Real-time Estimation of Fault Rupture Extent using Envelopes of Acceleration

Masumi Yamada[1], Thomas Heaton[2]

masumi@eqh.dpri.kyoto-u.ac.jp, heaton_t@caltech.edu

Abstract:

We present a new strategy to estimate the geometry of a rupture on a finite fault in real time for earthquake early warning. We extend the work of Cua and Heaton who developed the Virtual Seismologist (VS) method (Cua, 2005), which is a Bayesian approach to seismic early warning using envelope attenuation relationships. This paper extends the VS method to large earthquakes where fault finiteness is important. We propose a new model to simulate high-frequency motions from earthquakes with large rupture dimension: the envelope of high-frequency ground motion from a large earthquake can be expressed as a root-mean-squared combination of envelope functions from smaller earthquakes. We use simulated envelopes of ground acceleration to estimate the direction and length of a rupture in real time. Using the 1999 Chi-Chi earthquake dataset, we have run simulations with different parameters to discover which parameters best describe the rupture geometry as a function of time. We parameterize the fault geometry with an epicenter, a fault strike, and two along-strike rupture lengths. The simulation results show that the azimuthal angle of the fault line converges to the minimum uniquely, and the estimation agrees with the actual Chi-Chi earthquake fault geometry quite well. The rupture direction can be estimated at 10 seconds after the event onset and the final solution is achieved after 20 seconds. While this methodology seems quite promising for warning systems, it only works well when there is an adequate distribution of near-source stations.

[1]Kyoto University,
Gokasyo, Uji, 611-0011, Japan
[2]California Institute of Technology,
MC104-44, 1200 E. California Blvd. Pasadena CA 91125 USA

1 Introduction

Recently, with advances in data analysis and increased awareness of the seismic hazard, the topic of earthquake early warning has attracted more research attention and various early warning methods have been proposed from seismologists and engineers (Nakamura, 1988; Allen and Kanamori, 2003; Odaka *et al.,* 2003; Wu and Kanamori, 2005). Currently, the most ambitious system is the earthquake early warning system provided by the Japan Meteorological Agency, which is released to the public in October, 2007. The news of the system was broadcasted widely and attracted considerable public attention in Japan. The goal of seismic early warning is to initiate optimal mitigating actions based on the arrival time and amplitude of seismic waves predicted at a given location. To achieve this, an earthquake early warning system must collect and quickly analyze seismic data in a manner that can be used to predict future shaking. In principle, this could be achieved by using the present value of an approximately known wavefield as a boundary condition to predict future wavefields using Navier's equation (Baker *et al.*, 2005). However, from a practical viewpoint, there are advantages to data analysis schemes that involve characterization of the earthquake source. Predictions of future shaking can be achieved by utilizing the extensive existing work on predicting ground shaking from seismic sources. Ideally, an early warning system would provide the best estimate of slip in time and space that can be deduced from seismic data available at any given instant in time. We introduce a two-step strategy to accomplish this; 1) we determine the spatial and temporal extent of an ongoing rupture by analyzing waveform envelopes of high-frequency shaking, 2) we determine approximate slip from simple projections of long-period shaking onto the approximately known location of the rupture. It seems likely that true ground displacement may be very helpful in the real-time determination of fault slip. Since it is difficult to obtain displacement from inertial seismometers (Clinton, 2004), the future incorporation of real-time high-sample-rate GPS into early warning systems may be quite important. A more complete description of our strategy to determine distributions in real time is given by Yamada (2007).

Cua and Heaton developed the Virtual Seismologist (VS) method (Cua, 2005). It is a Bayesian approach to seismic early warning designed for modern seismic networks, and proposed for small to moderate earthquakes with ruptures that can be approximately modeled as a point source. The VS algorithm uses an envelope attenuation relationship and the predominant frequency content from the first few seconds after the P-wave arrival. The advantage of the VS method is its capacity to assimilate different types of information that may be useful to find quick and reliable estimates of magnitude and location (Cua, 2005). It gives the best estimate of an earthquake property in terms of a probability density function. The Bayesian approach is a scheme to emulate human capabilities to judge complex information by modeling uncertainty in a probabilistic way.

Our goal is to extend the VS method to large earthquakes where fault finiteness is important. Although there are some researchers trying to characterize fault finiteness in real time (Nakamura, 2001), most of earthquake early warning systems focus on estimating epicenters and magnitudes of earthquakes, not the fault geometry (Allen and Kanamori, 2003). However, for large earthquakes, rupture length can be on the order of tens to hundreds of kilometers, and the inhomogeneous slip distribution significantly affects the ground motion

amplitude at a site. For example, the fault rupture in the 1999 Chi-Chi earthquake was longer than 80 km, and the largest slip was recorded at the northern end of the fault. It would be difficult, if not impossible, to predict such large shaking at large distances from the epicenter when using a scheme that only characterizes the earthquake as a point source.

In this paper, we introduce a methodology to determine the direction and length of the rupture from high-frequency ground motions. We construct a model to describe high-frequency motions from earthquakes with large fault dimension. By applying this model to high-frequency ground motions in real time, it becomes possible to estimate the geometry of an ongoing earthquake.

2 Data

The data for this analysis is the strong motion dataset from the 21 September 1999 Chi-Chi Earthquake that occurred in central Taiwan (Lee *et al.,* 2001). The epicenter was located at 120.82 N, 23.85 E, with a focal depth of 8 km according to the Central Weather Bureau (CWB) of Taiwan (Shin and Teng, 2001). It is currently the largest well-recorded earthquake with moment magnitude 7.6. 441 strong motion stations recorded the main event, and 69 of those were at distances of less than 50 km from the epicenter. We use three component (NS, EW, and UD) strong motion records from the data set collected by CWB. They classified the recorded accelerograms into four quality groups based on the existence of absolute timing, pre-events, and defects. For this analysis, we use QA-class data (best for any studies), QB-class data (next best but no absolute timing) and a part of QC-class data (covering the principal strong motions but not having pre-event or post-event data) which includes the pre-event. The distribution of stations is shown in Figure 1.

Table 1 describes the crustal model for P-wave and S-wave velocity in central Taiwan (Ma *et al.,* 1996). P-wave and S-wave arrival time for the predicted envelope are computed with this 1-D layered crustal model. Since the original seismic records reported incorrect universal time, we use the data modified by Lee *et al.* (2001). They compared picked P-wave arrival times with computed theoretical P-wave arrival times. If the P-time residual was larger than 1 second for accelerograms at the distance within 50 km, they corrected the P-wave arrival time (Lee *et al.,* 2001). Therefore, the error of the time stamp of the modified data is less than 1 second.

We use envelopes of acceleration records in horizontal and vertical components. The horizontal components are calculated by the square root of sum of squares of the EW and NS components. Acceleration envelopes are obtained by taking the maximum absolute amplitude of the ground motion time history over a 1 second window. This one second sampling acceleration envelopes reduces the computational efforts.

3 Method

3.1 Ground Motion Models

Cua and Heaton examined over 30,000 seismograms in Southern California and developed relationships that predict waveform envelopes (Cua, 2005). Each seismogram is parameterized with a set of eleven parameters, and each parameter is characterised as a function of magnitude and epicentral distance. Therefore, the envelope ground motion $E(t|M,R)$ at time t can be expressed as a function of magnitude (M) and epicentral distance (R). The detail of the ground motion models are explained in the Appendix. An observed ground motion envelope and the best P-wave, S-wave, and ambient noise decomposition are shown in Figure 2. This ground motion model is similar to the traditional strong motion attenuation relationships (e.g., Boore *et al.*, 1993) with time domain transition. However, unlike attenuation relationships that have been developed primarily from large earthquakes, the data set of Cua and Heaton included earthquakes with magnitudes ranging between 2 and 7. In order to characterize this data, it was essential to utilize relationships in which the distance decay of amplitudes change as a function of magnitude, similar to that which was introduced by Campbell (1981). In particular, accelerations recorded close to a rupture saturate at magnitudes larger than 6, whereas distant sites do not demonstrate comparable saturation as a function of magnitude.

