Feasibility of Real-Time Broadband Waveform Inversion for Simultaneous Moment Tensor and Centroid Location Determination

Fumiko Tajima*, Charles Mégnin,
Douglas Dreger and Barbara Romanowicz

Berkeley Seismological Laboratory,
University of California at Berkeley

Submitted to *Bull. Seism. Soc. Am.*, August 13, 2001

*To whom correspondance should be addressed. U.C. Berkeley Seismological Laboratory, 207 McCone Hall, Berkeley, California 94720 (fumiko@seismo.berkeley.edu)

Abstract

We present a feasibility study for an automated system to simultaneously determine centroid source location and seismic moment tensor (MT) for regional earthquakes. This system uses continuous real-time waveforms in a time window that is continuously shifted forward with a short time interval ($\sim$20 sec), from a sparse network of broadband seismic instruments, and unlike the currently standard method, performs inversion, without prior knowledge of the location and origin time information. We tested 68 earthquakes with $M_L \geq$4.2, that occurred in northern and central California between 1993 and 1999 and have well-calibrated solutions obtained by an established procedure of MT determination. The solutions determined by the new system are compared with the well-calibrated solutions for performance appraisal. Results show that onshore earthquakes with $M_w \geq$4.5 can be detected and adequately characterized in terms of MT and approximate centroid location. The threshold value of variance reduction to detect an event is $\sim$ 70%. In contrast, events that occured off the Mendocino coast are not as reliably constrained, reflecting the unmodeled complexity of the transitional structure from ocean to the continent using the current 1-D model, as well as the gap in azimuthal coverage. Significant improvement of the computational efficiency is expected if the system is fully configured on a powerful PC with multiple CPU's of $\sim$1 Gbyte memory, and the monitoring of event location and MT determination over the grid can be updated within the time frame of shifting the time window ($\sim$20 sec).

1 Introduction

The Berkeley Seismological Laboratory (*BSL*) operates a regional network (Berkeley Digital Seismic Network, *BDSN*) of 25 very broadband stations to monitor regional seismic activity in northern and central California (Romanowicz et al. (1992); Romanowicz et al. (1994)). The real-time waveform data are continuously telemetered to *BSL* using frame-relay. The broad frequency band and high dynamic range of this network is particularly suitable for detailed characterization of earthquakes (centroid location, moment-tensor (MT), and source characteristics). The current system of automated MT determination at *BSL* has been developed as part of the *BSL*/U.S. Geological Survey (*USGS*) joint notification system, *i.e.*, the Rapid Earthquake

Data Integration (*REDI*) (Gee et al. (1996) and Pasyanos et al. (1996)). It relies on the availability of earthquake onset times and locations, provided to *BSL* by the *USGS* at Menlo Park, where they are determined in real-time from the dense array comprised primarily of short-period ($\sim$1 sec) vertical component sensors (Northern California Seismic Network (*NCSN*), Oppenheimer et al. (1992)).

Efforts aimed at developing a stand-alone system have recently been implemented at *BSL* with an introduction of azimuth and P and S arrival-time information into the earthquake location determinations and the development of an adaptative grid search method for event location (Dreger et al. (1998), Uhrhammer et al. (2001)). This method is based on the arrival time (*i.e.* high frequency) information of P and S waves and therefore does not take full advantage of the bandwidth of broadband waveform data. In addition, the automatic S phase arrival pick and azimuth determination are not sufficiently reliable to provide an accurate source location using the sparse *BDSN* network data alone.

With the present quality of broadband seismic data, it is possible to determine source location and MT of significant events independently of the short period data. Kawakatsu (1998) presented the basic concept for a fully automated system that continuously monitors the background seismic wavefield and determines MT solutions for virtual point sources distributed on a grid in the region of interest. Variance reductions ($VR's$) calculated on the grid are evaluated at each point in space and time so as to identify a point of maximum VR that may exceed a threshold value corresponding to a real earthquake. Such an approach will be particularly relevant in areas of occurrence of less frequent but potentially large earthquakes where the operation of more than a few broadband stations may not be affordable.

In the present study, we test the feasibility of a system that detects an event and simultaneously determines its parameters (centroid location, origin time and MT) as quickly as possible by monitoring continuous real-time broadband waveform data from a subset of the *BDSN* broadband stations. The purpose of this system is to be able to retrieve accurate moment tensor parameters independently of short period data even if the earthquake location that is obtained is only approximate. In addition we test the system for a possible false alarm using data recorded for major teleseismic earthquakes or noise data.

2 Inversion Algorithm

Following Kawakatsu (1998), we briefly describe the algorithm. A seismogram $d^k(t)$ recorded at station k can be expressed as the convolution of the response of the medium to an impulsive source $G_i^{sk}(t)$ for three fundamental source orientations (*e.g.* Jost & Herrmann (1989)), with the moment tensor elements m_i for a source s decomposed into 5 elements, or 6 when a non-zero trace is solved for (*e.g.* Dreger et al. (2000)).

$$\sum_i G_i^{sk}(t)\, m_i^s \;=\; d^k(t), \tag{1}$$

The normal equation based on (1) is given by:

$$\sum_i A_{ji}^s\, m_i^s \;=\; b_j^s \tag{2}$$

where

$$A_{ji}^s = \sum_k A_{ji}^{sk}\,,\;\; A_{ji}^{sk} \;=\; \int G_j^{sk}(t) G_i^{sk}(t) dt \tag{3}$$

$$b_j^s = \sum_k b_j^{sk}\,,\;\; b_j^{sk} \;=\; \int G_j^{sk}(t) d^k(t) dt \tag{4}$$

(for simplicity the superscript s is omitted hereafter). Since $\mathbf{A}$ is a square matrix, the least squares solution for the moment tensor is given by:

$$\widehat{\mathbf{m}} \;=\; \mathbf{A^{-1}}\,\mathbf{b} \tag{5}$$

$\mathbf{A}$ is an auto-correlation of Green's functions. Given a structural model, it can be readily calculated and stored in the computer memory. $\mathbf{A}$ must be also stored in addition to Green's functions, however, it represents only a minimal increase in required space. Then the moment tensor vector $\widehat{\mathbf{m}}$ can be obtained by a simple matrix multiplication once $\mathbf{b}$ is calculated. This calculation is almost equivalent to a cross-correlation calculation between Green's functions and data for a given source location.

Then, the model prediction error (or residual between data and synthetic seismograms)

$$Res \;=\; \sum_k \int \left(d^k(t) - \;\mathbf{A(t)}\widehat{\mathbf{m}} \right)^2 \, dt \tag{6}$$

can be easily obtained. In the standard MT inversion (e.g., Dreger & Helmberger (1993) and Romanowicz et al. (1993)), variance reduction

$$VR = \left[1 - Res \div \left(\sum_k \int \left(d^k(t)\right)^2 dt\right)\right] \times 100\,(\%) \tag{7}$$

is introduced to evaluate the fitness between data and synthetics. In the proposed automated MT (hereafter *AMT*) determination, $VR's$ will be evaluated at all the grid points (i.e., virtual sources) to find a location that gives the maximum VR for a given time window. Here, in a real-time monitoring system, the time window applied to the continuous data stream is to be repeatedly shifted forward by a certain time interval. If a threshold value of VR is given for an event detection, then this approach can automatically detect an event location and origin time, and simultaneously computes the MT. There are critical issues in this inversion scheme, however, regarding the spatial and temporal resolution and monitoring performance during a teleseism or aftershock sequence.

3 Feasibility of Automated Moment Tensor Determination

In what follows, we test the feasibility of the *AMT* determination described in the previous section using the complete waveform (CW) MT method (Dreger & Helmberger (1993); Romanowicz et al. (1993)) and recent waveform data from the *BDSN*. An important issue in evaluating the new algorithm is whether a real earthquake can be detected, and its spatial and temporal resolution is satisfactory. We test the resolution, simulate the *AMT*, and evaluate the results by comparing the solutions with 'ground truth' solutions available from the *REDI* system.

3.1 Data

First we summarize the regional earthquakes available from the REDI project in the last ten years. The *USGS* provides the hypocentral locations, on-set times and magnitudes using data from the dense short period network and the *BSL* determines MT solutions for events with $M_L \geq 3.5$ using broadband waveform data. There are $\sim$68 events with $M_L \geq 4.2$ (through 1999). After

eliminating noisy data, 60 events are selected for this study (see Table 1, column 1, and the epicenters plotted with stars in Figure 1). As the noise levels and/or the station distributions relative to sources vary across the state, the northern and central California region is divided into 5 subregions (see Fig. 1):

- Northern California (excluding the Geysers region),
- The Geysers,
- South of the San Francisco Bay (excluding the Mammoth region),
- Mammoth,
- Offshore Mendocino.

