Robust determination of earthquake source parameters and mantle attenuation

Yen-Ting Ko,[1,2] Ban-Yuan Kuo,[1] and Shu-Huei Hung[2]

Received 8 August 2011; revised 13 February 2012; accepted 14 February 2012; published 12 April 2012.

[1] An improved inversion technique is needed to effectively separate the frequency dependence of the source from the intrinsic attenuation of the medium. We developed a cluster-event method (CEM) in which clusters of nearby events, instead of individual events, pair with stations to form the basis for measurements of Q value and corner frequency (f_c). We assume that the raypaths from one cluster to a station share an identical Q while each event in the same cluster is allowed for only one f_c in the inversion process. This approach largely reduces the degrees of freedom to achieve a robust inversion. We use an optimization algorithm of simulated annealing to solve the nonlinear inverse problem. The CEM was applied to events at 70–150 km depths in the Japan subduction zone recorded by F-net. We show that the method proposed here leads to better constraints on both source parameters and attenuation. The resultant Q's in the mantle wedge increase from lower than 300 beneath the arc and back-arc to greater than 600 in the fore-arc region. The f_c's satisfy a self-similar scaling relationship with seismic moment of $M_0 \propto f_c^{-3}$ with a best fit stress drop of 21.9 ± 6.9 MPa in Madariaga's form. This contrasts to the stress drop of 1.4 ± 1.1 MPa for a global data set composed of prior measurements for crustal events. The results of this study agree with results from previous studies, except with an upward deviation due to higher corner frequencies and stress drops.

Citation: Ko, Y.-T., B.-Y. Kuo, and S.-H. Hung (2012), Robust determination of earthquake source parameters and mantle attenuation, *J. Geophys. Res.*, *117*, B04304, doi:10.1029/2011JB008759.

1. Introduction

[2] Study of attenuation involves measuring the spectral content of a waveform, which is more difficult than measuring the traveltime. A broadband power spectrum substantially above the noise level is required to extract the correct spectral characteristics of the seismic signals. The earthquake source has its own frequency dependence, sometimes characterized by a corner frequency f_c, which convolves with path effects to compose a seismogram. These path effects include focusing/defocusing and scattering during traveling blended with intrinsic attenuation to shape the spectrum at a seismic station.

[3] One strategy to discriminate the source and path effects is to eliminate either of the two with specific tools. Examining the ratio of the spectra recorded by two stations cancels the source contribution [e.g., *Roth et al.*, 1999; *Shito et al.*, 2004]. In studies of earthquake sources, the spectral ratio is usually calculated between 2 events to eliminate the path effects. Examples in this scheme are the empirical Green's function (EGF) method if one event is small enough [e.g., *Mueller*, 1985; *Mori and Frankel*, 1990; *Abercrombie and Rice*, 2005; *Viegas et al.*, 2010] and the determination of 2 corner frequencies rather than the conventional 2 parameters for source and path [e.g., *Tajima and Tajima*, 2007]. Although the spectral ratio mitigates the tradeoff problem, a reference station or event must be chosen, which may limit the applicability of the method in mapping large-scale attenuation structure of a region.

[4] Other approaches tackle the tradeoff problem during or after the inversion of the spectrum. The f_c can be obtained by fitting individual spectra recorded at each station, but this leads to multiple f_c's for one event. Despite f_c's are in principle path-dependent for rupturing of a finite fault, we show in this study that the multiple f_c's determined with paths are primarily artifacts of inversion due to tradeoffs between parameters. One way to circumvent this is to average all the f_c's over stations for each event and fix it in the second stage of the determination of Q [e.g., *Wang et al.*, 2010]. However, a better remedy inversion-wise is to implement the single f_c requirement in the inversion process for the one-event, many-station system, anticipating a reduced directivity effect averaged over stations. This scheme was first used by *Stachnik et al.* [2004] and adopted in later studies [e.g., *Rychert et al.*, 2008; *Pozgay et al.*, 2009]. Additional constraints of f_c including the existing moment-f_c relationship can be exercised to bracket the solution [e.g., *Pozgay et al.*, 2009].