3.2 Subsource Model

We introduce a one-dimensional source model to express the fault finiteness based on the ground motion models by Cua and Heaton (Cua, 2005). The fault surface is divided into subfaults, and each subfault is represented by a single point source, called "subsources" (Figure 3). To simplify the problem, we assume that the dimensions of all subsources are uniform. Each source nucleates, and the P- and S-waves are radiated when the rupture front arrives at the subsource.

The ground motion at a site is modeled as the combination of the responses of each subsource. An envelope of the ith subsource is expressed as follows:

$$E_i(t|M,R) = E(t - t_i|M_i, R_i(\theta)) \tag{1}$$

where M_i is the magnitude of the subsource i, $R_i(\theta)$ is the distance between the station and subsource i which is a function of the azimuthal angle (θ) of the rupture direction, and t_i is the time delay due to the rupture propagation.

For high-frequency motions with approximately random phase, the square root of the sum of the squares of the envelope amplitudes from each subsource provides a good estimation of an acceleration envelope (Cocco and Boatwright, 1993):

$$E_{total}(t|M,R) = \sqrt{\sum_{i=0}^{N1} E_i^2(t|M,R) + \sum_{i=1}^{N2} E_i^2(t|M,R)} \tag{2}$$

where N1 and N2 are number of the point sources to the north and the south and $E_{total}(t|M, R)$ is the estimated envelope as a function of time. $E_i(t)$ is actually a fairly complex function of time, magnitude, and distance, although its forward calculation is very fast since it only involves analytic functions (Cua, 2005; Cua and Heaton, 2006).

This model only works for high-frequency ground motions. Unlike longer-period ground motions, high-frequency motions seem to be insensitive to either radiation pattern (Liu and Helmberger, 1985) or directivity (Boatwright and Boore, 1982). Furthermore, near-source high-frequency motions saturate as a function of magnitude. That is, near-source high-frequency ground motions are independent of the amplitude of the slip for large earthquakes (Kanamori and Jennings, 1978; Cua and Heaton, 2006).

Heaton and Hartzell (1989) pointed out that the assumptions of a Brune (1970) source spectrum combined with constant stress drop leads to high-frequency energy radiated from a subfault that is independent of the slip on the subfault. A consequence of the fact that high-frequency near-source ground motions can be modeled as random noise whose amplitude is independent of slip is that the high-frequency radiated energy in earthquakes is proportional to the rupture surface area. This is consistent with the observation of Boatwright (1982) who showed that high-frequency spectral acceleration amplitudes are proportional to the root-mean-square (rms) dynamic stress drop and the square root of the rupture area. Our simple model for simulating high-frequency motions is also compatible with the observation of Hanks and McGuire (1981) that high-frequency ground accelerations are remarkably similar from one event to another. Subsources for our source model are evenly spaced, so the surface area and high-frequency radiated energy corresponding to each subsource are also constant. Based on this theoretical interpretation, we estimated the ground motion envelopes with the subsource model for the 1999 Chi-Chi earthquake.

Figure 4 (top) shows an example of predicted envelopes for vertical accelerations using the subsource model. It shows the envelopes of the vertical acceleration record for each subsource with magnitude 6.0. Figure 4 (bottom) shows the time history envelope of the accelerogram (vertical component) at the station C024, a station on the foot wall side and 10 km from the Chelungpu fault line (shown in Figure 1, southwest of the epicenter). Figure 4 also shows that the vertical acceleration envelopes predicted by the subsource model for the VS-FS (finite source) method fit the observed envelopes much better than the envelopes predicted by the single source model for the VS-PS (point source) method.

Even though the Ch-Chi rupture has large spatial variations in the amplitude of the slip, it appears that the high-frequency accelerations can be modeled as a sum of the radiation from a uniform tiling of the magnitude 6.0 subfaults, based on the random-phase assumption and saturation with regard to magnitude.

3.3   Comparison between Subsource Model and Observations

Figures 5 and 6 are a comparison of observed envelopes and predicted envelopes from the best fit source model (explained in the next section). The model consists of 12 subsources distributed along a line trending 17 degrees clockwise from north; there are 7 subsources north of the epicenter and 4 subsources to the south. The predicted acceleration envelopes

for this model agree well with the observed envelopes. Predicted envelopes of near-source stations have some discrepancy depending on the source process, but predicted envelopes of far-source stations fit the observation well.

The vertical predicted envelopes of the stations in the epicentral region (e.g. stations T078, T079, T084, and T089) are of particular interest. The predicted envelopes overestimate these observed envelopes for the first 10 seconds, but then underestimates the observed records 20 seconds after the earthquakes origin. The fact that the largest accelerations in the epicentral region occurred 20 seconds after the origin time seems to indicate that there may have been some rupture complexity in the hypocentral region; perhaps there was an early aftershock in the epicentral region 20 seconds after the first rupture. Although this feature is noteworthy, it does not have a significant effect on the inversions since the epicentral stations are less important for estimating azimuthal angle and length of the fault.

Note that there is a discrepancy between the predicted and observed horizontal envelopes of the stations located about 40 km north of the recognized northern terminus of the Chelungpu fault rupture (e.g. stations T045, T047, and T095). The observed records have a sharp pulse that appears about 40 seconds after the event onset in the records of stations. Shin and Teng (2001) suggested that these large accelerations were generated by a secondary rupture, perhaps on the Shihtan fault.

3.4 Estimating the Fault Parameters

We performed an inversion of acceleration envelopes to retrieve fault model parameters. We assume that the location of the epicenter is already estimated from the VS-PS method, and that the fault ruptures bilaterally from the epicenter with constant rupture velocity. Thus, the time delay for each subsource rupture is the distance from the epicenter divided by the rupture velocity. Therefore, parameters that we need to estimate from the observed data are the azimuthal angle (θ) of the rupture direction, N1 and N2, that are used to simulate each of the segments of the bilateral rupture. Three parameters are computed at each second using only the data available at that time.

The best estimate of the model parameters minimizes the residual sum of the squares (RSS) between observed ground motion envelopes and predicted envelopes from the subsource model. The misfit function as a measure of goodness of fit is defined as follows:

$$RSS(t) = \sum_{i=1}^{ns} \sum_{j=1}^{2} \sum_{k=1}^{t} (A_{ijk} - \hat{A}_{ijk})^2 \tag{3}$$

where ns is the number of stations, t is the time in 1 second intervals ($\Delta t = 1$) from the event onset, and A_{ijk} and $\hat{A}_{ijk}$ are observed and predicted envelopes of component j (horizontal and vertical) at station i at time $k\Delta t$.

This form of the misfit function tends to emphasize the importance of fitting stations with large accelerations. That is, distant stations have small observed and predicted accelerations and even if there are serious discrepancies in the ratio of the predicted and observed amplitudes, they will have little impact on the inversion. In the course of this study, we also tried inversions in which we defined misfit function in terms of log of amplitudes. This misfit

function emphasizes the ratio of predicted and observed amplitudes; large amplitude data is no more important than small amplitude data. However, we found that such a misfit function emphasized misfits in the coda for near-source data and furthermore the distant data was often not well explained by our simple descriptions of wave envelopes that have been developed to explain the "average" effects of waves propagating through the crust. That is, it is important to emphasize the data from the near-source records and a logarithmic misfit function was not appropriate to recover the timing and location of the rupture.

Our parameterization scheme has the advantage that we characterize the source with relatively few parameters (θ, N1, N2), none of which require high precision estimates. However, for this strategy to be effective, we will need to solve a nonlinear inverse problem in real time. In this study, we solve this nonlinear inverse problem by using the Neighborhood Algorithm (Sambridge, 1999a; Sambridge, 1999b). We recognize that other inverse techniques may ultimately be chosen for real-time applications. However, since the purpose of this study is to determine the effectiveness of our parameterization, we use the Neighborhood Algorithm to characterize and solve this nonlinear inverse problem.

The Neighborhood Algorithm is a direct search method for finding models of acceptable data fit in a multidimensional parameter space. We generate samples in the parameter space and draw the Voronoi cells for these samples. Voronoi cells are nearest neighbor regions defined under a suitable distance norm, and the shape and the size of each Voronoi cell is determined by the sample distribution in the parameter space. See Figure 1 as an example of Voronoi cell as used to define the nearest neighbors to seismic stations. We calculate the misfit function for each sample and choose the model with the minimum misfit. New samples are generated by performing a uniform random walk in the chosen Voronoi cell. By repeating these steps, we will find a set of samples that identifies those regions of the parameter space that provide the best fit to the data. This is an approach for constructing the posterior probability density function from the ensemble samples based on the Voronoi cell concept (Yamada, 2007).

