

Application of Seismic Array Processing to Earthquake Early Warning

L. Meng^{1,2*}, R. M. Allen² and J-P. Ampuero³

¹Department of Earth, Planetary and Space Sciences, UCLA. meng@ess.ucla.edu

²Berkeley Seismological Laboratory, U.C. Berkeley.

³Seismological Laboratory, California Institute of Technology.

Abstract

Earthquake early warning (EEW) systems that issue warnings prior to the arrival of strong ground shaking are essential in mitigating earthquake hazard. Currently operating EEW systems work on the assumption of a point source and are of limited effectiveness for large events, for which ignoring finite source effects results in magnitude underestimation. Here, we explore the concept of characterizing rupture dimensions in real time for EEW using small-aperture seismic arrays located near active faults. Back tracing the array waveforms allows the estimation of rupture size and directivity in real-time, which improves EEW capabilities for $M > 7$ earthquakes. We demonstrate the concept by implementing it in a simulated real-time environment and analyzing the 2004 Parkfield earthquake recordings by the UPSAR array and the 2010 El-Mayor Cucapah earthquake recordings by strong motion sensors in San Diego. We find it important to correct for the bias in back-azimuth induced by dipping structures beneath the UPSAR array, based on data from smaller events. Our real-time estimates of rupture length are 30% shorter than the length inferred from other studies, but still reasonable for EEW purpose. We attribute this difference to rupture directivity effects and to the limited field of view

of a single array. The accuracy of the approach may be improved with a network of arrays with overlapping fields of view. We demonstrate this by tracking the 2011 Tohoku earthquake rupture with two clusters of Hi-net stations in Kyushu and Northern Hokkaido. The obtained results are consistent with teleseismic back-projection results and yield reasonable estimates of rupture length and directivity. Compared to other proposed finite-fault EEW approaches, the array method is less affected by the coarseness of a GPS or seismic network and provides a high frequency characterization of the rupture which may yield a more suitable predictor of ground shaking for certain structures.

Introduction

Earthquake early warning (EEW) systems are a new mechanism for mitigation of seismic hazard based on issuing warnings prior to the arrival of strong ground shaking during an earthquake. All currently operating EEW systems treat earthquakes as a point source and use a magnitude-amplitude/frequency scaling. This approach works well for moderate to large earthquakes, but for the largest earthquakes ($M > \sim 7.5$) this can lead to a significant underestimate of the ground shaking due to an underestimation of the magnitude and the absence of information about the location and extent of the fault that has ruptured. This issue was highlighted by the EEW performance during the M9.0 Tohoku-Oki earthquake (*Hoshiba et al.*, 2011). Using only P-wave amplitude information, the magnitude estimated by the Japanese EEW system saturated at M8.1. Also, failing to take into account the finite rupture propagation, the system issued warnings of strong seismic intensity only for the Tohoku region. However, the Kanto region experienced much larger ground motions than those predicted. The example of the Tohoku-Oki earthquake demonstrates the need of characterizing the finite dimension in real time for the EEW of large earthquakes.

Among the ongoing efforts to determine the finite fault extent in real time, GPS approaches provide more reliable static displacements, and thus a static slip model, than seismic methods (Hudnut *et al.*, 2002; Allen and Ziv, 2011; Colombelli *et al.*, 2013). The FinDer approach is also proposed to determine linear fault geometry based on the difference in near/far field seismic waveforms, provided dense station coverage (Böse *et al.*, 2012).

Here, we explore the concept of imaging the rupture process of large earthquakes in real-time using clusters of dense seismic arrays located near an active fault (Meng and Ampuero, 2013). The arrays can be deployed in a zigzag fashion along the fault with each array covering a “effective field of view”. Back tracing the waveforms of an earthquake recorded by such arrays allows the estimation of the rupture directivity, size, duration, speed, and segmentation. The approach is now commonly used to image earthquakes recorded at teleseismic distance (Ishii *et al.*, 2005; Meng *et al.*, 2011) and regional distance (Vallée *et al.*, 2008; Meng *et al.*, 2012b), with some work on using it at local distances (Fletcher *et al.*, 2006; Allmann and Shearer, 2007; Honda and Aoi, 2009). Here, we develop key aspects of the implementation of array processing at local distance in real-time to estimate the rupture dimension of $M > 6$ earthquakes. The principle is analogous to the location and tracking of moving sources by antennas in a variety of military and civilian applications. Strong high-frequency (HF) seismic waves are typically assumed to radiate from the rupture front. Tracking the source of the HF seismic waves during large earthquakes recovers the migration of the rupture front. The trajectory of the rupture front marks the fault extent involved in the earthquake (Fig. 1). More generally, the approach provides the extent of the source area of HF radiation.

Method

The direction of arrival of waves propagating across an array can be determined by analyzing the phase of coherent seismic signals with array-processing techniques. Here, we adopt the correlation stacking technique, in which the normalized cross-correlation coefficients of all waveform pairs are beamformed (delayed and summed) instead of the waveforms themselves. This improves robustness against scattering and multi-pathing in the shallow crustal environment (*Borcea et al.*, 2002; *Fletcher et al.*, 2006). For a selected time window, the correlation coefficient cc_{ij} between stations i and j is defined by:

$$cc_{ij}(\theta) = \frac{\sum_t x_i(t)x_j(t - \tau_{ij}(\theta))}{(\sum_t x_i^2 \sum_t x_j^2)^{1/2}}$$

where x_i is the seismic signal recorded by the i -th station, t is time index, $\tau_{ij}(\theta)$ is the time delay of the wave arrival between the station i and j . The sums in the equation above encompass the duration of a running time window. We assume that the array aperture is much smaller than the epicentral distance, and approximate the impinging wave front as a plane. The delay as a function of the slowness vector θ of an impinging plane wave is given by $\tau_{ij}(\theta) = \theta \cdot (r_i - r_j)$, where r_i and r_j are the location vector of the i -th and j -th station. For fast implementation $\tau_{ij}(\theta)$ is rounded to the nearest discrete sample. We compared this approach to a more time-consuming implementation based on sub-sample spectral time-shifting and verified that it does not introduce notable errors for signals with high sampling rate in the frequency band that we are interested in.