In the *AMT* method, Green's functions ($G_i^{sk}(t)$ in eq. (1)) should be calculated for the paths from a grid of sources (s) to a number of stations (k) in advance. However, as the path calibration in northern California has not been completed yet, we use the approach of the standard MT inversion, and select a layered velocity model, either *SoCal* (Dreger & Helmberger (1993)) for the continental side or *GIL7* (Dreger & Romanowicz (1994)) for the coastal range of the state, according to the source (i.e., grid point) locations. *SoCal* is characterized by a crustal thickness of $\sim$ 35 km, and *GIL7* by a shallower Moho depth of $\sim$ 25 km and a slower upper crust than *SoCal*.

The existing set of Green's functions was produced using the frequency-wavenumber integration method of Saikia (1994) applied to the two multi-layered models. For each model, a catalog of Green's functions was constructed for a distance range between 30 and 700 km with an increment of 5 km and for source depths between 5 and 39 km with an increment of 3 km (note that the depth increment around $\sim$ 17 km is slightly different to avoid the layer boundary in *GIL7*), for a total of 2 (models) $\times$ 135 (distances) $\times$ 12 (depths) $= 3,240$ sets of 8 orthogonal elementary moment tensors. Figure 2 shows the grid points (i.e., these are 160 virtual point source locations) and the model regionalization, the models adopted, *GIL7* with open circles, and *SoCal* with shaded circles. The grid interval is set to be a half-degree.

We use five minute time-windows of data recorded at up to 6 *BDSN* stations BKS, CMB, HOPS, ORV, SAO, and YBH (see the triangles in Figure 2). These stations were selected to cover a broad range of azimuths in northern and central California, and because of their data quality. They provide

data in real time both to *BSL* and to the *BDSN* backup system situated at the Office of Emergency Services (*OES*) in Sacramento, to accomodate a possible disruption of the computer system at *BSL* due to an earthquake. The five minute time window can include all body and surface waves at any station for any event within the region. Figure 3 illustrates the waveforms recorded at five of the stations in a five minute time window containing an event.

3.2 Procedure of Analysis

To simulate the *AMT*, we use waveform data that include source signals of the events in Table 1. The waveforms in the given time window are inverted for all the virtual point sources distributed over the grid. The best fit origin time for a possible source is sought by cross correlating Green's functions of 120 sec with the data (see eq. (4)) to find the maximum of the $VR's$ (eq. (7)) at each of the grid points. After $VR's$ are calculated at all the points, the maximum is evaluated over the grid. Once an event is detected with VR exceeding a threshold value on the initial grid, a more precise location is sought by 'zooming' in the vicinity, over which a new mesh is set with a finer interval. Here the *AMT* system is illustrated using the seismograms from the 12/20/1994 Parkfield, CA earthquake (M_L=4.7, origin time 10:27:48). The successive inversions were performed in the period from 10:24:00 to 10:34:00 by shifting the time window forward every 20 seconds, initially on a coarse grid of 0.5° interval, and then on a finer grid of 0.25° around the location where a VR exceeded a threshold value.

Figure 4 shows four snap shot frames from this simulation. The seismograms from *BDSN* stations BKS, CMB, ORV, and SAO used for the analysis are shown at the bottom of each frame. In each frame, the waveforms on the left are the raw displacement data, and those on the right were filtered in the frequency band between 20 and 50 mHz, and used in the inversion. Here, only the vertical component traces are plotted while all three component data were used in the inversion. All the grid points are color coded to indicate the level of $VR's$. The best solution of the *AMT* (i.e., with the maximum VR) in each time window is shown with a red beach ball at the grid point, and for comparison the *REDI* MT solution with a blue beach ball and its epicenter with a red star.

Figure 4 a) is for the time window from 10:24:00 to 10:29:00 that does not contain much source signal (see the waveforms in the window), and the

maximum VR over the grid is 31.9%. The *AMT* solution obtained with this VR value is very different from the 'ground truth' solution provided by *REDI* (compare the *AMT* mechanism and location with the *REDI* solution). In Figure 4 b) the source signals are in the window (from 10:26:20 to 10:31:20), and the VR exceeds 90%. The mechanism of the detected event, which is located near the REDI epicenter, is similar to the REDI solution. Figure 4 c) shows the 'zooming in' on the 0.25° grid with a diameter of ∼80 km around the detected event location. The $VR's$ near the REDI location show high values of over 90%. In Figure 4 d) the source signals are disappearing from the time window, and the *AMT* and *REDI* solutions are diverging.

An animation in mpeg format that shows the successive inversion results for the entire period (from 10:24:00 to 10:34:00) by shifting the time window every 20 sec, is available upon request to the authors or at the URL:

ftp://quake.geo.berkeley.edu/outgoing/pepe/Parkfield.12_20_94.mpeg.gz

and the mpeg reader precompiled for many platforms is freely available from:

http://www.visiblelight.com/mpeg/software/index.htp

3.3 Resolution Tests

In the *AMT* practice, parameter resolution is an important issue when the seismic signals are within the time window. The *AMT* system in the present feasibility study uses the CW inversion method (Dreger & Helmberger (1993)) for MT determination. Here we first examine the temporal resolution in the CW/MT inversion scheme, and then test the resolution over the grid. The waveform data used for the tests are in a 5 minute time window that starts before the REDI origin time and includes real seismic signals for each event.

Temporal Resolution

The CW/MT method evaluates the cross-correlation between data and Green's function (120 sec) (eq. (4)) to estimate the event origin time that gives the maximum value of $VR's$ defined by eq. (7). An important issue is to examine if the CW/MT inversion scheme in the *AMT* can distinguish the maximum VR within the time window, and whether the origin time estimated with the maximum $VR's$ is comparable to the *REDI* origin time that is determined

with travel-time data from the dense short-period array. We performed successive inversions, for each event at the *REDI* epicenter (for a verification purpose), with the procedure that allowed the 'virtual' origin time to vary over a period of 2 min around the *REDI* origin time, with an increment of 1 sec. Here we applied a non-causal filter in the frequency band between 20 and 50 mHz, however, both the data and Green's functions are filtered in the same way that allows optimal parameter determination in the inversion. We also checked the effects of local maxima that could mislead the event detection capability (i.e., false alarm) during the continuous inversion.

Figure 5 shows an example of the temporal resolution tests using the data recorded for 1998/08/12, $M_L = 5.0$ San Juan Bautista event at stations BKS and CMB. The solid curve represents the $VR's$ as a function of 'virtual' origin time, and the *REDI* origin time is indicated with a thick vertical line. There are some local maxima of $VR's$ present in this time span, which are due to cycle shifts of the surface waves. However, the local maxima of $VR's$ were significantly lower than that of the global maximum (and below the threshold for an event detection). The time span of 2 min was sufficient to isolate the global maximum. As the 5 min time window is shifted forward by 20 sec in the present simulation, it is not likely that the local maxima would influence the successive inversions.

Generally we found that the estimated origin times are in good agreement with those by *REDI* except for events from the off-shore Mendocino region. The difference between the estimated and *REDI* origin times ΔT ranges between 1.8 and 4 sec for northern Californian events, between 1.5 and 2.7 sec for south of the San Francisco events, and 3.3 sec for the Geyser event. ΔT's for events in Mammoth are slightly larger ranging between 3.6 and 6.0 sec. ΔT's for off-shore Mendocino events scatter between -15 and 9.8 sec, poor temporal resolution, suggesting that Green's functions calculated with *GIL7* may be inadequate in this region.

Resolution over the Grid

The location resolution by the *AMT* is restricted by the grid interval and the adopted velocity model. We tested the *AMT* performance for all the events ($M_L \geq 4.2$) in Table 1 on two different grids, a coarse one with a 0.5° interval and a finer one with a 0.25° interval. As the present grid interval is still not fine enough, its location evaluation should also account for the *REDI* location relative to the nearest grid point. We classified the solutions

in three groups Q1, Q2, and Q3 according to the horizontal distance between the *AMT* location and the *REDI* location. If the location is determined at a grid point closest to the *REDI* location, the mislocation should be less than ~40 km, and Q1 is assigned. If the mislocation is greater than 40 km but less than or equal to 80 km, Q2 is assigned. Otherwise, Q3 is assigned for a mislocation of greater than 80 km. We also tested a suitable frequency band in the range between 10 and 100 mHz and summarized the results for on-shore events using the mislocation Q's in Table 2. The results for events in the off-shore Mendocino were poor in general, and not included in this table. The performance is generally more stable in the 20-50 mHz band, suggesting that the structural effects are not negligible in the broader range, especially at shorter periods.

