at normal incidence can be used.

The various WEs for 2-D media are given below:

Two-way wave equation

$$\frac{\partial^2 p}{\partial t^2} = \rho(x,z) \ c^2(x,z) \left[\frac{\partial}{\partial x} \left(\frac{1}{\rho(x,z)} \ \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial z} \left(\frac{1}{\rho(x,z)} \ \frac{\partial p}{\partial z} \right) \right] \tag{1}$$

Constant-density two-way wave equation

$$\frac{\partial^2 p}{\partial t^2} = c^2(x, z) \left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} \right)$$
(2)

Impedance-matching two-way wave equation

$$\frac{\partial^2 p}{\partial t^2} = c(x,z) \left[\frac{\partial}{\partial x} \left(c(x,z) \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial z} \left(c(x,z) \frac{\partial p}{\partial z} \right) \right]$$
(3)

One-way wave equation

$$\frac{\partial p}{\partial t} = c(x,z) \ FFT_{xz}^{-1} \left[sign(k_z) \ i \ (k_x^2 + k_z^2)^{1/2} P \right]$$

$$\tag{4}$$

where p is the pressure field, P is the pressure field double Fourier-transformed with respect to directions x and z, c is the velocity, ρ is the density, k_x and k_z are spatial wave numbers and FFT_{xz}^{-1} denotes a double inverse Fourier transform with respect to x and z. p, c and ρ are functions of the position.

The one-way WE cannot be directly implemented into finite-difference methods since it is implemented in the wavenumber domain.

Most of the computational work is needed for calculating the spatial derivatives. The spatial derivatives were computed by the Fourier method, which avoids numerical anisotropy and allows the use of coarse grids, i.e., only two grid points per shortest wave length.

In the WEs the time derivatives were replaced by the following finite-difference approximations:

$$\frac{\partial^2 p}{\partial t^2} \approx \frac{p(t+\Delta t) - 2p(t) + p(t-\Delta t)}{(\Delta t)^2}$$

and

$$\frac{\partial p}{\partial t}\approx \frac{p(t+\Delta t)-p(t-\Delta t)}{2\Delta t}$$

respectively. For extrapolation in time - be it forward or reverse in time - the resulting expressions are then solved for $p(t + \Delta t)$ or for $p(t - \Delta t)$ when modelled backwards in time.

NUMERICAL CASE STUDIES

The performance of the localization methodology described above is demonstrated using simple and complex subsurface models for which synthetic seismograms were computed by the Fourier method utilizing a code developed by Tessmer (Tessmer, 1990). We consider first a simple 2-D model consisting of four homogeneous layers separated by interfaces exhibiting lateral variations and a very complex 2-D subsurface structure, a part of the Marmousi model Versteeg and Grau (1991). The synthetic seismograms for the tests were computed by the Fourier method using the acoustic variable-density wave equation. These seismograms served as input data. In order to study the quality of estimating the source location and excitation time we apply noise to the input seismograms and use incorrect subsurface velocities for the reverse modelling to demonstrate the ability of the method to cope with such situations. Moreover, we perform tests with reduced numbers of receivers. The 3-dimensional version of the algorithm demonstrates that the method works well also for a 3-D subsurface structure.

Simple Model 2-D

The simple velocity model is shown in Fig. 1. The grid spacing is 10 m. The velocities vary between 2000 m/s and 3000 m/s. The explosive source is located at (360, 400) m. We use a Ricker-wavelet with a dominant frequency of 50 Hz and a high-cut frequency of 100 Hz. 2000 time steps with a time increment of 0.5 ms were calculated. 143 receivers are placed at the surface with a spacing of 10 m. The reverse propagation is performed using the one-way wave equation. The calculations are done on a numerical grid of 143 x 117 nodes. Surface seismograms for the above mentioned source position are shown in Fig. 2. Only every second trace is displayed. Since this is a synthetic example the exact location and excitation time are known. All times mentioned in the following text are relative to the exact source time, which is 0 s (zero time); i.e., negative numbers correspond to times prior to the real excitation time and positive numbers indicate times after the real excitation time.