The correlation coefficients are then summed over all station pairs to obtain the stacked correlation coefficient, $scc(\theta)$, as a function of the slowness vector of the candidate plane wave. The slowness vector contains information on both the source depth (apparent velocity) and its

lateral location (back azimuth). For EEW purposes, we are primarily interested in the lateral rupture dimension and focus on the use of the back azimuth. For a given back azimuth, we define the projected lateral location as the intersection of the great circle path and a candidate fault surface (possibly non-planar) extracted from a predefined fault catalog (e.g. the SCEC Community Fault Model). The epicenter location is provided in a real-time fashion by the conventional point source based EEW system, e.g. ElarmS (*Kuyuk et al.*, 2013). If the epicenter is outside the known fault zones, the fault surface can be approximated as a linear plane with the strike determined by interpolation of nearby faults.

The stacked correlation computed on the running windows is then analyzed as a function of the distance along the fault strike and time. At each time, the location of the peak stacked correlation is identified. If the peak values exceed a prescribed threshold (here, 3 times the average stacked correlation of the background noise), they are characterized as significant seismic radiators. In earthquake source imaging studies (*Fletcher et al.*, 2006; *Vallée et al.*, 2008; *Meng et al.*, 2012a), the trajectory of the seismic radiators in the distance-time domain is used to infer the rupture dimension and rupture speed. However, this analysis involves human intervention to identify the rupture trend and branching patterns, which is not readily applicable in real-time implementations. We thus propose a simple algorithm to characterize the rupture dimension and directivity based on the location information. We define the rupture length as the distance between the leftmost and rightmost seismic radiators along the fault plane. The location of these two end points with respect to the hypocenter provides the rupture directivity and type (unilateral or bilateral). The information on rupture speed is not required for EEW.

The M6.0 2004 Parkfield earthquake

We applied the proposed methodology to the 2004 M6 Parkfield earthquake in a simulated real-time environment. The earthquake is one of the few big events ($M > 6$) that are recorded by a local small-scale seismic array (UPSAR array, see Data and Resources Section). The array is composed of 13 stations with an aperture of about 1 km and spacing of roughly 200 m and is located at about 12 km from the San Andreas Fault. Fletcher et al., 2006 showed that the rupture kinematics can be retrieved from the UPSAR array data.

Here, before we examine the main event, we perform array analysis on the aftershocks. Figure 2a compares the hypocenters with the locations inferred from the back azimuth retrieved by the UPSAR array. The results show that the locations are biased towards the “center” of the rupture zone. Figure 2b further shows a linear dependence of the back-azimuth residue (difference between observed and theoretical) with respect to the theoretical azimuth. Fletcher et al., 2006 propose that this bias is caused by ray path bending due to the lateral velocity variation associated with the fault-damaged zone. However, this hypothesis requires a decreasing velocity gradient away from the fault, which is the opposite of what is observed in natural fault zones.

In the nuclear monitoring community, it has been long known that a dipping velocity structure beneath a seismic array can affect the incoming teleseismic wavefield and introduce deviations of azimuth and slowness from their expected values in a horizontally layered medium (Fig. 2b) (Niazi, 1966; Otsuka, 1966a, 1966b). In that context, the dipping structures are typically interpreted to be the Moho discontinuity (Greenfield and Sheppard, 1969). This dipping layer effect has not been addressed in the context of array analysis at local distance. Here, we propose that the azimuth bias at the UPSAR array is caused by the dipping interface between the shallow sediments and bedrock. Theory suggests a sinusoidal dependence of the back-azimuth residual as a function of the true back-azimuth across a 2π range (Bondar et al., 1999;

150 *Flanagan et al.*, 2012). This bias pattern can be used to constrain the orientation of the dipping
151 structure: the azimuth at which the residue crosses zero from negative to positive corresponds
152 to the dipping direction of the interface (*Lindquist et al.*, 2007).

153
154 Since our analysis involves only the seismicity in the Parkfield section of the San Andreas fault,
155 the azimuth coverage is limited to about ~ 90 degrees (Fig. 2b). Nevertheless, the back-azimuth
156 anomaly crosses zero from positive to negative at 43 degree (normal to the strike of the San
157 Andreas fault), suggesting that the interface dips away (223 degree from N) from the San
158 Andreas Fault. This is consistent with the topography of the UPSAR array (Fig. 2a). The dipping
159 layer is likely caused by the fault-parallel folding deformation associated with the long-term
160 deformation of the San Andreas fault (*Mount and Suppe*, 1987). The magnitude of the azimuth
161 bias depends on the angle of incidence, velocity contrast across the dipping interface and its dip
162 angle. The incidence angles of individual aftershocks are calculated based on a 1-D velocity
163 model in the Parkfield area (*Roux et al.*, 2005). The dip of the interface can be calculated by
164 Snell's law given the velocity across the interface (*Flanagan et al.*, 2012). The velocity contrast
165 between the sediment and bedrock is too shallow to be resolved by tomography. Here we
166 assume a ratio of 1:2 between the shear wave speed of the shallow sediments layer and the
167 basement rock. The inferred dip is 22 degree, consistent with that of the top layers imaged by
168 the tomography studies (e.g. *Thurber et al.*, 2006). This dip angle trades off with the assumed
169 velocity contrast across the dipping interface. The true angle might be larger for a smaller shear
170 wave speed ratio. While we cannot resolve this trade-off, for the purpose of correcting the back-
171 azimuth we do not need to.

172
173 Once the dipping structure is determined, we can apply a model-based azimuth correction to the
174 array analysis of the Parkfield mainshock, rather than the empirical correction derived by
175 Fletcher et al., 2006 based on fitting an ad hoc smooth function to the azimuth bias. Our model

based correction enables extrapolation beyond the azimuth range represented in the empirical dataset and, since the sinusoidal shape of the calibration function is established, fewer aftershocks are required to obtain reliable corrections.

A station-specific delay correction is also applied to take into account the local S-wave velocity variation across the array following Fletcher et al., 2006. In practice, these corrections can be obtained by averaging the travel time residues with respect to the best-fitting plane wave from several aftershocks. We find that this correction has a minor effect. The resulting difference in back-azimuth is much smaller than the dominant dipping layer effect.