Figure 6 (a) shows an example of spatial resolution tests using also the data recorded for the 1998 San Juan Bautista earthquake as above. The red mechanism shows the *AMT* solution at $121.5°W$ and $37.0°N$ with VR=93.8%, and $M_w = 5.1$, close to the *REDI* solution. In Figure 6 (b), the stability of the solution is examined by plotting the variation of the mechanisms around the *AMT* location on a finer ($\Delta x \sim 0.25°$) grid. The numbers above the beach balls are the $VR's$ determined at the locations. The distribution of $VR's$ displays a well-isolated peak with little variation in mechanism around it, surrounded by decreasing $VR's$ within a short distance. Figure 6 (c) shows the mechanisms obtained at the epicenter with the maximum VR for different depths between 5 and 39 km. The *REDI* depth is indicated by the thick vertical line at $\sim$ 8 km. This example shows that the mechanism remains relatively stable at the epicenter for different source depths, which is somewhat in agreement with the results of Kawakatsu (1998). However, MT inversion for different source depths are not always stable if the velocity structure is not well constrained or has strong variation in the region.

3.4 Summary

The S/N ratio is often poor in the bandpass filtered waveforms ($20 \leq$ f $\leq$ 50 mHz) of smaller events. Accordingly, the results are summarized only for the 30 events with $M_w \geq 4.5$ that have good S/N ratios. Table 3 shows the evaluation parameters obtained in this feasibility study on the 0.5° grid (center column) and on the 0.25° grid (right column) for the 21 onshore events. Thus, if a fixed bandwidth is employed, the minimum event that can be analyzed is approximately $M_w = 4.5$ in the *AMT* scheme.

μ is a measure to evaluate the difference between two moment tensor solutions $M_{ij}^{(1)}$ and $M_{ij}^{(2)}$ defined as:

$$\mu \equiv \sqrt{\frac{\sum_{i=1}^{3} \sum_{j=1}^{3} \left(\delta M_{ij}' \right)^2}{8}} \tag{8}$$

where $\delta M_{ij}' \equiv M_{ij}^{(1)'} - M_{ij}^{(2)'}$ and $M_{ij}' \equiv \frac{M_{ij}}{M_0}$, and M_0 is the scalar moment (Pasyanos (1996); Pasyanos et al. (1996)). The two focal mechanisms are considered to be essentially equivalent to each other if $\mu < 0.25$, and significantly different if $\mu > 0.50$. Examination of the results in Table 3 suggests that the mechanism consistency is good for events with high $VR's$ ($\geq 80\%$) in northern California and south of San Francisco bay regions, but less so in Mammoth and Geysers where geothermal related activities exist. The minimum VR level for an event detection ($M_w \geq 4.5$) is about 70%.

Table 4 presents the solutions for nine off-shore Mendocino earthquakes. A similar event detection threshold is obtained for these events, but the majority of events (seven out of nine) were assigned Q2 or Q3. The mechanism consistency between the *AMT* and *REDI* solutions varies with μ between 0.1 and 0.63. The poorer results for this region are likely due to unaccounted velocity structure.

The results in Tables 3 and 4 also allow us to examine the effects of the zooming-up in terms of improvement or deterioration of the solutions. This is summarized in Table 5 for three comparison parameters Δ, δz, and μ in the subregions. In principle the zooming should provide improvement in the solution parameters. It is confimed in northern California, the Geysers, and the south of San Francisco bay area. In Mammoth four out of ten events showed slight deterioration although the location quality is still Q1, and μ is <0.5. In the Mendocino region the quality of location determination is poor in general, seven out of nine events are assigned Q2 or Q3, and five out of nine events show deterioration in the zooming. This indicates that the propagation structure from ocean to the on-land stations is not represented by *GIL7*. Sufficient calibration for the propagation paths from virtual source locations to the stations will improve the resolution and accuracy attainable with 20 to 50 mHz waveforms in the *AMT* inversion.

4 Discussion

We presented the feasibility of a new automated system (*AMT*) that continuously monitors seismic waveforms and determines moment tensors using waveforms recorded by a sparse network of broadband stations without relying on any other location and onset time information. The advantage of the *AMT* is its independency of the information that is currently provided by dense short-period networks. Thus, it provides a standard capability in the event that the short period data are not available. To make this system useful, the spatial and temporal resolution of the parameter determination should be comparable to the standard system such as the *REDI*. Thus we tested the feasibility of the *AMT*, and also suggest the potential to improve computational efficiency of MT determination should the entire routine be fully configured on powerful PC's with multiple CPU's.

We used data of recent events in northern California, for which well calibrated MT solutions and other event parameters are available from the *REDI* project as well as waveforms from the *BDSN* archive. Although the tested system is still very preliminary, it is shown that it has ability to identify and characterize an event in a fully automated grid search scheme. The spatio- and temporal resolution can be satisfactory if the structural model is well constrained, and the grid is set with an appropriate grid interval. In the region considered, the system can determine the approximate centroid location on a given grid and MT for an on-shore event with $M_w \geq\sim 4.5$ with reasonable resolution in space and time. The suggested threshold value VR for an event detection is $\sim$70%.

Figure 7 shows the proportion of solution qualities Q1, Q2, and Q3 as a function of M_w, the top plot for on-shore events and the bottom for off-shore Mendocino events. The on-shore events all possess Q1 or Q2 solutions, with more events of Q1. The proportion of Q2 or Q3 is large for off-shore Mendocino events. In the off-shore Mendocino region, the currently used 1-D velocity model *GIL7* does not represent the transitional structure from ocean to continent well, making the *AMT* inversion unstable. A newly developed model for this region *Mend1* (Tajima et al. (2001)) will be tested in the *AMT* scheme in the near future.

Figure 8 examines the dependence of the results on the number of stations used for on-shore events in the *AMT*. The number of stations varied between 1 and 4. The top plot suggests that if two or more stations are used, the horizontal mislocations result in Q1 (i.e., $\Delta \leq 40$ km) for most of the events.

The vertical mislocation (δz) is less than 6 km for most of the events if two or more stations are used (middle plot). The maximum $VR's$ (bottom plot) are all above 70 % regardless of the number of stations as this is the result for events of $M_w \geq 4.5$ with good S/N ratios. The results suggest that the use of two or more stations is desiable and essential in the AMT.

There are a number of critical issues and problems to be resolved in order to make the AMT system practical.

Computation Speed

In the AMT determination the time window of waveforms is shifted forward by a prescribed amount Δt (ex. 20 sec in the present feasibility study). Therefore, the CPU time needed to compute $VR's$ for all the grid points and detect an event if any should be less or equal to Δt. Roughly speaking, the CPU time is proportional to the number of stations used, and on average about 1.5 sec on a 248 MHz Ultra-Sparc 2 workstation for a grid point with the depth varied using one station. The CPU time necessary to compute the $VR's$ over the entire grid using typically 3 stations is:

$$3\,(\text{stations}) \times 1.5\,(\text{seconds}) \times 160\,(\text{grid points}) = 12\,\text{min.}$$

At present, Green's functions for each grid point are read in from the disk, and this I/O process occupies significant portion of the computation time. The CPU time spent during this step is $\sim$0.4 sec. If Green's functions as well as the precomputed cross correlations of Green's functions (see Eq. (3)) are stored in the computer memory and the code is parallelized, the computation time can be cut down substantially. If the AMT is performed on a 10-node cluster of 500 Mhz PC workstations, the computation will be carried out in a much shorter time, somewhere within the time frame of the time shift of $\sim$ 20 sec, allowing continuous updates of AMT determination. As computer abilities continue to improve and computer prices drop, implementation of such a system will become more practical.

Path Calibration

To make the AMT system operational and computationally effective, path calibration will be necessary to account for the 3-D structure between virtual sources and stations. If Green's functions $G_i^{sk}(t)$ are calculated with well

constrained velocity models and stored in the memory, the computation time for eqs. (3) and (4) will be saved significantly, and the efficiency desired for the *AMT* system can be achieved. Sufficient path calibration will also enable the system to use a broader frequency band of waveforms with better resolution and stability. As knowledge about the 3D paths improves, the gridded structure of the methodology is ideally suited to incorporate such information.