Correct Velocities In the first test of reverse propagation the correct velocity model is used. The scan for maximum amplitudes yields the exact grid point of the source location. Snapshots showing the wavefront collapsing at the source location are shown in Fig. 3. The final snapshot shows an almost point-like wavefield. The areal width of the focus corresponds roughly to the the prevailing wavelength of the recorded data (about 60 m). The localization uncertainty is automatically obtained. For this ideal surface acquisition, an optimal result without spatial or timing errors is achieved.

Velocity Errors In this test of reverse propagation, the velocities used throughout the model were 10% higher than the correct velocities. Snapshots showing the wavefront collapsing at the source location are shown in Fig. 4. The scan for maximum amplitudes yields the exact horizontal position of the source. However, the vertical position is found 50 m above the real source position (i.e., at 350 m instead of 400 m) corresponding to about one dominant wavelength. Also the excitation time is not found correctly (maximum amplitude is found at 30 ms instead of 0 ms) due to wrong velocities. These deviations are expected, since an incorrect velocity model was used for the reverse modelling. The image with the highest amplitude is shown in Fig. 4 (bottom).

Another example of an incorrect velocity model would be a smoothed version of the velocities shown in Fig. 1. Such models are typically obtained from a tomography study. Performing the reverse modelling with a 100 times applied 3-point smoothing operator almost repeats the results for the unsmoothed velocities and we therefore do not reproduce the result here. The maximum amplitude of the reversely modelled sections is found at -1.5 ms and the determined location corresponds to the real source position.

Noisy Data In order to examine the effect of noise on the localization procedure a signal-to-noise ratio of 0.5 is chosen. The seismograms contaminated with noise are displayed in Fig. 5. The seismic event cannot be identified in the individual seismogram of the section. After reverse propagation of the wavefield, the estimated source position is found at (360, 390) m, i.e., the vertical position is wrong by a third of the dominant wavelength. The timing error is 3 ms. In view of the fact that the events cannot be visually identified in the seismograms, this is a remarkably good result which is explained by the stacking of energy

from the large number of receivers. The corresponding snapshots showing the wavefront collapsing at the source location are shown in Fig. 6. The focal area is clearly visible despite a higher noise level in the image.

Reduced number of receivers For realistic acquisitions with passive arrays the number of receivers is considerably smaller than in the previous examples. This also means that the observations are spatially aliased. In this example the number of receivers is only a tenth of that in the previous cases, i.e., 14 equally spaced receivers are used. The receiver spacing is 100 m. The estimated source position and its excitation time is still found correctly. However, the wavefront does not look as smooth as in the previous cases since the wavefield is heavily undersampled. It can be seen merely as the envelope of several wavefronts, however, the focal area is clearly imaged. The snapshots showing the wavefront collapsing towards the source location are shown in Fig. 7.

Complex Model

As a more complex model, a portion of the Marmousi model Versteeg and Grau (1991), is considered to test the algorithm in a geologically more difficult situation. In this subsurface model the velocities vary between 1500 m/s and 5500 m/s. The grid spacing is 8 m. The Ricker-wavelet has a maximum frequency of 90 Hz where the dominant frequency is 45 Hz. The position of the explosive source is at (2800, 1520) m. 5000 time steps with a time increment of 0.25 ms were calculated. The velocity model is shown in Fig. 8. The calculations are done on a numerical grid of 693 x 385 nodes. Seismograms for the above mentioned source position are shown in Fig. 9. Only every fifth trace is displayed.

Correct Velocities Despite the fact that the Marmousi velocity model is rather complex, the reverse modelling of the seismograms for the localization yields the exact source position and the correct excitation time. Snapshots of the wavefront are shown in Fig. 10. The excellent localization can be explained by the good coverage with receivers. The receiver spacing in this example is only 8 m. In another example below we consider a more realistic acquisition with a reduced number of geophones.

Velocity Errors For this test the velocities are chosen 10% higher than the correct velocities in the entire subsurface model. The scan for maximum amplitudes yields source coordinates of (2800, 1496) m. This is correct in horizontal direction. The vertical position, however, is about 2% incorrect in depth (1496 m instead of 1520 m). Also, the determined excitation time is incorrect by 76 ms. This is caused by the use of wrong velocities. Snapshots showing the wavefront collapsing towards the source location are shown in Fig. 11.

Reduced number of receivers Compared to the previous example here the number of receivers is reduced by a factor of 60, i.e., the receiver distance is now 480 m. The reverse modelling of the seismograms for the localization yields an almost accurate source position. Only the vertical coordinate is missed by 8 m. The excitation time error is 2 ms. Snapshots are shown in Fig. 12.