4 Results of the Simulation for the Chi-Chi Earthquake

We have run many different inversions by varying both the inversion parameters and the data sampling, such as the number of records used for the inversion, the components of the records, etc. In this section, we show the results of the main simulation. We use the horizontal and vertical records of the stations within 120 km of the epicenter. To simplify the problem, we assume each subsource has the same magnitude of 6.0 and the same spacing of 10km.

4.1 Result of the Fault Parameter Estimate

Figure 7 shows the estimation results for three parameters (azimuthal angle of fault line (θ), number of the point sources to the north (N1) and to the south (N2)). These three

parameters are computed at each second using only the data available at that time. The estimation is updated every second as the ground motion data are observed. Although it does a good job at characterizing the rupture length and timing, we see that it is difficult to resolve θ until 15 seconds after the event onset since the event can be approximated as a point source at the beginning. The estimated θ at 15 seconds is about -15 degrees and it increases gradually after 20 seconds due to a impulsive acceleration arrival at station C080 which is located at the south of the epicenter. Estimates of θ stabilize at about 14 degrees with respect to additional data after 27 seconds. There is an additional small shift at 45 seconds, at which point the inversion achieves its final solution of 17 degrees, which compares favorably with the observed average fault strike of the Chelungpu fault rupture (the fault trace is shown in Figure 1).

Since the subsources are equally spaced, the length of the fault is represented by the number of the point sources to the north (N1) and to the south (N2). Figure 7 (bottom) shows values of N1 and N2 as a function of time after the origin. From the figure, we can see the fault length grows bilaterally along the dashed black lines. Around 30 seconds, the rupture stops growing to the south. It also stops to the north temporarily, but it grows again around 40 seconds. This is due to the delayed high-frequency radiation at stations north of the Chenlungpu surface rupture and may have been caused by rupture on the Shihtan fault. Even though the result of the simulation fits the actual location of the fault accurately, the subsource model does not consider "rupture jumping dislocations" (i.e. the rupture at the adjacent active faults triggered by the main shock) (Shin and Teng, 2001). The final result shows 7 point sources to the north and 4 point sources to the south. This fault length is comparable to the total length from the Chelungpu fault to the Shihtan fault in Figure 1.

The best fit source model consists of 12 subsources distributed along a line trending 17 degrees clockwise from north; there are 7 subsources north of the epicenter and 4 subsources to the south. That is, the best fitting source parameter is given by (θ=17 degrees, N1=7, N2=4). The predicted acceleration envelopes for this model agree well with the observed envelopes, as shown in Figures 5 and 6. Predicted envelopes of near-source stations have some discrepancies depending on the source process, but predicted envelopes of far-source stations fit the observation well.

4.2   Sensitivity Analysis of Station Sampling and Rupture Velocity

Sensitivity of the station distribution and rupture velocity are also tested. We assume a constant rupture velocity of 2.0 km/s to construct the predicted envelopes from subsources. In order to check the sensitivity of the parameter estimate to the rupture velocity, we run four simulations with different rupture velocities. Figure 8 shows the estimated parameters, N1 and N2, for rupture velocities varying from 2.0 km/s to 3.5 km/s. The simulation results are relatively insensitive to the choice of rupture velocity. Since the most probable rupture velocity minimizes the misfit between observed envelopes and predicted envelopes, we can include this rupture velocity as a parameter to be estimated for future analysis.

In order to test the effect of station distribution, some of the stations are removed from the dataset. By randomly sampling from the total of 239 stations, we created two subsets of the dataset: subset A consists of 126 records with an even station code number and

subset B includes 56 records with a station code ending in 6 or 8 (e.g. T078). Although the station distribution is not homogeneous, the average station density is 214 km²/station for the subset A, and 482 km²/station for the subset B. The stations are located in an area of about 27000 km². Even though the station density is different, the estimated parameters are quite similar. In Figure 7, the time series of θ and N2 for full dataset, subset A, and subset B are almost the same. N1 for subsets A and B stays around 5 after 30 seconds, and the increase observed in full dataset due to the Shihtan fault rupture does not appear. The reason is that several near-source stations of the Shihtan fault have an odd number station code and are not included in this analysis (e.g. T045, T047, and T095). Considering that the rupture of the Shihtan fault is quite small compared to that of the Chelungpu fault, Subsets A and B can express the Chi-Chi earthquake rupture well. The VS-FS method for large earthquakes works well even if the station density is reduced to a quarter of the original density, as long as the station distribution is uniform.

4.3   Geometry of the Parameter Space

We have solved the optimization problem in parameter space (θ, N1, and N2) by the Neighborhood Algorithm. Here, we discuss the geometry of the parameter space.

Figure 9 shows the error surface of θ and N1 at a fixed N2 of 4 and assuming that all data is used in the inversion. The surface is smooth and has a deep and narrow valley at $\theta = 10$ degrees. The solution easily converges to this minimum. Figure 10 shows the error surface of N1 and N2 at a fixed θ of 17 degrees. The surface is very smooth in both N1 and N2 directions. The global minimum is very sensitive to the choice of the dataset, as shown in the results of subsets A and B.

Contour maps of the error surface of N1 and N2 at 10 second intervals are shown in Figure 11. θ is fixed at 17 degrees which is the optimal final solution. At 10 seconds, the minimum of this error surface is (N1, N2) = (0, 1). However, it is not the global minimum in the parameter space since $\theta = 17$ degrees is not the optimal solution at 10 seconds. At 20 and 30 seconds, the minimum of the error surface is at the maximum N1 and N2 in the possible parameter space even though θ is not optimal. There is a strong possibility that the rupture is still ongoing at this point. At 40 seconds, the minimum of the contour is around (N1, N2) = (5, 4) and it suggests that the rupture has stopped rupturing toward the south. After 40 seconds, the shape of contour map does not change much, and the elliptic shape of the smallest contour indicates that N2 is determined uniquely, but that considerable uncertainty about N1 remains.

5   Discussion and Conclusions

We outlined a strategy to estimate slip in time and space for an ongoing earthquake rupture. A key aspect of this strategy is to map the location of the rupture using envelopes of high-frequency acceleration data. Once the location of the rupture is estimated, long-period displacement data can be projected back onto the fault to determine the slip in real time (see Yamada, 2007).

Our strategy for using high-frequency radiation to determine the timing and length of the rupture relies on the observation that high-frequency seismic waves can be modeled as random-phase waves whose total radiated energy scales linearly with the rupture area. By using this assumption, we show that we can simulate the ground motion of a large earthquake by tiling the surface of the large event with smaller events and then summing the random phase signals from the smaller events. In our example of the Chi-Chi earthquake, we showed that a sum of 10 km interval magnitude 6.0 subevents provided a good prediction of the acceleration envelopes for this earthquake. In order to turn this simulation into a real-time inverse, we parameterize the rupture with as a linear alignment of magnitude 6.0 earthquakes. We then invert for the azimuth angle of the alignment as well as two integers N1 and N2, which are the number of additional 10 km patches in the positive and negative directions from the epicenter, respectively.

The best estimate of the model parameters minimizes the residual sum of the squares between observed ground motion envelopes and predicted envelopes from the subsource model (in equation (3)). This misfit function with linear amplitudes of ground motions can provide better estimates than that of logarithmic amplitudes, since the linear misfit function tends to emphasize the importance of fitting stations with large amplitudes.

Our study of the Chi-Chi data set indicates that it is more difficult to determine rupture length than it is to determine rupture azimuth. Furthermore, for this method to work well, an adequate near-source station distribution is important.

We are currently investigating alternate strategies to determine the temporal evolution of the rupture. In particular, we have developed a method to distinguish whether or not a particular seismic record is from a station within 10 km of a rupture surface (Yamada *et al.,*, 2007). Real-time mapping of an going rupture using this strategy becomes a simple matter of tracking the spatial evolution of near-source seismic stations. Although this strategy appears promising, it requires adequate station coverage to track near-source stations.