[5] Recently, *Shearer et al.* [2006] proposed stacking of event spectra within bins of similar seismic moment before

[1]Institute of Earth Sciences, Academia Sinica, Taipei, Taiwan.
[2]Department of Geosciences, National Taiwan University, Taipei, Taiwan.

0148-0227/12/2011JB008759

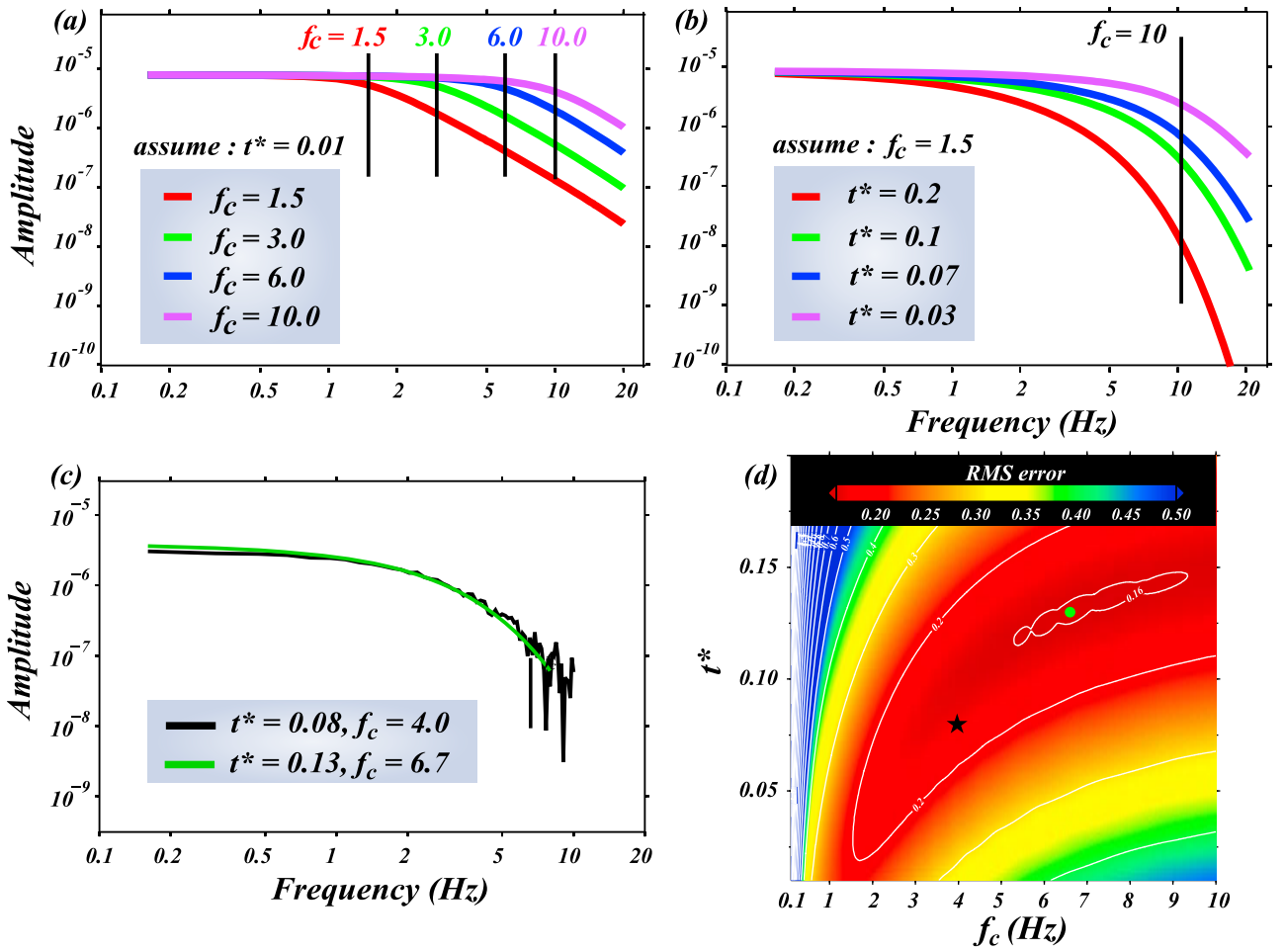

Figure 1. Illustration of the tradeoff between t^* and f_c with synthetic spectra computed from equation (1). Both parameters act to bend the spectral amplitude with increasing frequency. Example spectral curves are for (a) $t^* = 0.01$ and $f_c = 1.5$, 3, 6, and 10 Hz and for (b) $f_c = 10$ Hz, and $t^* = 0.2$, 0.1, 0.07, and 0.03. (c) Synthetic spectra for (0.08, 4.0) (black line, with random error) and (0.13, 6.7) (green) are indistinguishable. The former is taken as data for the grid-search in Figure 1d. (d) The misfit map shows a curve-shaped low misfit valley typical of fitting the spectrum with t^* and f_c. The minimum is located at (0.13, 6.7) (green dot), instead of (0.08, 4.0) (black star). The tradeoff results in no robust solution on the misfit map.

computing the EGF. *Mayeda et al.* [2007] applied the spectral ratio method on *S* coda and demonstrated that the path effect can be canceled more efficiently than for direct seismic phases. These improved methods, however, are not designed to extract both source characteristics and intrinsic