We applied the correlation stacking to the S waves recorded during the 2004 Parkfield main event in the frequency band of 0.5 - 8 Hz with a sliding window of 1 s. The S-minus-P travel times at such close distance is shorter than the rupture duration, which prevents us from exploiting P waves to track the whole rupture. The dominant frequency of the ground acceleration signals is higher than 1 Hz. The bandwidth is narrower and the window is longer compared with Fletcher et al., 2006 to obtain a more robust image for EEW purposes. The unilateral rupture pattern is well constrained as the leftmost point is close to the hypocenter and the rightmost point is much further to the right. Figure 3 compares the source imaging of the 2004 Parkfield earthquake with and without dipping layer correction. Ignoring the dipping structure leads to a rupture length of only 5 km (Fig. 3a), significantly shorter than the typical rupture length of a M6 earthquake. Once corrected, the rupture length is 9 km (Fig. 3b), consistent with the estimate by Fletcher et al., 2006 based on the same dataset. Both estimates are shorter than the 13 km distance between the hypocenter and the northernmost subevents identified by back-projection using all local stations (*Allmann and Shearer, 2007*). One reason for this difference is the directivity effect. The unilateral northwestward rupture propagates towards the array initially, but away from the array once it is passed the latitude 35°55'. This

directivity effect enhances the seismic radiation of the southern rupture and weakens the northern rupture recorded at UPSAR array. The northernmost part of the rupture thus appears too weak to be imaged by UPSAR, which results in an underestimation of the rupture length. In this case, a single array is not sufficient to monitor the entire rupture area. A network of multiple small-scale arrays with partially overlapping fields of view would therefore be required to balance off the directivity effect.

The 2010 M7.2 El Mayor-Cucapah earthquake

As the largest recent earthquake in Baja California, the 2010 El Mayor-Cupacah earthquake is another example that can be used to demonstrate the application of array processing for EEW. It was well recorded by a group of 13 strong motion stations in San Diego, with an aperture of about 30 km (see Data and Resources Section). The array is at more than 150 km from the rupture and outside the area of strong motion, but we can use this example as an additional test and illustration of the general approach proposed here to estimate earthquake rupture dimension. We process the S wave data in the frequency range of 0.2 to 1 Hz with sliding windows of 10 s. Since the array is not optimally oriented with respect to the fault trace, i.e. it is not perpendicular to the fault, the resolution of the array is limited. The rupture was more complicated than the 2004 Parkfield event, rupturing bilaterally with a complex back-and-forth time history. Using our array techniques in simulated real-time, a general bilateral rupture trend can be identified (Fig. 4), although the seismic radiators associated to the two rupture fronts appear intermittently. The length of the northern branch is 60 km, consistent with finite fault models (e.g. *Wei et al.*, 2011). However, the southern rupture appears as only 20 km long, significantly shorter than that of the finite fault models. This is again due to the directivity effect: the southeastward rupture appears weak because the rupture moves away from San Diego. In addition, the resolution of the southern part is also lower as the great circle path of the array is

even more oblique to the fault strike than the northern rupture. This example again demonstrates the value of a network of multiple arrays along the length of active faults for real-time application of this method. Overall, the 80 km rupture length is shorter than the 120 km rupture length estimated by Wei et al., 2011 but is still a reasonable estimate for EEW purpose, and better than treating the event as a point source.

The 2011 M9.0 Tohoku Earthquake

The previous two examples of application of array processing for EEW demonstrate the need of multiple arrays. Very few large earthquakes have been recorded by multiple small-scale arrays in close distance. Nevertheless, we can upscale the problem and proof the concept with a megathrust earthquake recorded by a regular seismic network. We consider the 2011 M9.0 Tohoku earthquake, recorded by the dense Hi-net seismic network over the entire Japan.

We apply the array analysis to two clusters of Hi-net stations, one in Kyushu and one in northern Hokkaido (Fig. 5-left). The arrays are at about 1000 km and 700 km, respectively, from the hypocenter. The inter-station spacing is approximately 20 km. This setting is scaled up by roughly 100 times compared to the UPSAR / 2004 Parkfield earthquake example (average spacing of 200 m and epicentral distance of 12 km). The data is processed in the frequency band from 0.1 to 0.25 Hz. This frequency range is less than 100 times lower than that adopted in the Parkfield example. We chose this relatively high frequency band in order to compare the results with the teleseismic array back-projection results. Analysis at a lower frequency band reproduces the large-scale features of the rupture consistently. We use the Pn phase, which carries the highest frequency content and proves to be well suited for earthquake source imaging at regional distance (Meng et al., 2012b). Although the Tohoku rupture has been imaged by array processing with individual arrays and networks at local and regional scale

(Honda *et al.*, 2011; Nakahara *et al.*, 2011; Maercklin *et al.*, 2012; Roten *et al.*, 2012), we provide the first effort to jointly image this earthquake with multiple arrays.

We performed the array analysis on recordings of the two arrays separately, and then at each epoch we intersected the directions defined by the two back-azimuths to determine the location of the seismic radiation source. The significance is defined as the sum of both correlation coefficients. We apply a static correction to the back azimuths so that the initial source is located at the hypocenter. This correction is approximately 2 degrees for the Hokkaido array and 10 degrees for the Kyushu array. At this scale this azimuth discrepancy is more likely due to lateral velocity variations rather than topography or dipping structures. The larger bias observed at the Kyushu array may reflect the lateral velocity contrast between the island arc and back-arc basin.

In general, the joint array analysis recovers the rupture process of the Tohoku earthquake reasonably well (Fig.5-right). Both the spatial extent and rupture history are consistent with the results of teleseismic back-projection (e.g. Ishii, 2011; Koper *et al.*, 2011; Meng *et al.*, 2011; Yao *et al.*, 2011; Zhang *et al.*, 2011). The along-strike rupture length of 300 km and the southward directivity towards the Kanto area are captured adequately for EEW purposes. This example of the Tohoku earthquake illustrates how multiple seismic arrays enable a better characterization of the rupture length and directivity.

Discussion

The array-based EEW approach can be compared with other methods characterizing the finite dimensions of earthquake rupture in real time. Based on the measurement of the static field, the GPS-based approaches (*Allen and Ziv, 2011; Colombelli et al., 2013*) can potentially provide valid slip estimate upon the arrival of S-waves. However, with limited spatial resolution and the smoothing regularization applied in the inversion, the boundary of the actual slip area can be only roughly defined. An additional approach makes use of the “Zipper” GPS network composed of pairs of GPS stations, one on each side of the fault, along the length of the southern San Andreas Fault. It provides a direct measure of surface slip, but requires GPS stations along every fault to be monitored (*Hudnut et al., 2002*). Similarly, by discriminating near field from far field stations, the Finder approach (*Böse et al., 2012*) estimates the “perimeter” of the rupture zone. The boundary may again be smeared unless extremely dense station coverage is available. The array approach on the other hand, provides a sharper definition of the rupture boundary achieved by exploiting the high frequency content of the seismic wavefield. In the 2004 Parkfield example, we processed the seismograms up to 8 Hz (wavelength of 37 m, assuming a shear wave speed of 300 m/s at shallow depth), implying a much sharper spatial and temporal resolution.