Another challenge in this step is the availability of computer memory. If the horizontal and vertical grid size are set to be 0.1° and 3 km, respectively, there will be

$$50 \times 50 \times 13 = 32,500 \text{ (grid points)}$$

for a 5° × 5° × 39 km area. The total size of Green's functions for each station requires

$$120\,(\text{points for 2 min data}) \times 32,500\ (\text{points}) \times 8\,(\text{MT elements}) \times 4\,(\text{bytes}) = \sim 125\ (\text{Mb}).$$

If up to six stations should be used, the memory size will be about 750 Mb, and be slightly increased should the auto correlated Green's functions be also in the memory. This memory requirement is affordable if a dedicated powerful PC is available. In fact Dr. H. Kawakatsu and his colleagues at the Earthquake Research Institute at the University of Tokyo just started experimenting this scheme by implementing the package of the *BSL* CW/MT codes on a SGI 1100 PC with multiple CPU's, each with 1 Gbyte memory. While the path calibration efforts for the region of their interest will be substantial, the code modification for the *AMT* to run on the PC will be straightforward, and the computation efficiency can be tested before long.

False Event Detection?

An important issue in the *AMT* practice is to avoid false event detections. We carried out experiments using waveforms during a seismically quiet period and tested if $VR's$ could exceed a threshold value, i.e., $\sim$ 70% to detect an event. The highest $VR's$ obtained never exceeded 30%, indicating that background noise alone is not likely to be interpreted as an earthquake.

Another possible source of false triggering is a large teleseimic event. We performed a test using data from several teleseismic events with $M \geq 7.0$ such as the 9 June 1994 Bolivia earthquake. The maximum $VR's$ obtained in

the AMT for these earthquakes didn't exceed 30%, either. This experiment indicates that Green's functions calculated for the regional structure do not model teleseismic waveforms with high $VR's$ due to the huge differences in phase velocity, and that interpreting teleseismic events as regional events in northern California is not likely to occur.

Effective Use of *AMT*

Figure 9 illustrates the time frame of the AMT as compared with the present standard system. The standard system at *BSL* as part of *REDI* determines MT's when it receives an email for an event with $M_L \geq 3.5$ from USGS. The time frame between an event occurrence and MT determination is within 10 min in the present *REDI* system. If the AMT system monitors seismic wavefield continuously, and updates the regional MT search over the grid every 20 sec, the time interval between the origin time and MT determination will be shortened substantially, by several minutes or more.

As an effective application of the AMT method, a tsunami warning system can be considered as the event location and MT can be determined early enough before the tsunami waves arrive at the coasts. Here, the interval of the P-wave first arrival at the station (T_1) and the tsunami arrival at the coast (T_t in Fig. 9) is roughly estimated as

$$\Delta T = r_2/c - r_1/v \tag{9}$$

where r_1 is the distance between the epicenter and the station, r_2 the distance between the epicenter and the coast, c is the speed of tsunami wave propagation, v is the P wave speed, and ΔT is the interval between the P-wave first arrival and tsunami wave arrival. The tsunami phase velocity is estimated as $c = \sqrt{gd}$ where d is water depth and g is gravity acceleration (see the equation in P. 148 in Lay & Wallace (1995)). If r_1 and r_2 are roughly $\sim$ 200 km, c is $\sim$ 200 m/s, and v is $\sim$ 8 km/s, then ΔT is $\sim$ 975 sec (or $\sim$ 16 min). As the AMT can characterize the earthquake within several minutes of the P-wave arrival (T_3 - T_1), there is still $\sim$ 10 min before the tsunami arrival after the location and MT determination.

In the occurrence of a large earthquake with significant spatial extent, it is also useful to know the earthquake centroid in addition to the high frequency hypocenter. The search scheme over the grid in the AMT may have useful application to estimating the fault rupture dimension, and offer a

starting point to rapidly produce a map of regional ground shaking (Kaverina et al. (1998) and Dreger & Kaverina (2000)), and possibly the location of aftershocks.

Acknowledgements: C.Trifu (AE), L. Zhu and an anonymous reviewer provided helpful review comments that improved the manuscript substantially. We thank H. Kawakatsu for providing us with his preprint on the present subject before its publication, fruitful discussion, and collaboration. We also thank R. Uhrhammer, H. Tsuruoka, T. Urabe and Y. Nakayama for comments and discussion. The figures were prepared using GMT-3.3 (Wessel & Smith (1991)).

References

Dreger, D. S. & Helmberger, D. V., 1993, Determination of source parameters at regional distances with three-component sparse network data, *J. Geophys. Res.*, *98*, 8,107–8,125.

Dreger, D. S. & Kaverina, A., 2000, Seismic remote sensing for the earthquake source process and near-source strong shaking: a case study of the October 16, 1999 Hector Mine earthquake., *Geophys. Res. Lett.*, *27*, 1,941–1,944.

Dreger, D. S. & Romanowicz, B., 1994, Source characteristics of events in the San Francisco Bay region, *U.S. Geol. Surv.*, *Open-File Rept. 94-176*, 301–309.

Dreger, D. S., Uhrhammer, R. A., Frank, J., & Romanowicz, B., 1998, Regional and far-regional earthquake locations and source parameters using sparse broadband networks: A test on the Ridgecrest sequence, *Bull. Seism. Soc. Am.*, *88*, 1,353–1,362.

Dreger, D. S., Tkalčić, H., & Johnston, M., 2000, Dilatational processes accompanying earthquakes in the Long Valley Caldera, *Science*, *288*, 122–126.

Gee, L., Dreger, D. S., Neuhauser, D., & Romanowicz, B., 1996, The REDI program, *Bull. Seism. Soc. Am.*, *86*, 936–945.

Jost, M. L. & Herrmann, R. B., 1989, A student's guide to and review of moment tensors, *Seism. Res. Lett.*, *60, No 2*, 37–57.

Kaverina, A., Dreger, D. S., & Antolik, M., 1998, Source process of the 21 April 1997 Santa Cruz Island earthquake (M_w 7.8), *Geophys. Res. Lett.*, *25, No 21*, 4,027–4,030.

Kawakatsu, H., 1998, On the realtime monitoring of the long-period seismic wavefield, *Bull. Earthq. Res. Inst.*, *73*, 267–274.

Lay, T. & Wallace, T. C., 1995, *Modern Global Seismology*, Academic Press, New-York.

Oppenheimer, D., Klein, F. W., & Eaton, J. P., 1992, The fist 20 years of CALNET, the norhtern California seismic network, *U.S. Geol. Surv., Open-File Rept. 92-0209*, 33.

Pasyanos, M. E., 1996, Regional moment tensors and the structure of the crust in central and north California, *Ph.D. dissertation, University of California at Berkeley*, pp. xx.

Pasyanos, M. E., Dreger, D. S., & Romanowicz, B., 1996, Toward real-time estimation of regional moment tensors, *Bull. Seism. Soc. Am., 86, No 5*, 1,255–1,269.

Romanowicz, B., Gee, L., & Uhrhammer, R. A., 1992, Berkeley Digital Seismic Network: a broadband network for northern and central California, *IRIS Newslett., XI*, 1–5.

Romanowicz, B., Dreger, D. S., E., P. M., & Uhrhammer, R. A., 1993, Monitoring of strain release in central and northern California using broadband data, *Geophys. Res. Lett., 20*, 1,643–1,646.

Romanowicz, B., Neuhauser, D., Bogaert, B., & Oppenheimer, D., 1994, Accessing northern California earthquake data via the internet, *EOS Trans. AGU, 75*, 258–261.

Saikia, C. K., 1994, Modified frequency-wavenumber algorithm for regional seismograms using Filon's quadrature, modelling of Lg waves in eastern North America, *Geophys. J. Int., 118*, 142–158.

Tajima, F., Dreger, D., & Romanowicz, B., 2001, Berkeley moment tensor determination and path calibration for the transitional structure from ocean to continent in the Mendocino region, *Bull. Seism. Soc. Am., to be submitted.*

Uhrhammer, R. A., Dreger, D. S., & Romanowicz, B., 2001, Best practice in earthquake location using broadband three-component seismic waveform data, *Pure Appl. Geophys., 158*, 259–276.

Wessel, P. & Smith, W. H. F., 1991, New version of the Generic Mapping Tool released, *EOS Trans. AGU, 72* **441**, 445–446.

Source region	Number of ($M_L \geq 4.2$)	Tested Events ($M_w \geq 4.5$)
Northern California	8	4
Geysers	1	1
South of San Francisco Bay	12	6
Mammoth	25	10
Mendocino (offshore)	14	9
Total	60	30

Table 1 Number of events used in the study. Column 1: events ($M_L \geq 4.2$), from the initial pool; and column 2: events ($M_w \geq 4.5$ and recorded at two stations or more) used for the summary.