attenuation of the medium in one inversion. In this work, we present the cluster-event method (CEM) to simultaneously determine f_c and Q from "cluster"-station pairs. In addition to enforcing f_c as the single-valued parameter for each event, the CEM inversion scheme groups raypaths based on event "clusters" to reduce the degrees of freedom of the inversion and suppress the tradeoff effects. In this study we focus on demonstration and evaluation of the method in the Japan subduction zone. A better determination of f_c also allows us to re-examine the f_c-moment scaling problem.

2. The Issue of Tradeoff

[6] We constructed synthetic spectra to examine the contributions of different parameters to a spectrum. A displacement spectrum $A(f)$ with a Brune-type source model [*Brune*, 1970] can be written as [e.g., *Anderson and Hough*, 1984]

$$A(f) = \frac{\Omega_0 \exp(-\pi f t^*)}{1 + \left(\frac{f}{f_c}\right)^2}. \quad (1)$$

where

$$\Omega_0 = C \cdot M_0, \qquad C = \frac{{}_cU_{\phi\theta}}{4\pi\rho v_p^3 R}, \quad (2)$$

$$t^* = \int_{path} \frac{d\tau}{Q}. \quad (3)$$

In equations (1) and (2), f is the frequency and Ω_0 is the low-frequency asymptote of the spectrum related to seismic moment M_0 through a frequency independent constant C, which is composed of v_p, the P wave velocity near the source, ρ, the density of the source material, R, the source-receiver