The view of the rupture obtained from high frequency array analysis is not necessarily correlated to the low frequency aspects of the rupture process. In general, the array sees the last strong HF sub-event that stands out of the coda wave of the previous HF sub-events. In other words, the array sees only the strongly radiating parts of the rupture, but cannot resolve a smooth rupture arrest. This partially explains why the rupture length of the Parkfield imaged by the high-frequency array studies (*Fletcher et al., 2006; Allmann and Shearer, 2007; this study*) is shorter than the rupture length estimated by finite fault studies (*Liu and Archuleta, 2006*) and the aftershock zone (*Waldhauser et al., 2004*). On the other hand, if the rupture terminated abruptly, with a strong stopping phase, the high frequency rupture size might be as large as or

even larger than the low frequency size. This is the situation of the 2010 Haiti earthquake where the high frequency slip ruptured beyond the terminal edge of static slip area (Meng *et al.*, 2012b). In either of these cases, the high frequency rupture process imaged by small-aperture arrays is more directly relevant for the prediction of high frequency ground motion. Wei *et al.*, 2012 show that most of the ground motion of the Tohoku earthquake can be explained by only 18% of total slip located in the deep portion of the megathrust close to the Honshu Island. This deep slip is imaged by the array analysis at high frequency and its location delineates the bottom of the long period slip region. In the example of the catastrophic Haiti earthquake, the westward high frequency rupture is closer to Port-au-Prince (Meng *et al.*, 2012) than the low frequency slip, which might have contributed to the unusually high ground motion experienced during this event (Hough *et al.*, 2010). Thus, this HF array approach can potentially become an essential component of an EEW system, as it can constrain estimates of ground motion at the high frequencies which matter to residential areas with low story buildings (e.g. the Los Angeles area) as well as critical infrastructure and lifelines (O'Rourke *et al.*, 2012). For this purpose, an empirical ground motion prediction approach can be developed based on the distance and power of the HF subevents identified by array analysis.

In terms of the practicality of a real-time implementation, the array analysis is an imaging procedure instead of an inversion, it does not suffer from ill-posedness and underdeterminacy. It also requires fewer assumptions about the velocity model and rupture kinematics and is therefore more suitable to be implemented in real-time. In the other existing network-based finite fault EEW approaches (GPS and Finder), information about earthquake magnitude and slip is only available following the S wave arrival at network stations. In contrast, the array approach can provide a continuous update of rupture size as the earthquake unfolds, potentially saving seconds.

The calibration of the anomalous azimuth due to dipping layer beneath the array is important for the array processing based EEW approach. The dipping structure under the UPSAR array may not be unique since fault-parallel folding is common near large strike-slip faults (*Mount and Suppe, 1987*). Since we propose to deploy small-scale arrays close to active faults, it is likely that these corrections are needed for all future deployments. For the purpose of the correction, a temporary high sensitivity array is needed to record the calibration events (small local events or teleseismic events) prior to the deployment of the permanent strong motion array. For regions with low seismicity, the model-based dipping-layer correction is particularly advantageous, as it requires less calibration events than the empirical correction.

The EEW system based on the array approach still faces challenges. Particularly, a fault strike needs to be assumed based on a fault map, which is problematic if the earthquake occurs on previously unrecognized fault, although the chance of a large earthquake ($M > 7$) on an unknown fault system is low. The assumption of the fault strike might be released by combining the processing of multiple arrays and triangulating the rupture location, as demonstrated here by the Tohoku earthquake example. While site-specific back-azimuth bias are usually calibrated with aftershock data in offline processing (*Honda et al., 2008*), for an EEW implementation a first-order static correction can be achieved by forcing the initial back-azimuth estimates to match the hypocenter location. This procedure produces reasonable results in the Tohoku example. The examples of the Parkfield and El-Mayor Cucapah earthquakes shown here demonstrate the incompleteness of the source imaging by a single array due to the rupture directivity effect and limited effective field of view. This results in underestimation of the rupture length. A system of multiple arrays with partially overlapping field of view is a potential solution. The joint “fields of view” can cover a large region and allow averaging of the rupture length estimates within the overlapping target segments to improve robustness. The bias due to the directivity effect can be mitigated with arrays located at different orientation with respect to the rupture direction.

Another challenge of all finite-fault EEW methods, including the array-based EEW approach demonstrated here, is the size of the blind zone, defined as the region that cannot receive a warning before the arrival of S waves. The radius of the blind zone is predominantly dictated by the threshold of earthquake size at which a warning is issued. It takes about 10 s for an earthquake to grow up to M 6.5 (with a rupture length of 30 km, assuming a rupture speed of 3 km/s) and roughly 3 s for the S wave to arrive at an array 10 km away from the fault (assuming a S-wave speed of 3.5 km/s). These 13 s needed to declare a warning result in a blind zone of at least 45 km, significantly larger than that of the current CISM ShakeAlert EEWS in California (Kuyuk and Allen, 2013). We note that this blind time is dominated by rupture duration, rather than by wave travel time. Using P waves instead of S waves would not reduce it significantly (and is challenged by S minus P travel times shorter than rupture duration). Although the issue of the blind zone is common to all finite-fault EEW methods, these systems are not intended to provide the first warning, but to send updates of rupture growth after the ShakeAlert warnings.

There is still more development necessary before arrays can be deployed for the purpose of EEW. It will be important to develop optimal designs for the array geometry and position, identify the best processing parameters (e.g. frequency band and length of sliding window), and quantify the uncertainties. Validation by synthetic tests with realistic source and path effects will help the further development of the real-time implementation strategy. Still, the initial tests shown here constitute a proof of concept for a local-array-based approach to track earthquake rupture for the purpose of EEW.

Conclusion

We developed a methodology for real-time characterization of earthquake rupture length and directivity based on seismic array data. The strategy we propose can potentially prompt more reliable earthquake early warning for large earthquakes. We implemented the proposed methodology to analyze the mainshock in a simulated real-time environment. We demonstrate the concept on the 2004 Parkfield earthquake recordings by the UPSAR array. We find it important to correct for the bias in back-azimuth induced by dipping structures beneath the array. After calibrating the dipping-layer effect with data from smaller events, we obtained an estimated rupture length of 9 km. This is comparable, but a little shorter, than the distance of 13 km between the two main high frequency subevents identified by back-projection using all local stations. We attribute this difference in rupture length to the rupture directivity that affects the amplitude of high frequency radiation towards UPSAR. In another example, we processed the ground accelerations of the 2010 El-Mayor Cucapah earthquake recorded by a group of strong motion sensors in San Diego. We find the rupture length estimate is again perturbed by the different directivity effects of the two fronts of this bilateral rupture and limited by the field of the view of a single array. Our simulated real-time estimate of 80 km, a little less than the 120 km estimate from other studies, is still reasonable for EEW purpose. A network of arrays with overlapping fields of view can potentially improve the robustness and accuracy, which we demonstrated by jointly imaging the 2011 Tohoku earthquake with two clusters of Hi-net stations in Kyushu and Hokkaido: the recovered rupture process is consistent with teleseismic back-projection imaging results and the inferred rupture length of 300 km and southward directivity are satisfactory for EEW purposes. We compared the benefits of the array-based EEW approach with other real-time finite-fault EEW solutions based on GPS and seismic data. The array method constrains high frequency aspects of the rupture that are complementary to the current EEW system and a more suitable ground motion predictor in certain cases.