Appendix: Ground Motion Models

Cua and Heaton examined over 30,000 seismograms in Southern California and developed relationships that predict waveform envelopes as a function of magnitude, distance and station corrections (Cua, 2005). First, the envelopes of the ground motions are modeled as a combination of the envelopes of P-wave, S-wave, and ambient noise.

$$E_{observed}(t) = \sqrt{E_P^2(t) + E_S^2(t) + E_{ambient}^2} + \epsilon \tag{4}$$

where

$$
\begin{aligned}
E_{observed}(t) &= \text{envelope of observed ground motion} \\
E_P(t) &= \text{envelope of P-wave} \\
E_S(t) &= \text{envelope of S-wave and later-arriving phases} \\
E_{ambient} &= \text{ambient noise at the site} \\
\epsilon &= \text{difference between predicted and observed envelope}
\end{aligned}
$$

The ambient noise, $E_{ambient}$, for a given time history is modeled as a station constant. The P- and S-wave envelopes, $E_P(t)$ and $E_S(t)$, are defined by a rise time (t_{rise_P} and t_{rise_S}), a constant amplitude (A_P and A_S), a duration (Δt_P and Δt_S), and two decay parameters (γ_P and γ_S) and (τ_P and τ_S) respectively. The general form of the envelope function is:

$$E_{i,j}(t) = \begin{cases} 0 & ; \quad t < T_i \\ \frac{A_{i,j}}{t_{rise_{i,j}}}(t - T_i) & ; \quad T_i \le t < T_i + t_{rise_{i,j}} \\ A_{i,j} & ; \quad T_i + t_{rise_{i,j}} \le t < T_i + t_{rise_{i,j}} + \Delta t_{i,j} \\ A_{i,j}\frac{1}{(t - T_i - t_{rise_{i,j}} - \Delta t_{i,j} + \tau_{i,j})^{\gamma_{i,j}}} & ; \quad t \ge T_i + t_{rise_{i,j}} + \Delta t_{i,j} \end{cases} \quad (5)$$

where

$$i \quad = \quad \text{P-, S-wave}$$
$$T_i \quad = \quad \text{P-, S-wave arrival times}$$
$$j \quad = \quad \text{acceleration, velocity, and displacement}$$
$$\text{(horizontal and vertical components)}$$

They parameterized each seismogram as a set of eleven parameters (five for the P-wave envelope, five for the S-wave envelope, and one for the ambient noise). Furthermore, each parameter is described by magnitude, distance, log of distance, and site dependent constants based on the traditional strong motion attenuation relationships (Campbell, 1981; Boore *et al.*, 1993). The functional forms which describe the P- and S- wave envelope functions are given below:

$$\log_{10} A_{ij} = a_{ij}M + b_{ij}(R_1 + C_{ij}(M)) + d_{ij}\log_{10}(R_1 + C_{ij}(M)) + e_{ij} + \epsilon_{ij}, \quad (6)$$
$$\log_{10} B_{ij} = a_{ij}M + b_{ij}R_1 + d_{ij}\log_{10}R_1 + e_{ij} + \epsilon_{ij}, \quad (7)$$

where

$$i = \text{P-, S-wave,}$$
$$j = \text{horizontal and vertical acceleration, velocity, and displacement,}$$
$$A_{ij} = \text{ground motion envelope amplitude,}$$
$$B_{ij} = \text{rise time } (t_{rise}), \text{ duration } (\Delta T), \text{ and decay parameters } (\tau, \gamma),$$
$$M = \text{local magnitude } (M_w \text{ for } M > 5.0),$$
$$R = \text{ epicentral distance in km for } M < 5,$$
$$\text{closest distance to fault for } M > 5.0 \text{ (when available),}$$
$$R_1 = \sqrt{(R^2 + 9)},$$
$$C_{ij}(M) = (\arctan(M - 5) + 1.4)(c_{1ij}\exp(c_{2ij}(M - 5))),$$
$$a_{ij}, b_{ij}, c_{1ij}, c_{2ij}, d_{ij}, e_{ij} = \text{regression constants,}$$
$$\epsilon_{ij} = \text{statistical (or prediction) error, } \sim NID(0, \sigma^2).$$

The A_{ij}s are the ground motion envelope amplitudes (P- or S-wave) from fitting equations 4 and 5 to the observed ground motion envelopes in the database. The B_{ij}s are the parameters characterizing the envelope function (t_{rise}, ΔT, τ, and γ). Coefficients in equations 6 and 7 are determined by regression analysis of the database using the Neighborhood Algorithm (Cua, 2005).

Acknowledgments

References

- Allen, R. M., and H. Kanamori (2003). The potential for earthquake early warning in Southern California, Science, 300, 685-648.
- Baker, T., R. Granat, and R. W. Clayton (2005). Real-time Earthquake Location Using Kirchhoff Reconstruction, Bull. Seism. Soc. Am. 75, 699-707.
- Boatwright, J. (1982). A dynamic model for far-field acceleration, Bull. Seism. Soc. Am. 72, 1049-1068.
- Boatwright, J., and D. M. Boore (1982). Analysis of the ground accelerations radiated by the 1980 Livermore Valley earthquakes for directivity and dynamic source characteristics, Bull. Seism. Soc. Am. 72, 1843-1865.
- Boore, D. M., W. B. Joyner, and T. E. Fumal (1993). Estimation of response spectra and peak accelerations from western North American earthquakes: An interim report, U.S. Geol. Surv. Open-File Rept. 93-509, 72 pp.
- Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes, J. Geophys. Res., 75, no.26, 4997-5009.
- Campbell, K. W. (1981). Near-source attenuation of peak horizontal acceleration, Bull. Seism. Soc. Am. 71, 2039-2070.
- Cocco, M. and J Boatwright (2004).The envelopes of acceleration time histories, Bull. Seism. Soc. Am. 83, 1095-1114.
- Clinton, J. F. (2004). Modern digital seismology : instrumentation, and small amplitude studies in the engineering. Ph.D. thesis, Department of Civil Engineering, California Institute of Technology.
- Cua, G. (2005). Creating the Virtual Seismologist: developments in earthquake early warning and ground motion characterization. Ph.D. thesis, Department of Civil Engineering, California Institute of Technology.
- Cua, G. and T. H. Heaton (in press). The Virtual Seismologist method: a Bayesian approach to seismic early warning. In "Seismic early warning", eds: P. Gasparini, J. Zschau, Spring Heidelberg.

- Hanks, T. C. and R. K. McGuire (1981). The Character of High-frequency Strong Ground Motion, Bull. Seism. Soc. Am. 71, 2071-2095.
- Heaton, T. H. and S. H. Hartzell (1989). Estimation of Strong Ground Motions from Hypothetical Earthquakes on the Cascadia Subduction Zone, Pacific Northwest, Pageoph 129, 131-201
- Ji, C., D. V. Helmberger, D. J. Wald and K. F. Ma (2003). Slip history and dynamic implication of 1999 Chi-Chi earthquake. J. Geophys. Res., 108(B9), 2412, doi:10.1029/2002JB001764.
- Kanamori, H. and P. C. Jennings (1978). Determination of local magnitude, ML, from strong motion accelerograms, Bull. Seism. Soc. Am. 68, 471-485.
- Lee, W. H. K., T. C. Shin, K. W. Kuo, K. C. Chen, and C. F. Wu (2001). CWB Free-Field Strong-Motion Data from the 21 September Chi-Chi, Taiwan, Earthquake, Bull. Seism. Soc. Am. 91, no.5, 1370-1376.
- Liu, H. L., and D. V. Helmberger (1985). The 23:19 aftershock of the 15 October 1979 Imperial Valley earthquake: more evidence for an asperity, Bull. Seism. Soc. Am.75 ,689-708
- Ma, K. F., J. H. Wang, and D. P. Zhao (1996). Three-dimensional seismic velocity structure of the crust and uppermost mantle beneath Taiwan, J. Phys. Earth, 44, 85-105.
- Nakamura, Y. (1988). On the urgent earthquake detection and alarm system (UrEDAS), Proc. of the 9th World Conference on Earthquake Engineering, Tokyo-Kyoto, Japan. VII. 673-678.
- Nakamura, Y. (2001). Fault movement recorded by UrEDAS during the 2001 Geiyo Earthquake, Japan, Proc. of the 26th JSCE Earthquake Engineering Symposium, Sapporo, Japan. 1409-1412 (in Japanese).
- Odaka, T., K. Ashiya, S. Tsukada, S. Sato, K. Otake, and D. Nozaka (2003). A New Method of Quickly Estimating Epicentral Distance and Magnitude from a Single Seismic Record, Bull. Seism. Soc. Am. 93, no.1, 526-532
- Sambridge, M. (1999a). Geophysical Inversion with a Neighbourhood Algorithm-I. Searching a Parameter Space. Geophys. J. Int. 138, 479-494.
- Sambridge, M. (1999b). Geophysical Inversion with a Neighbourhood Algorithm-II. Appraising the Ensemble. Geophys. J. Int. 138, 727-746.
- Shin, T. C. and T. L. Teng (2001). An Overview of the 1999 Chi-Chi, Taiwan, Earthquake, Bull. Seism. Soc. Am. 91, no.5, 895-913.
- Wessel, P., and W. H. F. Smith (1991). Free software helps map and display data, EOS. Trans. AGU 72, 441, 445-446.
- Wu, Y. M. and H. Kanamori (2005). Experiment on an onsite early warning method for the Taiwan Early Warning System, Bull. Seism. Soc. Am. 95, 347-353.
- Yamada, M. (2007). Early warning for earthquakes with large rupture dimension. Ph.D. thesis, Department of Civil Engineering, California Institute of Technology.
- Yamada, M., T. H. Heaton and J. L. Beck (2007). Real-time Estimation of Fault Rupture Extent Using Near-source versus Far-source Classification, Bull. Seism. Soc. Am. (in press).