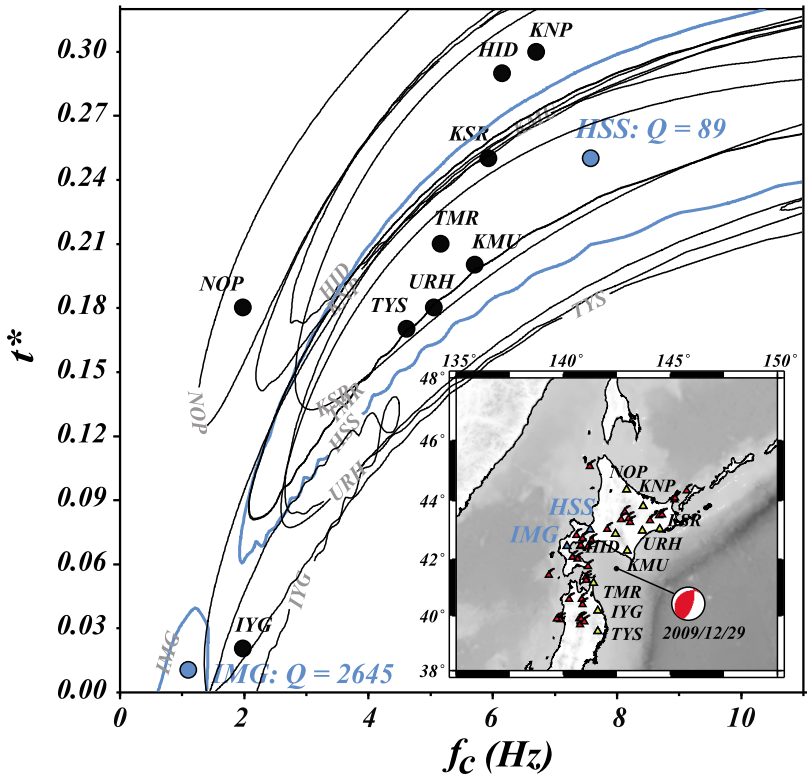

Figure 2. Illustration of wide-spread, distinct f_c's obtained for one event. Solutions (t^*, f_c) were determined by individually fitting the spectra recorded by 11 F-net stations for a Japan subduction zone event (2009/12/29, depth 80 km). Red triangles are arc volcanoes in this region. Stations IMG and HSS are back-arc affiliated and spatially close, but their resolved Q values differ significantly. The uncertainty of the solution (thin contour) illustrates the tradeoff between the two parameters.

distance, and ${}_cU_{\phi\theta}$, the spherical average of radiation patterns (0.52 for P) [e.g., *Aki and Richards*, 1980; *Abercrombie*, 1995]. The exponential term in equation (1) describes the along-path attenuation, in which t^* is the attenuation operator defined in the path integral in equation (3). M_0 is a constant parameter that does not tradeoff with t^* and f_c significantly and can be determined independently as described in section 6.3.

[7] High-temperature laboratory experiments [*Jackson et al.*, 2002] as well as seismic attenuation study in mantle wedge [*Stachnik et al.*, 2004] suggested that the attenuation operator is frequency-dependent in the form of

$$t^* = t_0^* f^{-\alpha}, \tag{4}$$

where t_0^* is the t^* value at 1 Hz, and α accounts for the frequency dependence of intrinsic attenuation. In this study, α is fixed at an experimental value of 0.27 [*Jackson et al.*, 2002, 2006]. Inversion of α as a free parameter is usually non-robust, and a few studies have reported inversion results of α in the range of 0–0.4 with only marginal preferences over α [e.g., *Stachnik et al.*, 2004; *Rychert et al.*, 2008; *Pozgay et al.*, 2009].

[8] Figure 1 shows that, while t^* and f_c in principle shape the spectral decay in different fashions, their effects can be similar and hard to distinguish in the presence of noise. In this example, the "data" is generated by blending the spectrum for t^* = 0.08 and f_c = 4.0, or (0.08, 4.0) with the Gaussian random noise which mimics the scatter of data at high frequencies well. We perform a grid-search against this spectrum in the (t^*, f_c) space, yielding a misfit map with a long, low-value valley, indicating a serious tradeoff between the parameters. Solutions along this valley explain the data equally well in a statistical sense, and therefore neither of the two can be determined in a robust way. The synthetic test outputs a best fit model of (0.13, 6.7), which is quite distant from the input (0.08, 4.0) (Figure 1d).