Data and Resources

408

409 The Center for Engineering Strong Motion Data (CESMD) is accessed to download the strong
410 motion data of the El Mayor-Cucapah earthquake. (<http://strongmotioncenter.org/>, last accessed
411 April 2013). The seismograms of the Parkfield earthquake recorded at UPSAR array is available
412 at the U.S. Geological Survey national strong-motion project (<http://nsmp.wr.usgs.gov/>, last
413 accessed February 2013). The broadband seismograms of the Tohoku earthquake is acquired
414 From the Japanese Hi-net (<http://www.hinet.bosai.go.jp>, last accessed February 2014).

415

416

417 **Acknowledgements**

418

419 This work was funded by the Gordon and Betty Moore Foundation through Grant GBMF3024 to
420 UC Berkeley. We thank Nathan Simmons for valuable discussions about the dipping layer effect.
421 We thank Paul Spudich and Joe Fletcher for providing the back azimuth estimates of the 2004
422 Parkfield earthquake and its aftershocks.

423

424

425

426 **References**

427 Allen, R. M., and A. Ziv (2011), Application of real-time GPS to earthquake early warning,
428 *Geophys. Res. Lett.*, 38(16), n/a–n/a, doi:10.1029/2011GL047947.

429 Allmann, B. P., and P. M. Shearer (2007), A high-frequency secondary event during the 2004
430 Parkfield earthquake., *Science*, 318(5854), 1279–83, doi:10.1126/science.1146537.

431 Bondar, I., R. North, and G. Beall (1999), Teleseismic Slowness-Azimuth Station Corrections for
432 the International Monitoring System Seismic Network, *Bull. Seismol. Soc. Am.*, (August),
433 989–1003.

434 Borcea, L., G. Papanicolaou, C. Tsogka, and J. Berryman (2002), Imaging and time reversal in
435 random media, *Inverse Probl.*, 18(5), 1247–1279.

436 Böse, M., T. H. Heaton, and E. Hauksson (2012), Real-time Finite Fault Rupture Detector
437 (FinDer) for large earthquakes, *Geophys. J. Int.*, 191(2), 803–812, doi:10.1111/j.1365-
438 246X.2012.05657.x.

439 Colombelli, S., R. M. Allen, and A. Zollo (2013), Application of real-time GPS to earthquake
440 early warning in subduction and strike-slip environments, *J. Geophys. Res. Solid Earth*,
441 118(May), n/a–n/a, doi:10.1002/jgrb.50242.

442 Flanagan, M., S. Myers, and N. Simmons (2012), Model-based corrections to observed back
443 azimuth and slowness observations from a dipping Mohorovicic discontinuity, in *Monitoring*
444 *Research Review, Albuquerque, NM, United States, Sep 18 - Sep 20, 2012*, vol. LLNL-
445 CONF-, p. LLNL–CONF–563592.

446 Fletcher, J. B., P. Spudich, and L. M. Baker (2006), Rupture Propagation of the 2004 Parkfield,
 447 California, Earthquake from Observations at the UPSAR, *Bull. Seismol. Soc. Am.*, 96(4B),
 448 S129–S142, doi:10.1785/0120050812.

449 Greenfield, R., and M. Sheppard (1969), The Moho depth variations under the LASA and their
 450 effect on $dT/d\Delta$ measurements, *Bull. Seismol. Soc. Am.*, 409–420.

451 Honda, R., and S. Aoi (2009), Array Back-Projection Imaging of the 2007 Niigataken Chuetsu-
 452 oki Earthquake Striking the World's Largest Nuclear Power Plant, *Bull. Seismol. Soc. Am.*,
 453 99(1), 141–147, doi:10.1785/0120080062.

454 Honda, R., S. Aoi, H. Sekiguchi, and H. Fujiwara (2008), Imaging an asperity of the 2003
 455 Tokachi-oki earthquake using a dense strong-motion seismograph network, *Geophys. J.*
 456 *Int.*, 172(3), 1104–1116, doi:10.1111/j.1365-246X.2007.03702.x.

457 Honda, R., Y. Yukutake, H. Ito, M. Harada, T. Aketagawa, and A. Yoshida (2011), A complex
 458 rupture image of the 2011 off the Pacific coast of Tohoku Earthquake revealed by the
 459 MeSO-net, *Earth, Planets Sp.*, 63(7), 583–588, doi:10.5047/eps.2011.05.034.

460 Hoshiba, M., K. Iwakiri, N. Hayashimoto, and T. Shimoyama (2011), Outline of the 2011 off the
 461 Pacific coast of Tohoku Earthquake (M w 9.0)—Earthquake Early Warning and observed
 462 seismic intensity—, *Earth, Planets Sp.*, 63(7), 547–551, doi:10.5047/eps.2011.05.031.

463 Hough, S. E., J. R. Altidor, D. Anglade, D. Given, M. G. Janvier, J. Z. Maharrey, M. Meremonte,
 464 B. S.-L. Mildor, C. Prepetit, and A. Yong (2010), Localized damage caused by topographic
 465 amplification during the 2010 M 7.0 Haiti earthquake, *Nat. Geosci.*, 3(11), 778–782,
 466 doi:10.1038/ngeo988.

467 Hudnut, K. W., G. J. Anderson, A. Aspiotes, N. E. King, R. A. Moffitt, and K. F. Stark (2002),
 468 GPS FAULT SLIP SENSORS, *Asia-Pacific Econ. Coop. Symp. Conffronting Urban*
 469 *Earthquakes - Seism. Early Warn.*, Nov.(28-29).

470 Ishii, M. (2011), High-frequency rupture properties of the M w 9.0 off the Pacific coast of Tohoku
 471 Earthquake, *Earth, Planets Sp.*, 63(7), 609–614, doi:10.5047/eps.2011.07.009.