Figures and Tables

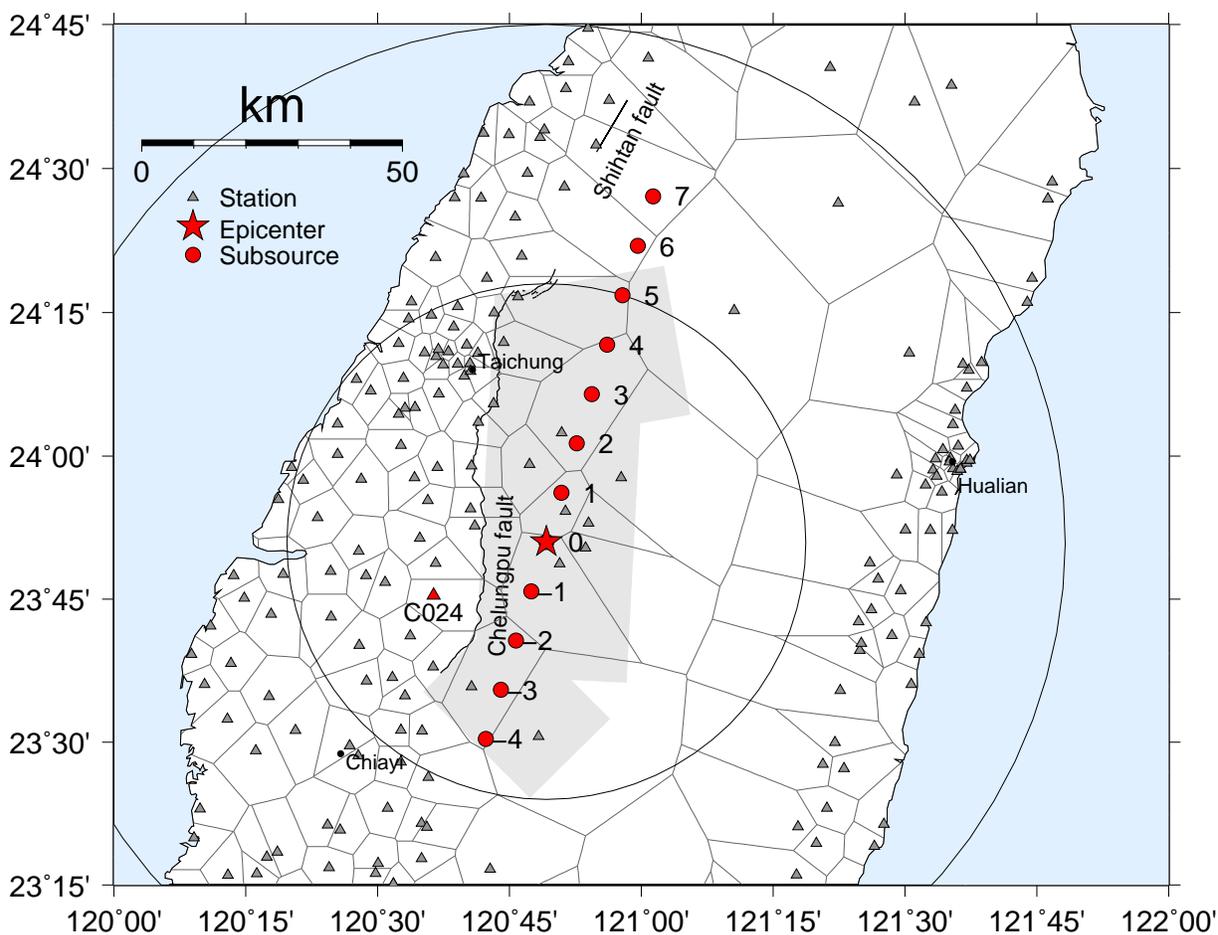| Thickness(km)	V_p(km/s)	V_s(km/s)
1.0	3.50	2.00
3.0	3.78	2.20
5.0	5.04	3.03
4.0	5.71	3.26
4.0	6.05	3.47
8.0	6.44	3.72
5.0	6.83	3.99
Half-space	7.28	4.21

Table 1: P-wave and S-wave velocity model in central Taiwan (Ma *et al.*, 1996).