It is important to note, that for correct reverse modelling the geometrical spreading losses should be removed, which is not correctly handled by the current reverse modelling approach. We use time-dependent weights for a crude correction of spreading losses. If such corrections are not applied the largest amplitudes are observed directly at the receivers for sparse receiver spacing. A more detailed consideration of this issue is given in the discussion.

Simple 3-dimensional Model

To test the three-dimensional implementation of the localization algorithm a 2.5-dimensional version of the velocities for the simple model (Fig. 1) is considered. The original 2-D velocity model is extended laterally into the y-direction leading to a 2.5-D velocity distribution. The model is made up of $140 \times 140 \times 110$ grid nodes with a spacing of 10 m. The explosive source is located at (360, 760, 490) m. A reduced number of receivers, in this case at every 20th grid node, is used. In total the seismograms of 25 receivers with an

average spacing of 200 m are fed into the numerical model as boundary conditions. In Fig. 13 snapshots at different times for the xz- and xy-plane containing the source location are shown. The algorithm delivered the correct horizontal locations and excitation time. The vertical position was incorrect by only one grid cell.

DISCUSSION AND CONCLUSIONS

We have presented a new technique for the localization of seismic events and the estimation of their excitation times. The examples show that the method works extremely well if a sufficiently high coverage with receivers is present. The algorithm still yields very good estimates even if the data are very noisy, here S/N = 0.5. This makes the method attractive especially for weak onsets which are not detectable on single traces. As one would expect, the quality of the localization and excitation time estimate degrades if the subsurface velocities are not exactly known. A smoothing of the velocities does not degrade the localization. The numerical experiments have shown that the algorithm works very well even for source radiation not visible in the individual seismogram of the array if a sufficient number of receivers is available.

Even in rather complex subsurface models the localization works surprisingly well. A sparse receiver coverage, however, may lead to wrong localizations if no spreading corrections are applied. The reversely propagated wavefield interferes constructively at the source position and the highest amplitudes should be observed at this point. However, if we use only very few receivers, i.e., a heavily undersampled wavefield, we observe the highest amplitudes at the receiver locations. The wavefields back propagated from the receivers experience spreading losses in the reverse modelling and consequently display lower amplitudes than at the receiver position. A possible solution is to restrict the search area (or volume) to locations which are sufficiently far from the receivers and where constructive interference of the wavefield is already present. This, however, requires a priori information on the possible source location. This can be exploited when observing aftershocks or induced seismicity.

The physically more appealing solution is to account correctly for the removal of geometrical spreading during back propagation. The spreading losses collected during the propagation from the source to the receivers need to be removed in the reverse modelling in order to obtain the correct radiated amplitude. In the current localization method based on the wave equation this is crudely implemented by a simple time scaling of the images at every time step. Another option is the removal of geometrical spreading using a wavefield continuation method for back propagation which is based on high frequency asymptotics. The removal of geometrical spreading in reflection data is known from true amplitude migration Schleicher et al. (1993); Tygel et al. (1992).

The potential to update the velocities based on the focussing of the event are not investigated yet. The experience with migration velocity analysis in applied seismics, which is also based on the focussing of energy at a common depth point, is motivating to progress in this direction. Having the source location and the excitation time estimated by the method described above a mapping of reflectors below the source location can be performed in a pre-stack reverse-time migration manner or with another pre-stack depth migration algorithm.

An extension of the localization procedure to three dimensions is straight forward and was verified by a numerical example. The computational effort for the downward continuation using the wave equation increases considerably. Downward continuation techniques based on high frequency asymtotics may help with this issue.

The range of possible applications of the presented techniques appears to be very broad. Currently, many research initiatives are installing large passive arrays with 100+ receivers. These arrays will record the wavefield for extended periods and it is almost impossible to check every recorded trace for a possible seismic event. The method used in this paper can be applied almost in real time if a fast downward continuation method is used since no picking of arrivals is required. The data recorded by the array are stored in blocks. Each block represents a certain time, e.g., 10 s. This wavefield is then reversed in time and processed according to the method described above. The results can be displayed on a screen and inspected for focussed energy. Activity in the subsurface can be visualized almost instantaneously as bright spots on the screen.