Frequency band	20-50 mHz (50-20 s)			10-100 mHz (100-10) s		
Location Q	Q1	Q2	Q3	Q1	Q2	Q3
No Cal	4	2	2	2	1	5
Geysers	1	0	0	1	0	0
So Bay	10	1	1	5	4	3
Mammoth	13	8	4	4	3	18
Total	28	11	7	12	8	26
(%)	(61)	(24)	(15)	(26)	(17)	(57)

Table 2 Comparison of inversion stability of the on-shore events between two different frequency bands. Numbers of events ($M_L \geq 4.2$) that fall in the respective quality bins as well as the corresponding percentages in the parentheses are shown. Q1 for $\Delta \leq 40$ km, Q2 for $40\,\text{km} < \Delta \leq 80$ km, and Q3 for $\Delta > 80$ km where Δ is the lateral distance between the *AMT* and *REDI* locations. Note that events with noisy data (one in northern California and two in south of the San Francisco Bay) were removed from the initial pool.

				0.5° grid					0.25° grid				
Date	Time	REDI (Lat)	Location (Long)	M_w	VR (%)	Δ (km)	δz (km)	μ	M_w	VR (%)	Δ (km)	δz (km)	μ
Northern California													
11/15/1995	20:33:59.76	39.62	-120.06	4.5	79.3	63.5	8.0	0.42	4.5	71.3	14.6	8.0	0.22
07/21/1998	08:38:49.88	40.62	-122.40	4.5	87.5	19.1	10.0	0.25	4.5	73.8	16.1	1.0	0.18
10/30/1998	09:53:31.39	39.29	-120.04	4.8	92.4	23.6	13.4	0.14	4.8	93.4	5.6	13.4	0.06
08/18/1999	01:06:18.93	37.91	-122.69	4.5	89.7	19.4	1.3	0.15	4.5	89.7	19.4	1.3	0.15
Geysers													
09/04/1995	14:16:17.66	38.68	-122.74	4.8	85.8	41.7	16.3	0.38	4.8	94.3	7.4	10.3	0.31
South of San Francisco Bay													
08/11/1993	22:33:04.00	37.31	-121.68	4.7	92.7	37.9	1.6	0.14	4.7	93.2	9.1	4.7	0.05
11/14/1993	12:25:34.87	35.95	-120.50	4.8	92.9	22.5	-3.7	0.14	4.8	92.3	5.3	-0.7	0.12
12/20/1994	10:27:47.17	35.92	-120.46	5.0	92.1	18.9	-4.1	0.21	5.0	90.5	9.7	-4.1	0.15
04/23/1995	08:41:36.63	36.60	-121.20	4.9	93.9	29.	3.2	0.16	4.9	95.3	12.0	0.2	0.12
05/21/1996	20:50:20.16	37.36	-121.72	4.8	75.3	44.5	18.9	0.55	4.8	75.3	44.5	18.9	0.55
08/12/1998	14:10:25.10	36.75	-121.46	5.1	93.8	28.	1.7	0.18	5.1	94.8	3.6	4.7	0.18
Mammoth													
11/02/1997	08:51:52.29	37.89	-118.14	5.2	90.8	45.0	-5.0	0.20	5.2	91.6	19.7	-8.0	0.14
11/22/1997	12:06:55.98	37.64	-118.92	4.5	77.1	16.7	15.6	0.52	4.5	78.3	19.5	-3.4	0.12
11/22/1997	17:20:35.14	37.64	-118.94	4.9	85.8	16.2	3.3	0.29	4.9	85.7	22.9	6.3	0.32
11/22/1997	18:10:59.45	37.63	-118.95	4.5	90.2	15.0	5.8	0.16	4.5	92.4	23.	8.8	0.14
11/30/1997	21:17:05.42	37.63	-118.95	4.9	86.7	15.	3.9	0.23	4.9	86.	15.	3.9	0.31
12/31/1997	20:36:47.34	37.63	-118.87	4.8	92.4	35.8	-1.6	0.23	4.8	92.4	35.8	-1.6	0.23
03/20/1998	10:43:03.14	37.91	-118.13	4.8	93.0	55.7	-2.2	0.29	4.8	94.5	20.4	-2.2	0.29
06/09/1998	05:24:40.16	37.59	-118.80	4.9	92.1	70.4	1.2	0.31	4.9	93.8	37.8	4.2	0.25
07/15/1998	04:53:19.25	37.56	-118.81	5.0	94.3	28.0	-1.2	0.11	5.0	94.8	34.8	-1.2	0.39
05/15/1999	13:22:10.66	37.53	-118.82	5.5	85.4	28.2	-0.6	0.13	5.5	85.8	7.	-0.6	0.14

Table 3 Summary of results for onshore events ($M_w \geq 4.5$) for 0.5° and 0.25° grid spacing. The subcolumns are the M_w determined by the *AMT*, VR, Δ (horizontal mislocation) that is the lateral distance between *AMT* and *REDI* locations, the vertical mislocation δz that is the depth difference between the *AMT* and *REDI* solutions. Δ is the horizontal mislocation (km) between the *REDI* and the *AMT* solutions. δz is the difference in depth (km) between the solutions. μ is a measure to evaluate the difference between *REDI* and *AMT* solutions defined in Eq. (8).

				0.5° grid					0.25° grid				
Date	Time	REDI (Lat)	Location (Long)	M_w	VR (%)	Δ (km)	δz (km)	μ	M_w	VR (%)	Δ (km)	δz (km)	μ
Mendocino													
10/23/1993	18:45:51.05	40.50	-126.76	4.8	80.1	64.3	27.9	0.32	5.1	81.3	70.3	33.9	0.32
06/19/1994	10:39:32.85	40.35	-124.47	4.9	87.4	48.2	7.8	0.09	4.9	87.7	75.7	7.8	0.26
09/01/1994	15:15:44.09	40.43	-126.71	6.9	76.2	60.6	18.3	0.09	6.9	77.5	81.7	18.3	0.09
12/26/1994	14:10:29.17	40.74	-124.30	5.4	77.2	33.4	-2.5	0.07	5.4	82.8	29.4	0.5	0.10
02/19/1995	04:03:14.27	40.59	-125.83	6.6	90.4	53.5	18.7	0.17	6.6	91.0	19.1	18.7	0.11
03/01/1995	10:55:36.33	40.64	-125.67	4.6	83.6	32.1	0.0	0.20	4.6	84.4	50.3	0.0	0.11
08/08/1996	09:19:04.72	40.35	-124.75	4.7	85.9	65.6	21.0	0.14	4.7	87.0	64.6	18.0	0.11
01/22/1997	07:17:16.75	40.27	-124.39	5.6	62.3	184.9	3.4	0.70	5.6	69.7	220.1	15.4	0.63
11/27/1998	00:43:48.93	40.66	-125.32	5.4	84.7	46.0	12.7	0.30	5.4	84.7	46.0	12.7	0.30

Table 4 Same as Table 3 for offshore Mendocino events ($M_w \geq 4.5$) for 0.5° and 0.25° grid spacing.

	Δ			δz			μ	
Event location	Imp./ Unaf.	Det.	Mean (km)	Imp./ Unaf.	Det.	Mean (km)	Imp./ Unaf.	Det.
Northern California	4	0	17.5	4	0	2.3	4	0
Geysers	1	0	34.3	1	0	6	1	0
South of San Francisco Bay	6	0	16.1	4	2	-0.1	6	0
Mammoth	6	4*	9.0	7	3	0.	6	4**
Mendocino	4	5	-7.6	7	2	-2	7	2
Total	21	9*	2.3	23	7	0.2	24	6

Table 5 : Summary of the results after 'zooming'. Imp. = Improved; Unaf = unaffected; Det. = Deteriorated. First column: for the horizontal distance Δ. Second column: for the vertical distance δz. Third column: for the MT difference parameter μ. Improvement in each parameter is indicated by a positive number. * indicates that the location deterioration is within Q1 criterion ($\Delta \leq 40$ km), and ** that $\mu is still within 0.5$.