Figure 1: The fault geometry and the station distribution of the Chi-Chi earthquake. The shaded area around the epicenter displays the map projection of the fault geometry proposed by Ji *et al.* (2003). Small circles indicate the location of subsources determined from the simulation in this paper. The area within 50 km and 100 km from the epicenter is shown by large circles. Stations used in this analysis are shown by solid triangles. The polygon surrounding each station is the Voronoi cell for the station.
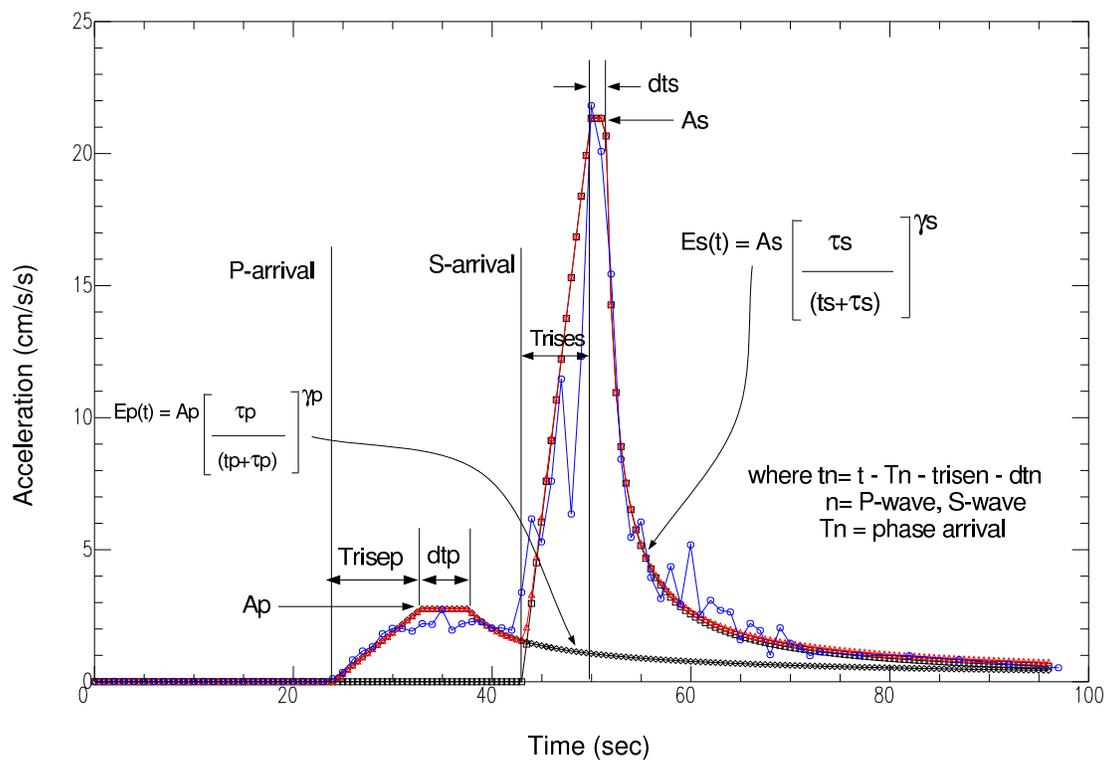
Figure 2: Observed envelope for accelerogram and P-wave and S-wave envelopes for the ground motion model defined in equation (5) (Cua, 2005).
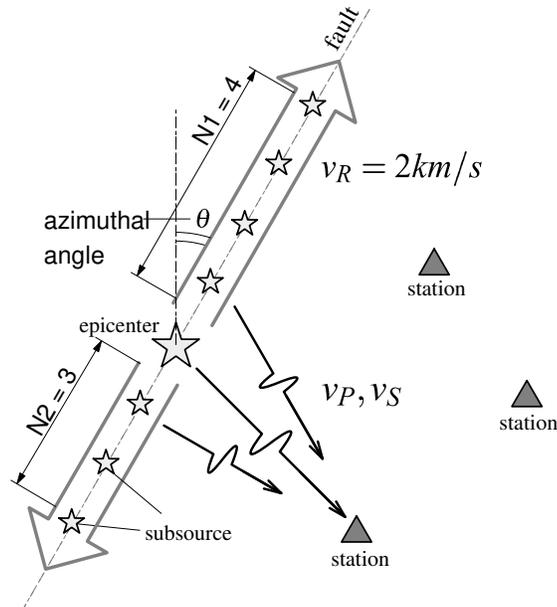
Figure 3: Schematic diagram of the multiple source model. The fault rupture is assumed to propagate from the epicenter at the constant velocity v_R. The fault is parameterized by θ, N1, and N2, where θ is the azimuthal angle of the fault, N1 and N2 are the number of subsources north and south of the epicenter, respectively. The ground motion at a station is expressed as a combination of the envelope from each subsource.
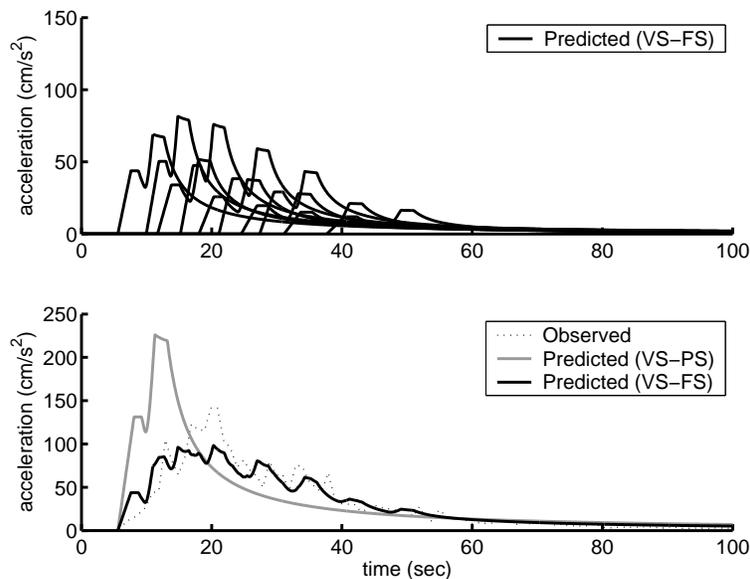
Figure 4: Envelopes of vertical acceleration recorded at the station C024 for the Chi-Chi earthquake. Top: predicted envelopes of the vertical acceleration record for each subsource with magnitude 6.0. Bottom: Observed envelope (in dotted black line), and predicted envelopes of the point source model in VS-PS method (in solid gray line) and of the multiple source model in VS-FS method (in solid black line).
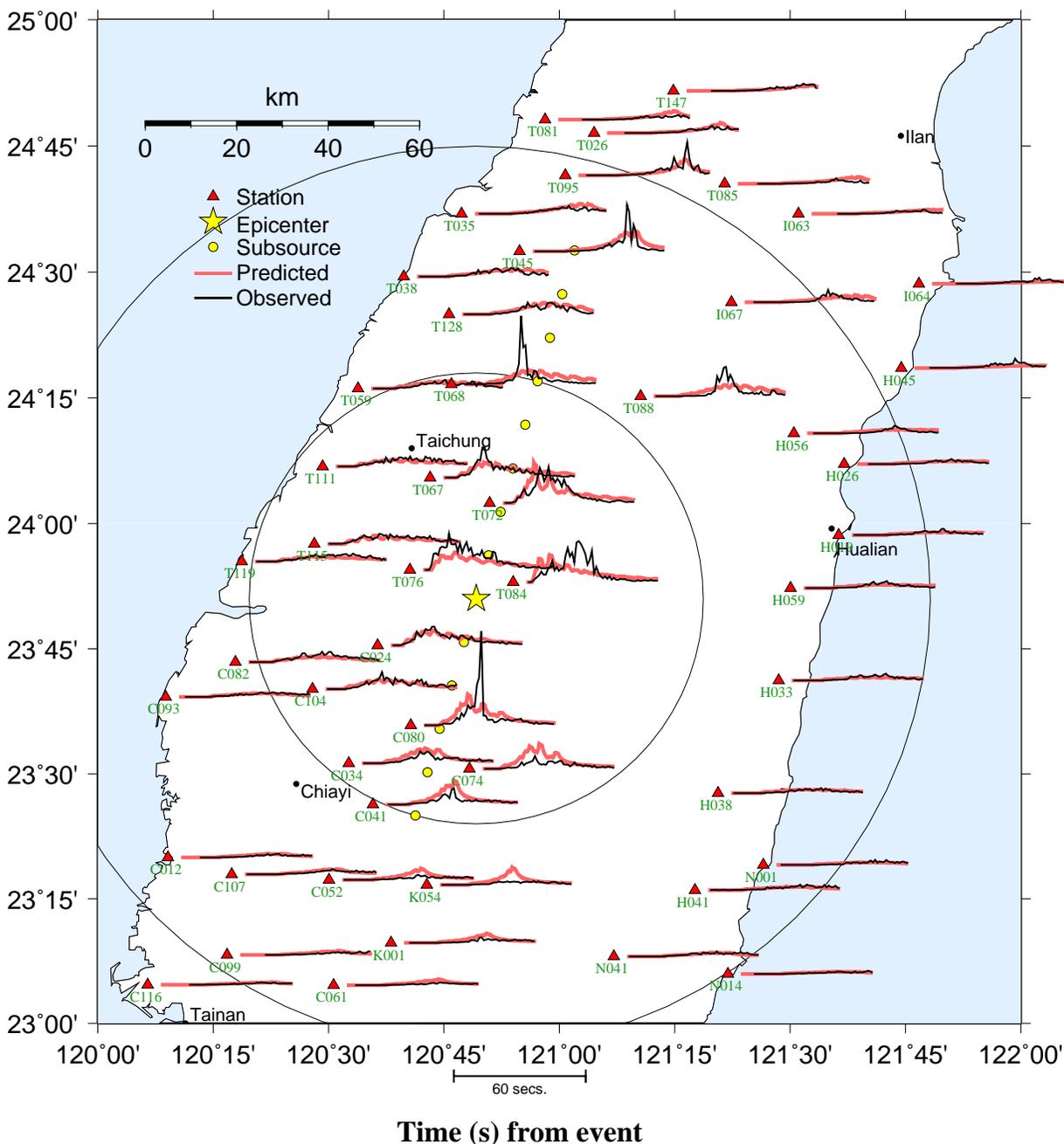
Figure 5: Predicted and observed envelopes in horizontal component. The red and black lines are the predicted and observed envelopes, respectively. The locations of the subsources estimated from the best fit model are shown in a small yellow circles. The area within 50 km and 100 km from the epicenter are shown by large circles. Only characteristic records of the stations are shown in this figure.