472 Ishii, M., P. M. Shearer, H. Houston, and J. E. Vidale (2005), Extent, duration and speed of the
 473 2004 Sumatra-Andaman earthquake imaged by the Hi-Net array., *Nature*, 435(7044), 933–
 474 6, doi:10.1038/nature03675.

475 Koper, K. D., A. R. Hutko, T. Lay, C. J. Ammon, and H. Kanamori (2011), Frequency-dependent
 476 rupture process of the 2011 M w 9.0 Tohoku Earthquake: Comparison of short-period P
 477 wave backprojection images and broadband seismic rupture models, *Earth, Planets Sp.*,
 478 63(7), 599–602, doi:10.5047/eps.2011.05.026.

479 Kuyuk, H. S., and R. M. Allen (2013), Optimal Seismic Network Density for Earthquake Early
 480 Warning: A Case Study from California, *Seismol. Res. Lett.*, 84(6), 946–954,
 481 doi:10.1785/0220130043.

482 Kuyuk, S., R. M. Allen, H. Brown, M. Hellweg, I. Henson, and D. Neuhauser (2013), Designing a
 483 network-based earthquake early warning system algorithm for California: ElarmS-2, *Bull.*
 484 *Seismol. Soc. Am.*, *In press*.

485 Lindquist, K. G., I. M. Tibuleac, and R. a. Hansen (2007), A Semiautomatic Calibration Method
 486 Applied to a Small-Aperture Alaskan Seismic Array, *Bull. Seismol. Soc. Am.*, 97(1B), 100–
 487 113, doi:10.1785/0120040119.

488 Liu, P., and R. Archuleta (2006), Kinematic Inversion of the 2004 M 6.0 Parkfield Earthquake
 489 Including an Approximation to Site Effects, *Bull. Seismol. Soc. Am.*, 96(4B), S143–S158,
 490 doi:10.1785/0120050826.

491 Maercklin, N., G. Festa, S. Colombelli, and A. Zollo (2012), Twin ruptures grew to build up the
 492 giant 2011 Tohoku, Japan, earthquake., *Sci. Rep.*, 2, 709, doi:10.1038/srep00709.

493 Meng, L., and J.-P. Ampuero (2013), Design and Validation of A Network of Seismic Arrays for
 494 Earthquake Early Warning, *In Prep.*

495 Meng, L., A. Inbal, and J.-P. Ampuero (2011), A window into the complexity of the dynamic
 496 rupture of the 2011 Mw 9 Tohoku-Oki earthquake, *Geophys. Res. Lett.*, 38(7), L00G07,
 497 doi:10.1029/2011GL048118.

498 Meng, L., J.-P. Ampuero, J. Stock, Z. Duputel, Y. Luo, and V. C. Tsai (2012a), Earthquake in a
 499 maze: compressional rupture branching during the 2012 M(w) 8.6 Sumatra earthquake.,
 500 *Science*, 337(6095), 724–6, doi:10.1126/science.1224030.

501 Meng, L., J.-P. Ampuero, a. Sladen, and H. Rendon (2012b), High-resolution backprojection at
 502 regional distance: Application to the Haiti M 7.0 earthquake and comparisons with finite
 503 source studies, *J. Geophys. Res.*, 117(B4), B04313, doi:10.1029/2011JB008702.

504 Mount, V. S., and J. Suppe (1987), State of stress near the San Andreas fault : Implications for
 505 wrench tectonics, *Geology*, 15, 1143–1146, doi:10.1130/0091-7613(1987)15<1143.

506 Nakahara, H., H. Sato, T. Nishimura, and H. Fujiwara (2011), Direct observation of rupture
 507 propagation during the 2011 off the Pacific coast of Tohoku Earthquake (M w 9.0) using a
 508 small seismic array, *Earth, Planets Sp.*, 63(7), 589–594, doi:10.5047/eps.2011.06.002.

509 Niazi, M. (1966), Corrections to apparent azimuths and travel-time residuals for a dipping
510 Mohorovicic discontinuity, *Bull. Seismol. Soc. Am.*, 56(2), 491–509.

511 O'Rourke, T., S.-S. Jeon, S. Toprak, M. Cubrinovski, and J. K. Jung (2012), Underground
512 Lifeline System Performance during the Canterbury Earthquake Sequence, in *15th World*
513 *Conference on Earthquake Engineering, Lisbon, Portugal, Lisbon*.

514 Otsuka, M. (1966a), Azimuth and slowness anomalies of seismic waves measured on the
515 Central California Seismographic Array, part I: Observations, *Bull. Seismol. Soc. Am.*,
516 56(1), 223–239.

517 Otsuka, M. (1966b), Azimuth and slowness anomalies of seismic waves measured on the
518 Central California seismographic array, Part II: Interpretation, *Bull. Seismol. Soc. Am.*, (3),
519 655–675.

520 Roten, D., H. Miyake, and K. Koketsu (2012), A Rayleigh wave back-projection method applied
521 to the 2011 Tohoku earthquake, *Geophys. Res. Lett.*, 39(2), n/a–n/a,
522 doi:10.1029/2011GL050183.

523 Roux, P., K. Sabra, P. Gerstoft, and W. A. Kuperman (2005), P-waves from cross-correlation of
524 seismic noise, *Geophys. Res. Lett.*, 32(19), L19303, doi:10.1029/2005GL023803.

525 Thurber, C., H. Zheng, F. Waldhuser, J. Hardebeck, A. Michael, and D. Eberhart-Philips (2006),
526 Three-Dimensional Compressional Wavespeed Model, Earthquake Relocations, and Focal
527 Mechanisms for the Parkfield, California, Region, *Bull. Seismol. Soc. Am.*, 96(4B), S38–
528 S49, doi:10.1785/0120050825.

529 Vallée, M., M. Landès, N. M. Shapiro, and Y. Klinger (2008), The 14 November 2001 Kokoxili
530 (Tibet) earthquake: High-frequency seismic radiation originating from the transitions

531 between sub-Rayleigh and supershear rupture velocity regimes, *J. Geophys. Res.*,
532 113(B7), B07305, doi:10.1029/2007JB005520.

533 Waldhauser, F., W. Ellsworth, D. P. Schaff, and A. Cole (2004), Streaks, multiplets, and holes:
534 High-resolution spatio-temporal behavior of Parkfield seismicity, *Geophys. Res. Lett.*,
535 31(18), L18608, doi:10.1029/2004GL020649.

536 Wei, S. et al. (2011), Superficial simplicity of the 2010 El Mayor–Cucapah earthquake of Baja
537 California in Mexico, *Nat. Geosci.*, 4(9), 615–618, doi:10.1038/ngeo1213.