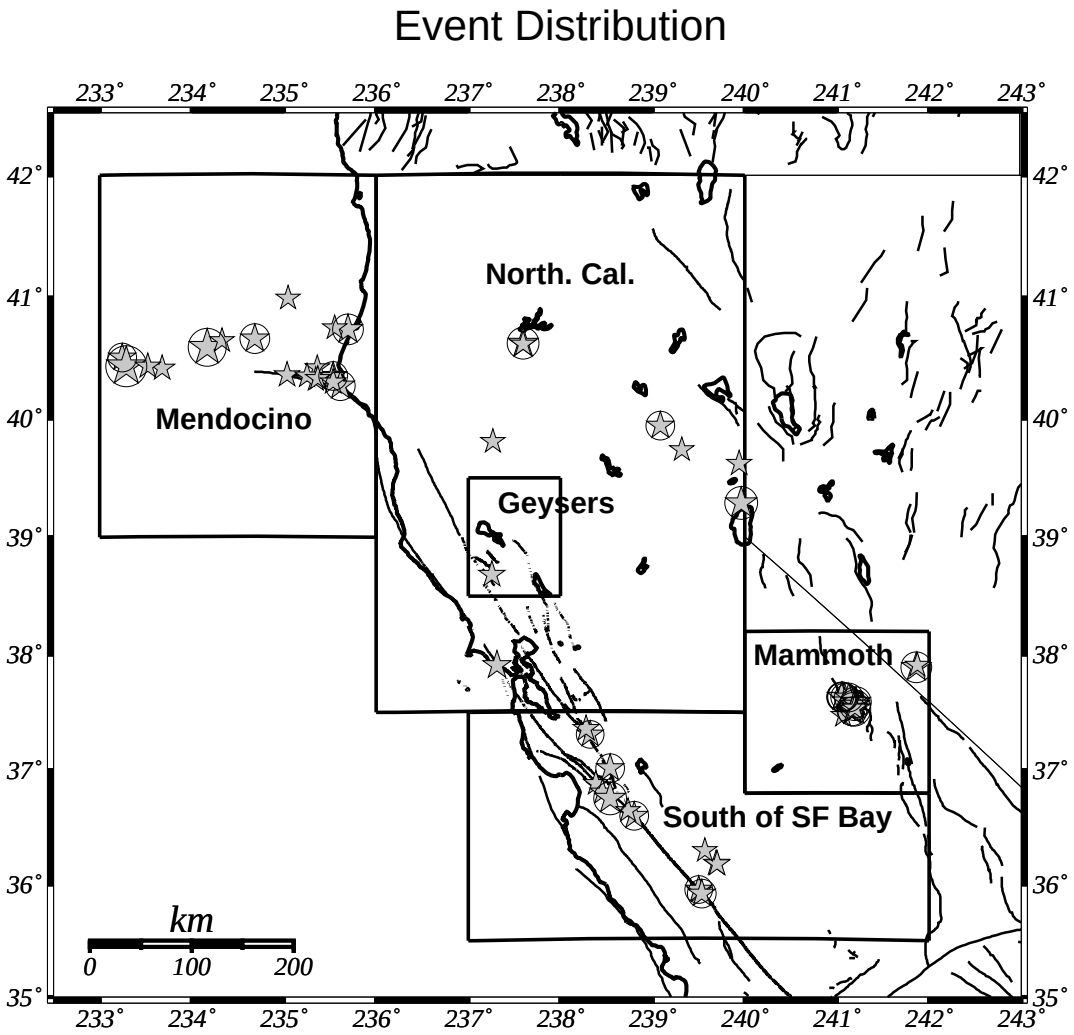

Figure 1 : Events used in this study are plotted at the REDI-determined epicenters in the five subregions. The circles indicate 32 events with $M_w \geq 4.5$.

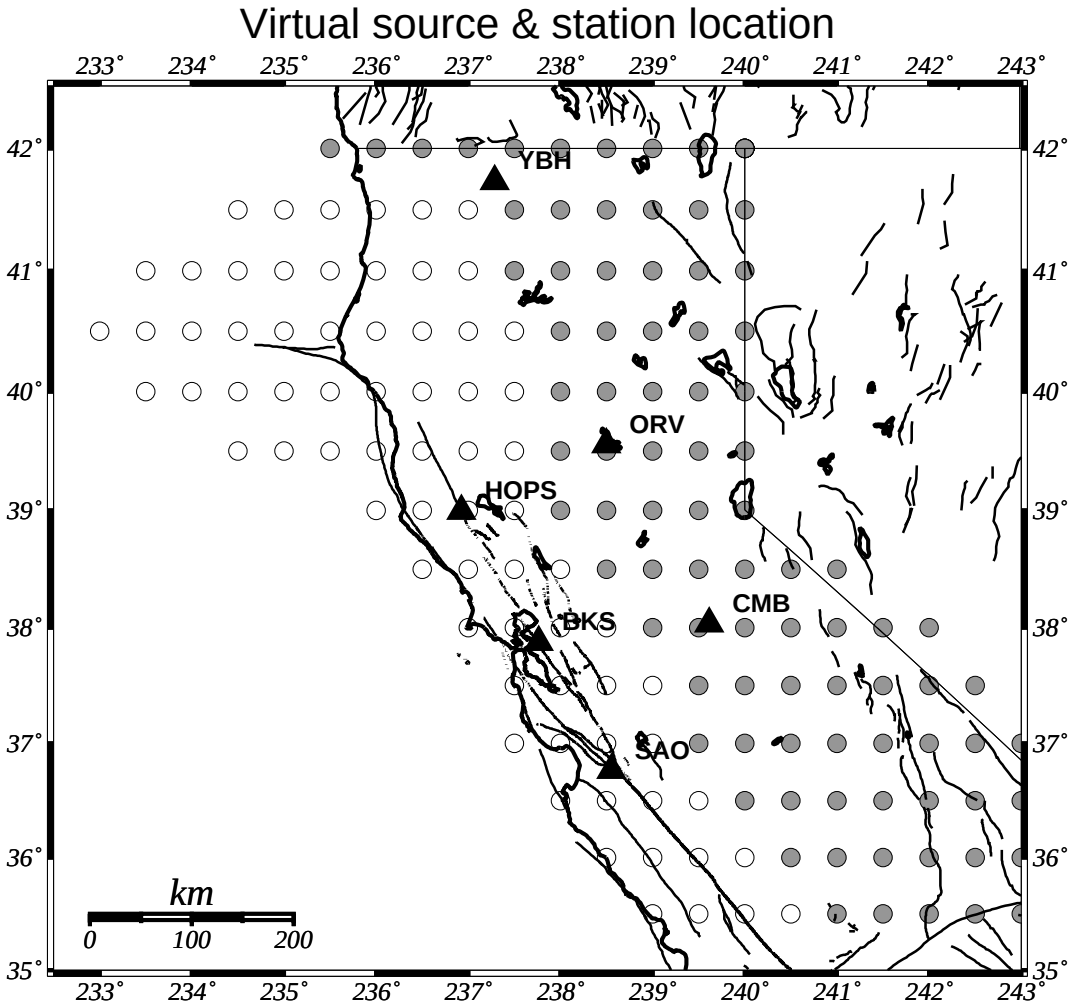

Figure 2 : Virtual source points (shown with 160 small circles) on the grid and station locations (triangles). The source locations are distinguished between the models: open circles for GIL7 and shaded circles for SoCal.