Figure 6: Predicted and observed envelopes in vertical component. The red and black lines are the predicted and observed envelopes, respectively. The locations of the subsources estimated from the best fit model are shown in a small yellow circles. The area within 50 km and 100 km from the epicenter are shown by large circles. Only characteristic records of the stations are shown in this figure.
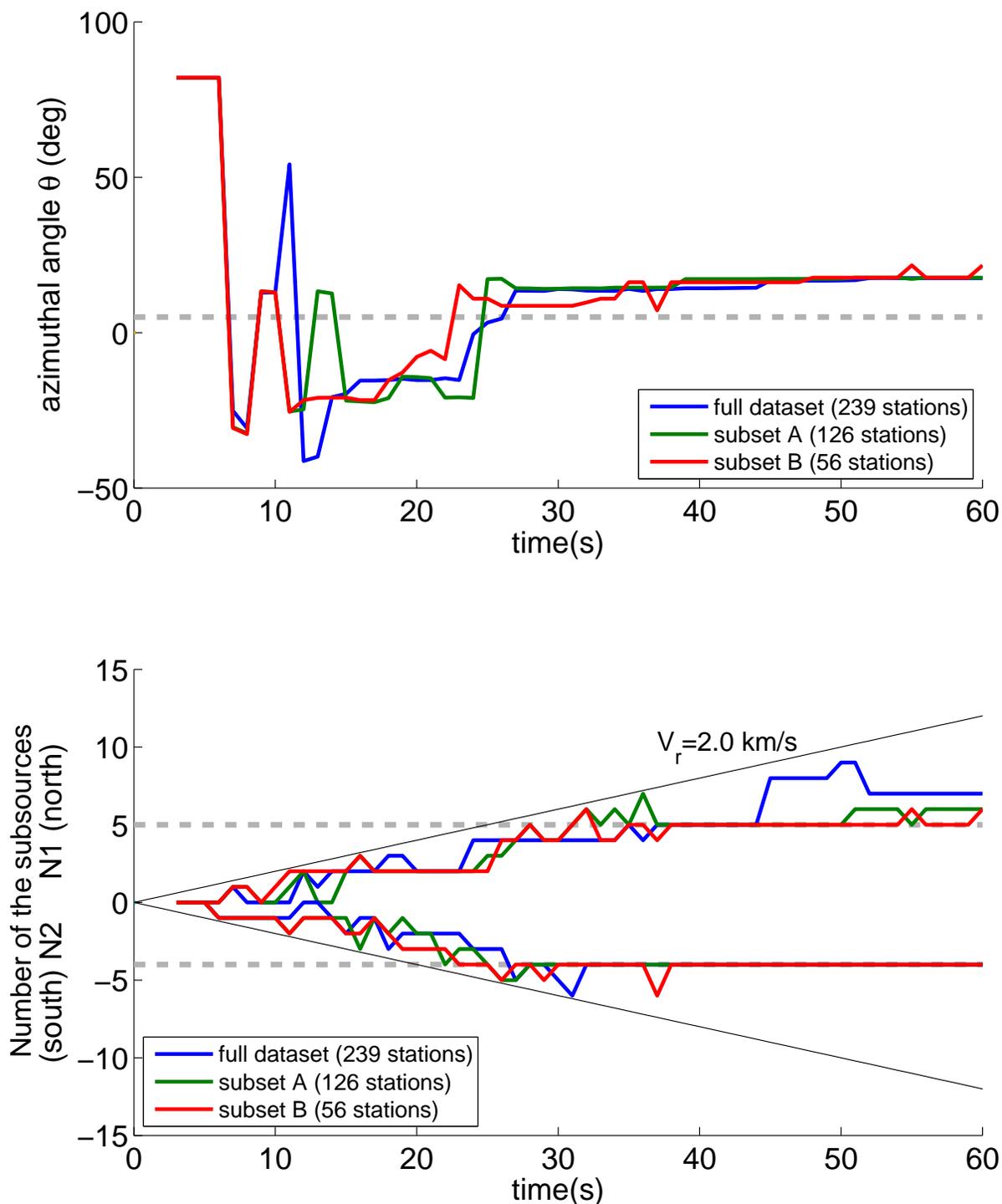
Figure 7: Time series of the estimated parameters, θ, N1, and N2 for each dataset. Time is relative to the origin. The parameters are computed at each second using only the data available at that time. The broken lines are the best estimates based on the fault model proposed by Ji *et al.* (2003). Top: time series estimations for θ. Bottom: time series estimations for for N1 and N2. The solid thin lines are the upper limits for N1 and N2 for the rupture velocity 2 km/s.
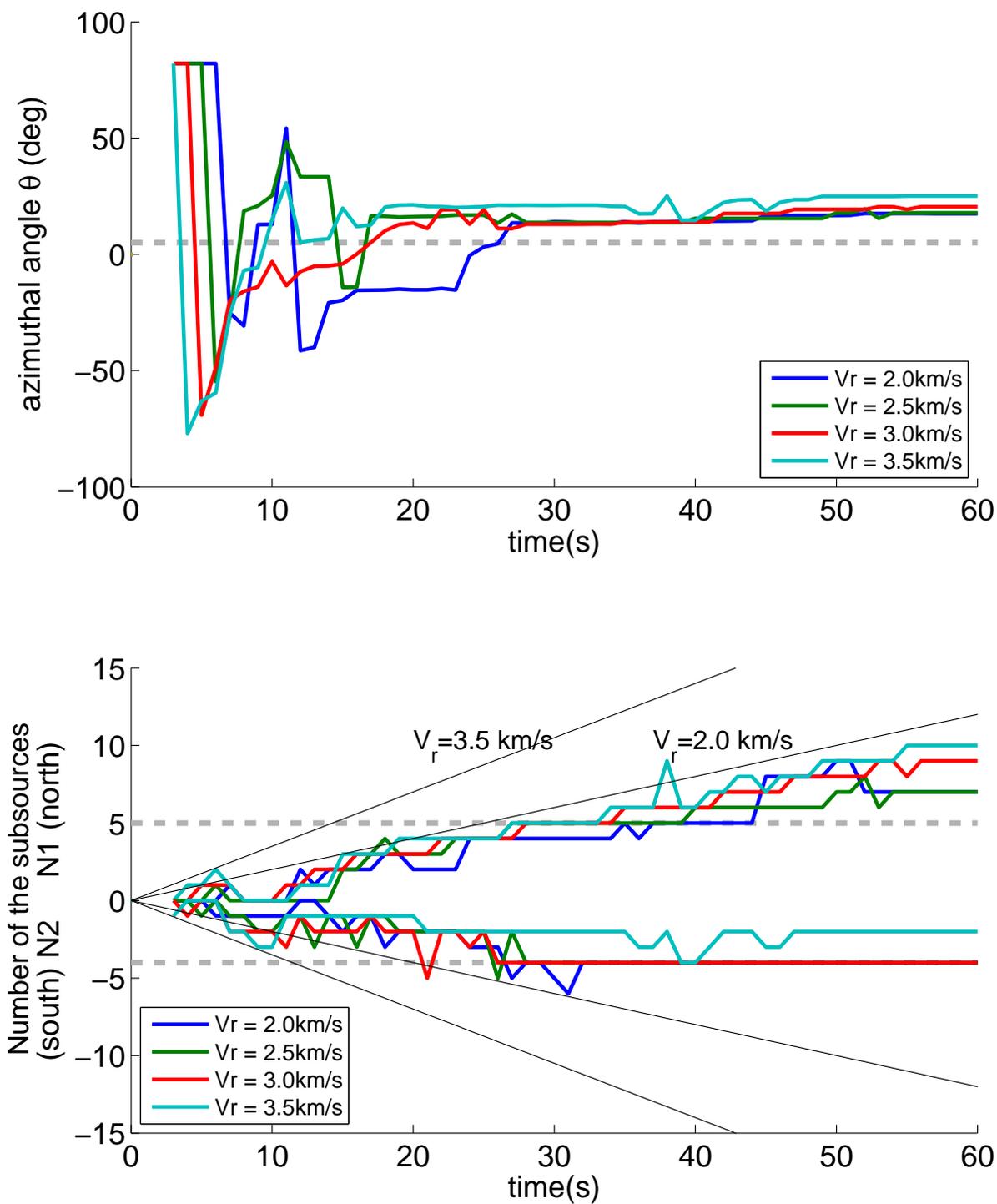
Figure 8: The estimated parameters, N1, and N2 for different rupture velocities. The solid thin lines are the upper limits for N1 and N2 for the rupture velocity 2 km/s and 3.5 km/s. The broken lines are the best estimates based on the fault model proposed by Ji *et al.* (2003). Time is relative to the origin. The parameters are computed at each second using only the data available at that time.
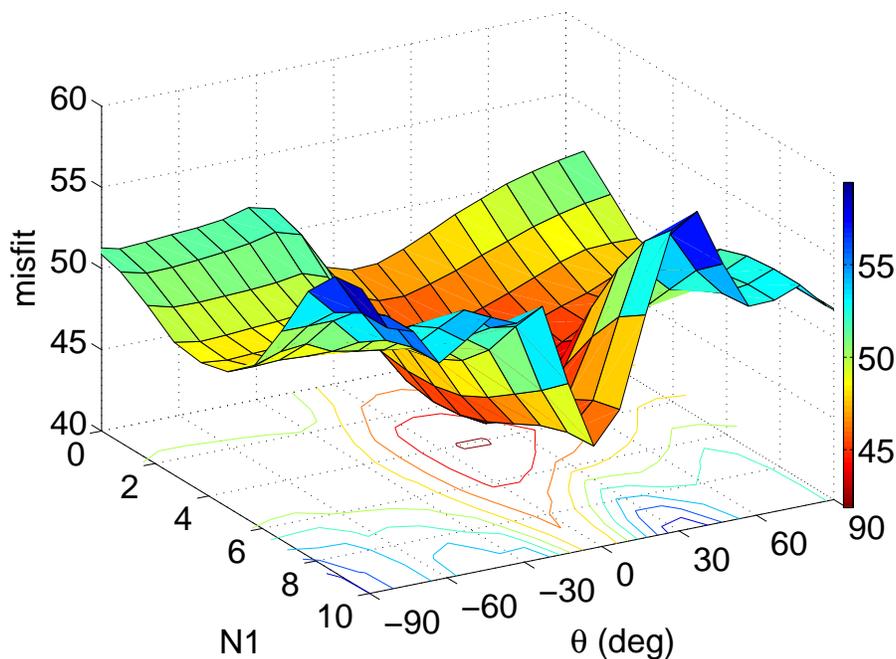
Figure 9: Error surface of θ and N1 at the fixed N2 = 4 at 60 seconds after the origin time. Since the surface is peaked around $\theta = 0$, it is easy to converge in θ. However, the optimal N1 will change easily depending on the misfit function (see equation (3)).
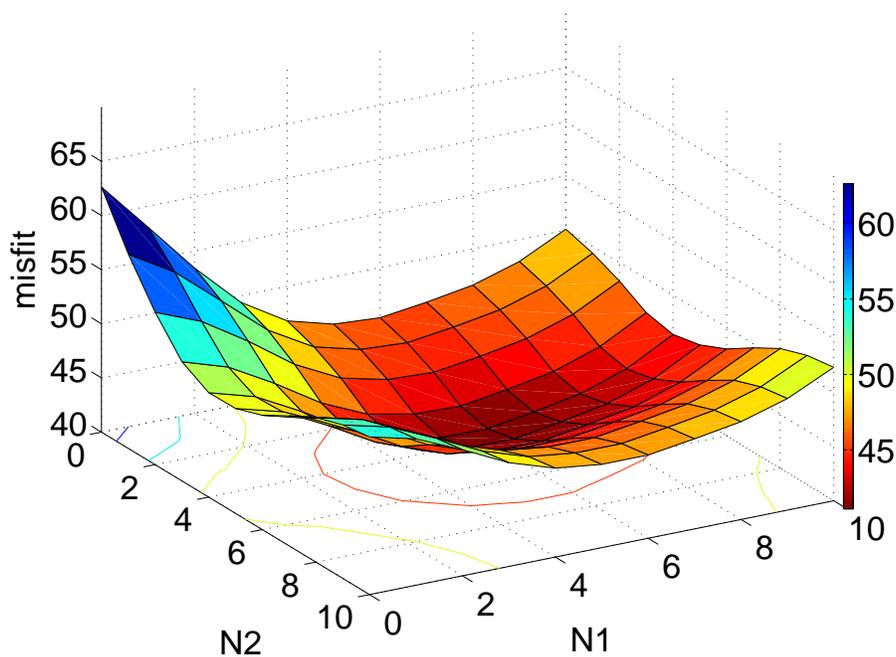