538 Wei, S., R. Graves, D. Helmberger, J.-P. Avouac, and J. Jiang (2012), Sources of shaking and
539 flooding during the Tohoku-Oki earthquake: A mixture of rupture styles, *Earth Planet. Sci.*
540 *Lett.*, 333-334, 91–100, doi:10.1016/j.epsl.2012.04.006.

541 Yao, H., P. Gerstoft, P. M. Shearer, and C. Mecklenbräuker (2011), Compressive sensing of the
542 Tohoku-Oki Mw 9.0 earthquake: Frequency-dependent rupture modes, *Geophys. Res.*
543 *Lett.*, 38(20), n/a–n/a, doi:10.1029/2011GL049223.

544 Zhang, H., Z. Ge, and L. Ding (2011), Three sub-events composing the 2011 off the Pacific
545 coast of Tohoku Earthquake (M w 9.0) inferred from rupture imaging by back-projecting
546 teleseismic P waves, *Earth, Planets Sp.*, 63(7), 595–598, doi:10.5047/eps.2011.06.021.

547

548

Figure and caption:

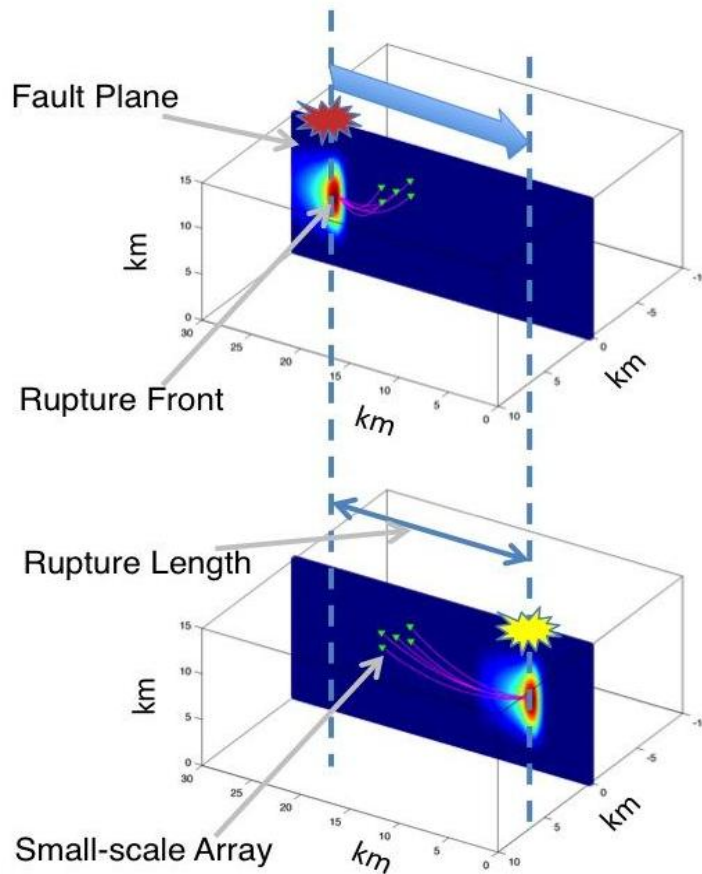


Figure 1. Concept of earthquake rupture imaging with a small-scale seismic array. The advancing slip pulse is the region of high slip velocity indicated by warm colors on the fault plane. The red and yellow “bang” symbols represent surface projections of the rupture front. The green triangles are the arrays stations. The pink curve is the ray path of the incoming seismic waves. The dashed lines mark the spatial extent of the rupture.

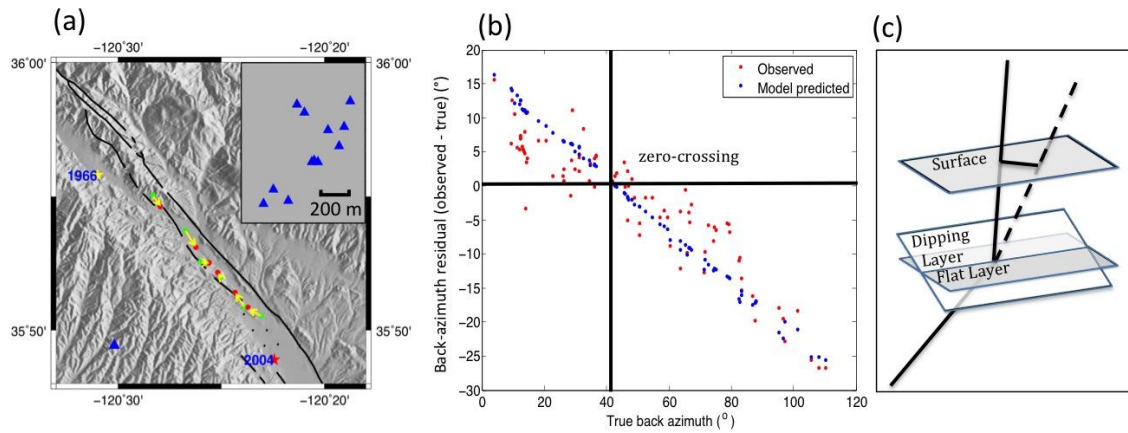


Figure 2. (a) Map view of the Parkfield area. The green dots are the epicenters of 5 sample aftershocks of the 2004 Parkfield earthquake. The red dots are the projected location of the aftershocks along a linear fault plane with azimuth retrieved by array analysis. The yellow arrows mark direction of the bias. The red and yellow stars are the epicenters of the 1966 and 2004 event. The blue triangle is the location of the UPSAR array with the distribution of the stations in the inset map. (b) the red dots are back-azimuth residual (observed - true) as a function of true back azimuth, the blue dots are the anomalous azimuth predicted by a 22 degree dipping interface toward 223 degree (c) Sketch of ray path (dash line) perturbed by a dipping layer and the expected ray path through horizontal layer (solid line).