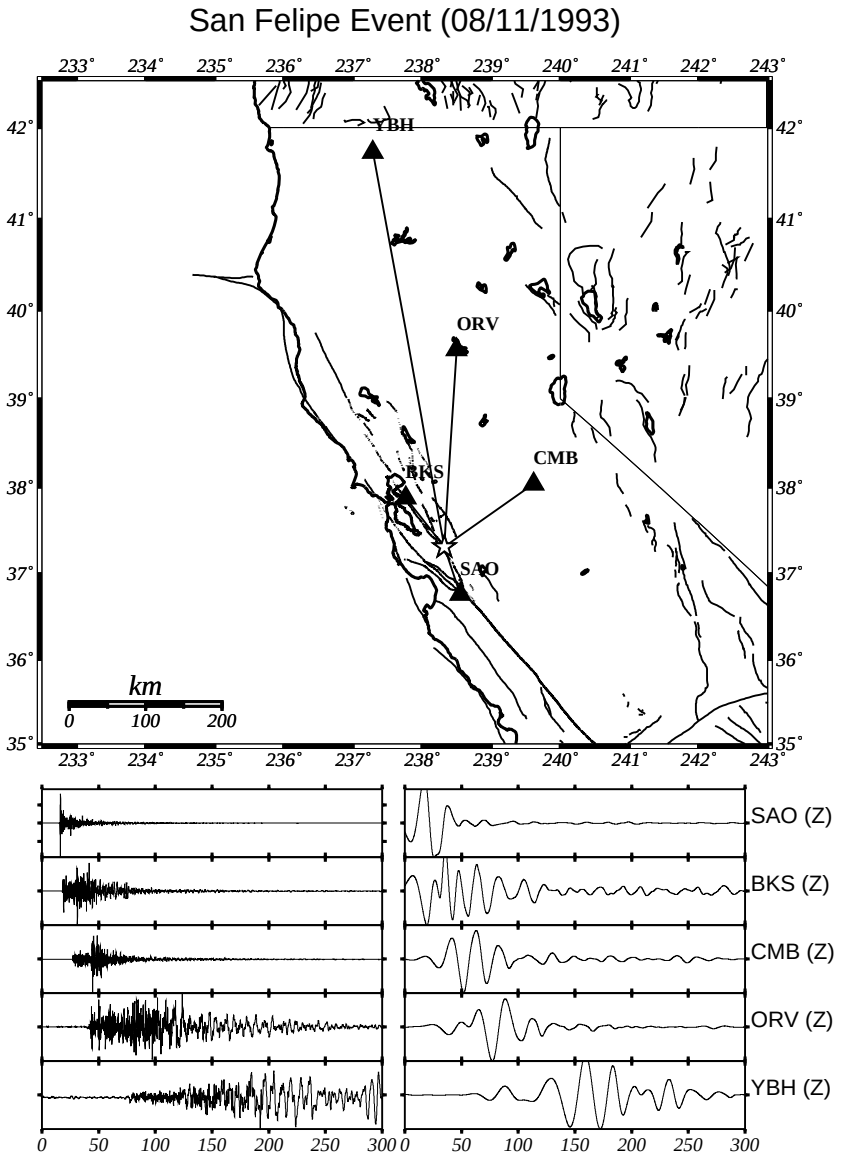

Figure 3 : (Top) Source and station location paths for the 8/11/1993 San Felipe earthquake ($M_L = 4.6$). (Bottom) (left) Raw waveforms and (right) waveforms filtered in the 20-50 mHz band.

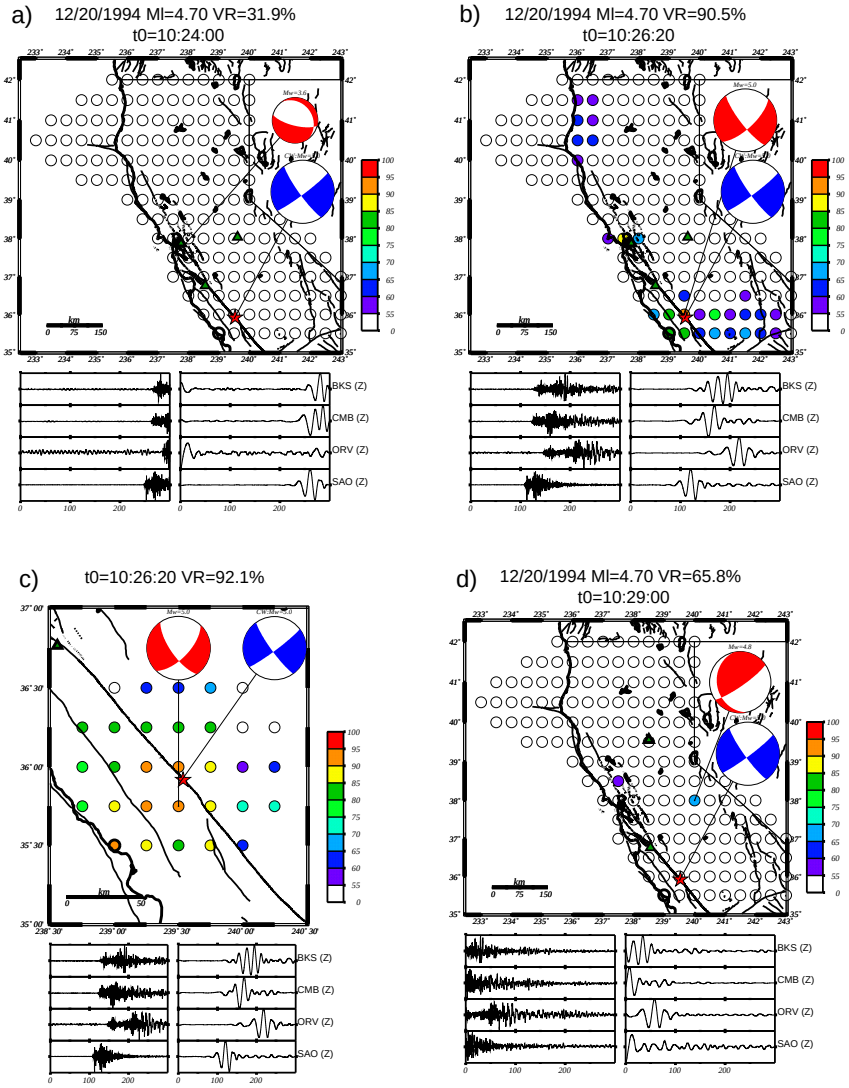

Figure 4 : Snapshots of the *AMT* simulation for the 12/20/1994 Parkfield event (M_L=4.7, origin time 10:27:48). Waveforms in the five minute window that is shifted every 20 sec are inverted for all the grid points (color coded to show the level of VR) from 10:24:00.00 to 10:34:00.00. The begining time of the window (T0) and the maximum VR in each window are listed above each frame. The *AMT* solution is shown with a red beach ball, and the *REDI* solution with a blue beach ball at the red star. a) Source signals barely enter the window; b) much of the source signals are included in this time window; c) *AMT* inversion performed over a finer grid (Δx $\sim$0.25°) in a smaller area around the detected location; and d) the front part of the seismic wavetrain is outside of the widow (see the text for details).

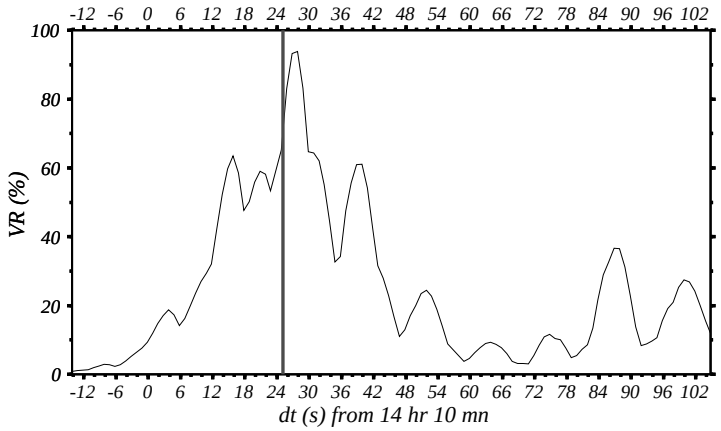

Figure 5 : Example of a temporal resolution test of the *AMT* inversion using the data from the 1998/08/12, $M_L = 5.4$ earthquake. The waveform data in a 5 minute time window are cross-corelated with Green's function to find a best 'origin time' which is shifted forward every 1 sec assuming the *REDI* location. The variance reduction is shown as a function of virtual 'origin time. The thick vertical line is the *REDI*-determined origin time.