Figure 10: Error surface of N1 and N2 at the fixed $\theta = 17$ at 60 seconds after the origin time. Since the surface is smooth in both N1 and N2 direction, the optimal solution is sensitive to a small disturbance.
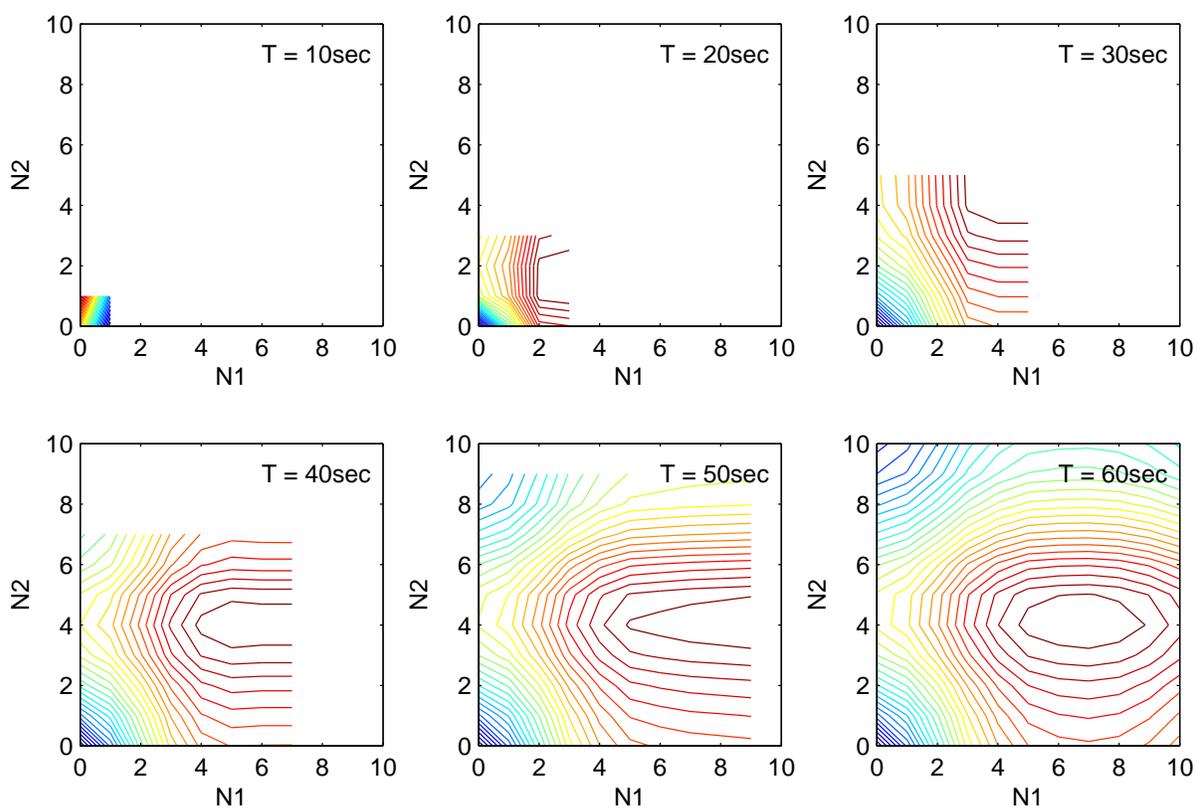
Figure 11: Contour maps of the error surface of N1 and N2 at the fixed $\theta = 17$. The maps are shown in 10 second intervals. The blank area in the boxes is the region where there is no solution due to the constraint that the rupture velocity is less than 2 km/s.