The new approach might be also particularly valuable for areas with very weak seismic activity or poor recording conditions (low S/N ratio), for production monitoring in oil and gas reservoirs, for monitoring

of hydrofracs and seismic swarm events. The methodology can be also applied in areas which up to now appeared to be seismically inactive. The apparent inactivity might be a result of poor recording conditions and weak events which can not be detected in the individual seismograms of the array but may be visible in the stacked superposition of the whole acquisition obtained by the reverse modelling.

ACKNOWLEDGMENTS

We wish to thank the members of the Applied Geophysics Group in Hamburg for continuous discussion on the subject. We also thank Andrew Dobson for proofreading this manuscript. This work is partially supported by the sponsors of the WIT consortium.

REFERENCES

- Baysal, E., Kosloff, D., and Sherwood, J. (1984). A two-way non-reflecting wave equation. *Geophysics*, 49:132–141.
- Baysal, E., Kosloff, D., and Sherwood, J. W. C. (1983). Reverse-time migration. *Geophysics*, 48:1514–1524.
- Kao, H. and Shan, S.-J. (2004). The source-scanning algorithm: mapping the distribution of seismic sources in time and space. *Geophys. J. Int.*, 157:589–594.
- Kosloff, D. and Baysal, E. (1982). Forward modeling by a Fourier method. *Geophysics*, 47:1402–1412.
- Pujol, J. (2004). Earthquake location tutorial: Graphical approach and approximate epicentral location techniques. *Seis. Res. Let.*, 75(1):63–74.
- Schleicher, J., Tygel, M., and Hubral, P. (1993). 3D true-amplitude finite-offset migration. *Geophysics*, 58:1112–1126.
- Tessmer, E. (1990). Seismische Modellierung unter Berücksichtigung der freien Oberfläche mithilfe einer spektralen Tschebyscheff-Methode. PhD thesis, Universität Hamburg.
- Thurber, C. and Rabinowitz, N., editors (2000). *Advances in Seismic Event Location*, Modern Approaches in Geophysics, Lod, Israel. Geophysical Institute of Israel.
- Tygel, M., Schleicher, J., and Hubral, P. (1992). Geometrical spreading corrections of offset reflections in a laterally inhomogeneous earth. *Geophysics*, 57:1054–1063.
- Versteeg, R. and Grau, G. (1991). The Marmousi experience. In *Proceedings*. EAEG, Zeist, The Netherlands.
- Waldhauser, F. and Ellsworth, W. (2000). A double-difference earthquake location algorithm: Method and application to the Northern Hayward Fault, California. *BSSA*, 90:1353–1368.



Figure 1: Simple velocity model.



Figure 2: Seismograms from a source at (360, 400) m for the simple velocity model.



Figure 3: Snapshots for the reversely modeled data of the simple velocity model. The maximum amplitude is found for the image at 0 ms, i.e., for the exact excitation time.



Figure 4: Snapshots of the reversely modelled data for the simple velocity model using 10% too high velocities. The maximum amplitude is found at 30 ms, i.e., 30 ms too late.



Figure 5: Seismograms from a source at (360, 400) m for the simple velocity model with noise added to the data (S/N = 0.5). No events can be identified in the individual seismogram.



Figure 6: Snapshots of the reversely modelled data with S/N = 0.5 for the simple velocity model. The maximum amplitude is found in the image at 3 ms, i.e., 3 ms too late.



Figure 7: Snapshots of the reversely modelled data of the simple velocity model with $10 \times$ less receivers compared to Fig.3.



Figure 8: Velocities for a portion of the Marmousi model.



Figure 9: Seismograms from a source at (2800, 1520) m for the Marmousi model.



Figure 10: Snapshots of the reversely modelled data for the Marmousi velocity model. The maximum amplitude is observed at the exact excitation time.



Figure 11: Snapshots of the reversely modelled data of the Marmousi velocity model using 10% too high velocities.



Figure 12: Snapshots of the reversely modelled data for the Marmousi velocity model with $60 \times$ fewer receivers compared to Fig.10. The maximum amplitude is found at -2 ms, i.e., 2 ms too early.



Figure 13: Snapshots of the reversely modelled data for the 3-D simple model using 25 equally spaced receivers. The maximum amplitude is observed at the exact excitation time.