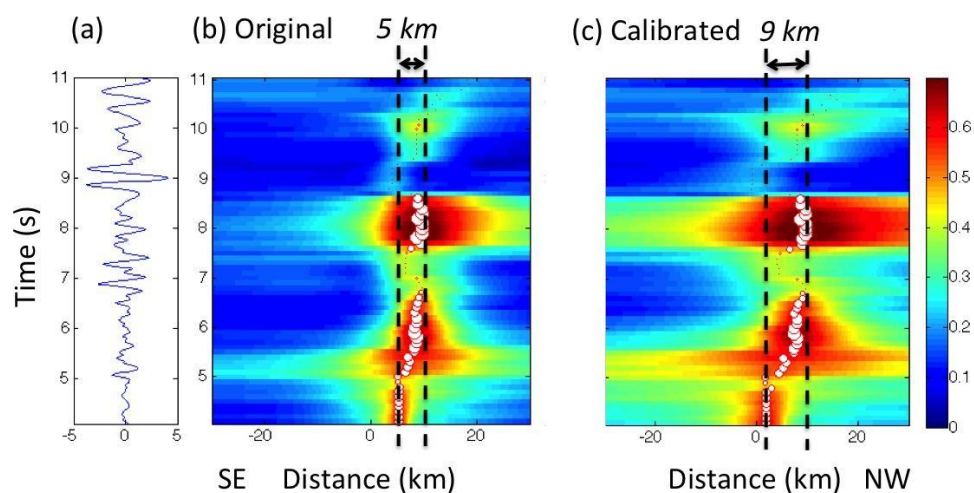


Figure 3. Stacked correlation coefficient as a function of along-strike position with respect to the hypocenter and time is shown in color. Panel **a** shows the seismogram at the central station of the UPSAR array. Panels **b** and **c** show the stacked correlation coefficient as a function of source position (projected along the fault strike, relative to the epicenter) and time (of the center of the sliding window relative to the event origin time), without and with correction for the dipping interface effect, respectively. The white dots are the locations of maximum correlation at each time sized by the correlation coefficient. The dashed lines mark the beginning and end of the HF energy source and give estimates of the rupture size of 5 km (b) and 9 km (c).

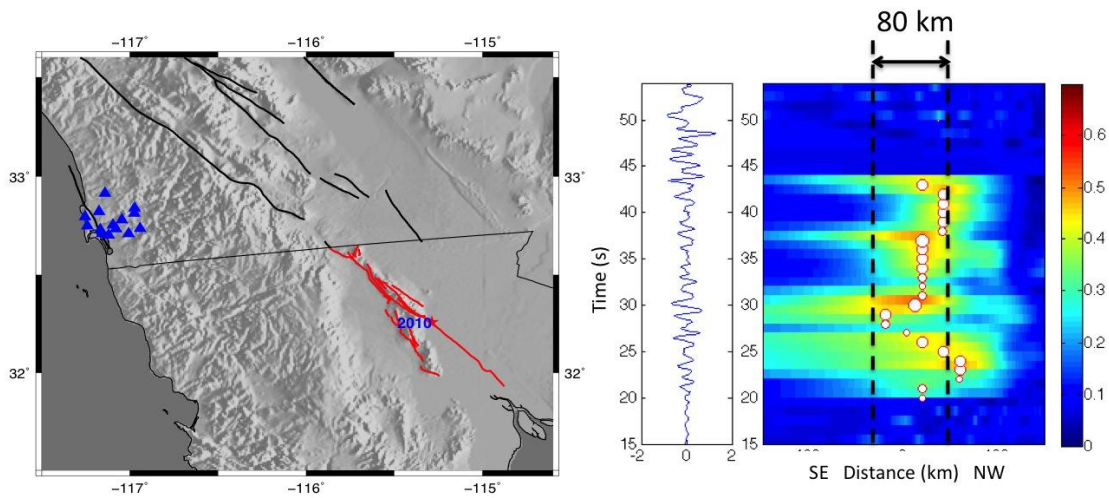


Figure 4. El Mayor Cucapah earthquake imaged by the San Diego strong motion array. The map shows the San Diego strong motion array (blue triangles) and the fault traces of the earthquake (red). The panel on the right is the correlation stacking analysis of the mainshock plotted in the same fashion as figure 3.

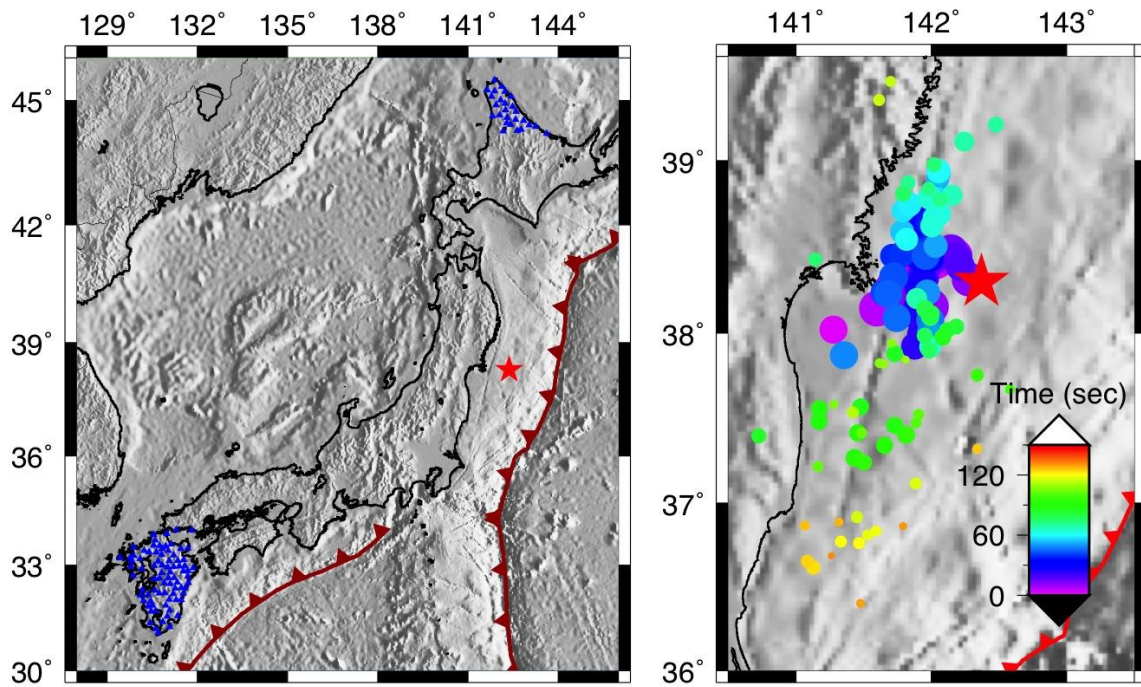


Figure 5. Tohoku earthquake imaged by clusters of Hi-net stations in Kyushu and Northern Hokkaido. The map shows the selected stations (blue triangles) and trenches (dark red). The red star denotes the hypocenter. The circles in the right panel are the locations of seismic radiation found by combining the back-azimuths determined by both arrays, colored by rupture time and sized by the sum of the correlation coefficients from the two arrays.