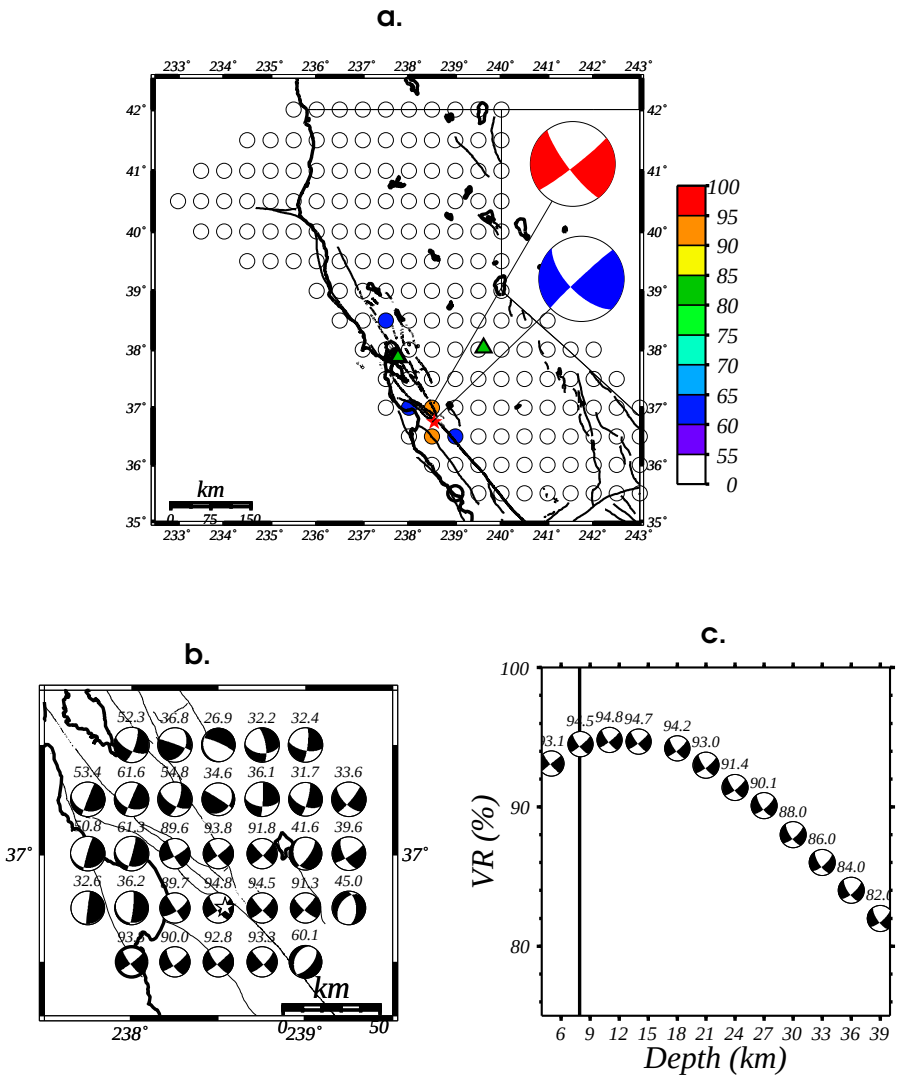

Figure 6 : a. Example of variance reduction distribution on a coarse grid (with the interval of 0.5° for the 1998/08/12, $M_L = 5.4$, San Juan Bautista earthquake. The grid points are color coded to show the level of $VR's$. The red beach-ball points to the location with the maximum variance reduction of 89.7%. The red star shows the *REDI* epicenter with its solution (blue beach ball). The green triangles are the stations used for this event. b. VR distribution on a finer grid (i.e., 'zoomed in' with Δ x=0.25°). The numbers above the mechanisms show the $VR's$ at the location and the red star the REDI location. c. The $VR's$ as a function of depth.

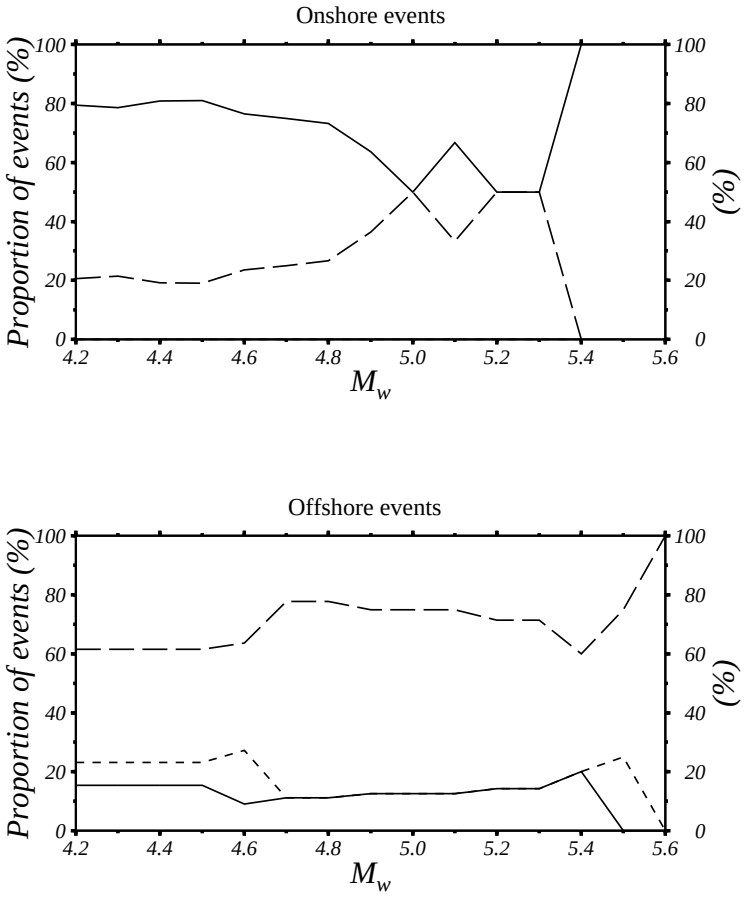

Figure 7 : Proportion of events with Q1 (solid line), Q2 (long dashed line) and Q3 (short dashed line) as a function of M_w. (top) onshore events; and (bottom) off-shore mendocino events. Note that all Q3 events are located off-shore Mendocino.

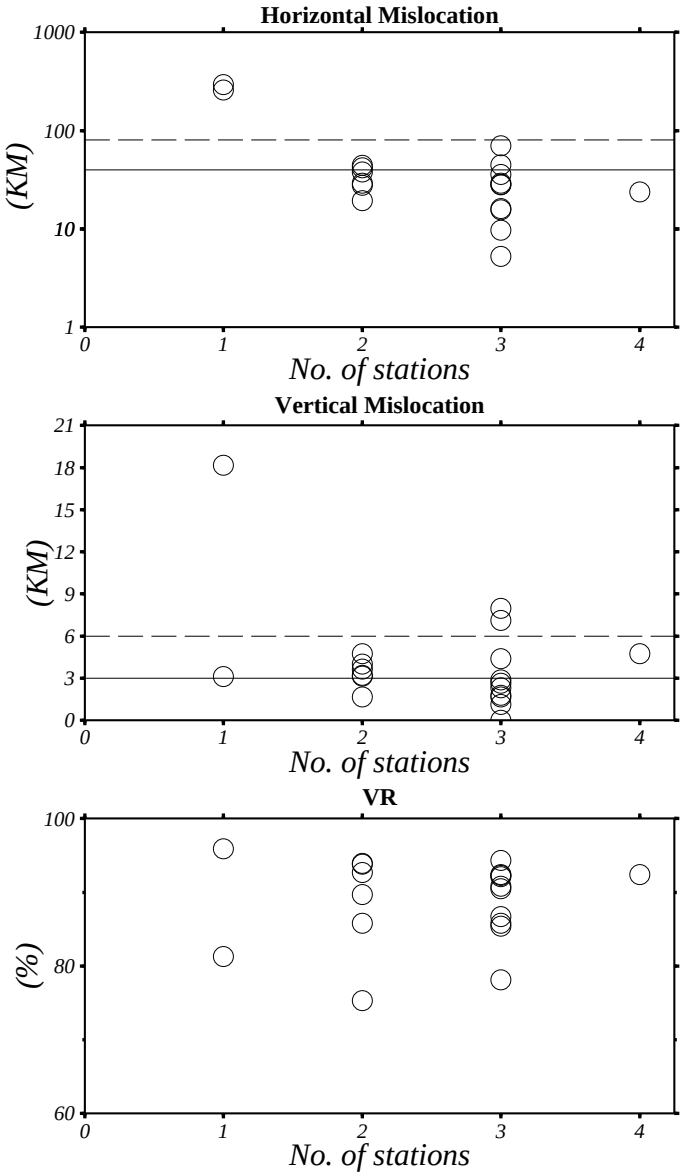

Figure 8 : Effects of the number of stations on the inversion results using the on-shore events. Top: the horizontal mislocation. Middle: the vertical mislocation. Bottom: the maximum variance reduction VR. Q1 solutions are plotted below the solid line, Q3 solutions above the dashed line, and Q2 solutions are in between.

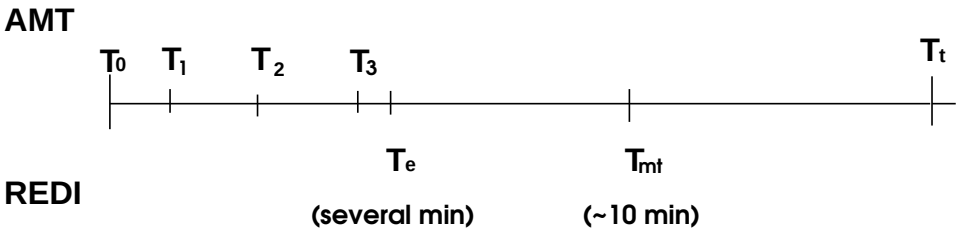

Figure 9 : Timing of event occurrence at T_0, P-wave arrival at T_1, grid search for location and MT determination starting at T_2, and completing at T_3 in the AMT system, and an event notification arrival by email at T_e and MT determination at T_{mt} in the REDI system. T_t illustrates tsunami arrival time for an epicentral distance of $\sim$200 km.