[9] Figure 2 elaborates the tradeoff with real data. We analyzed P waves recorded by 11 F-net stations for a subduction zone event in Japan. Fitting individual spectrum through a grid search yields 11 sets of (t^*, f_c) that scatter from 0.01 to 0.31 for t^* and from 1 to 8 Hz for f_c. The uncertainty interval is characterized by a misfit value 5% larger than the minimum [*Sonley and Abercrombie*, 2006]. The 2 extreme Q values, 89 and 2645, were determined for 2 stations, HSS and IMG, respectively, which are both located in the back-arc environment of the subduction system and are closer to each other than most of the other pairs of stations. This exemplifies the tradeoff problem in the sense that the 2 parameters are highly mutually substitutable in minimizing the misfit. Taking the average of the scattered measurements as the representative f_c of the event may obviate the over-parameterization problem [e.g., *Wang et al.*, 2010], but a more fundamental treatment should be sought to secure a correct recovery of the physical property of the earthquake.

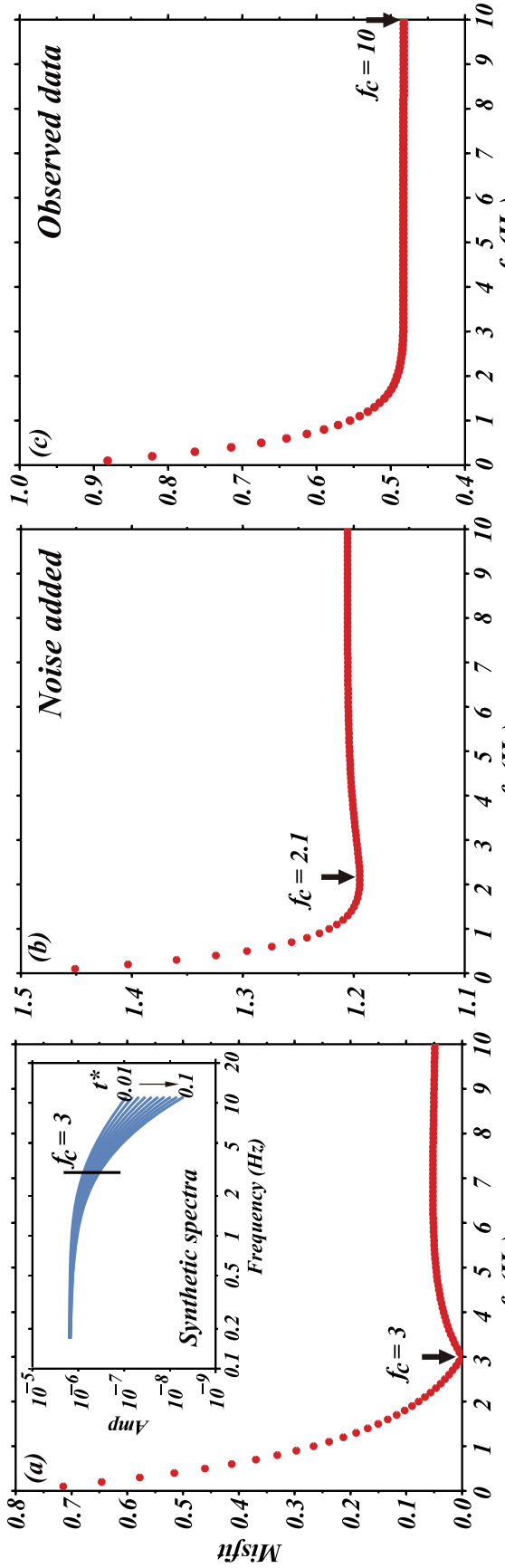

Figure 3. Test of inversion using equation (5), *Stachnik et al.*'s [2004] approach. Each point at the curve is the misfit value for a given f_c and the Q value found to minimize the misfit at that f_c. Black arrow points out the numerical optimal solution in each case. (a) 10 synthetic curves without random errors for f_c = 3 Hz and t^* evenly distributed between 0.01 and 0.1. The input f_c is correctly recovered. (b) The same set of curves as in Figure 3a but with random errors. Location of minimum is ambiguous and statistically insignificant. (c) Inversion for the real data in Figure 2 provides a one-sided constraint with a numerical minimum at the frequency upper limit.

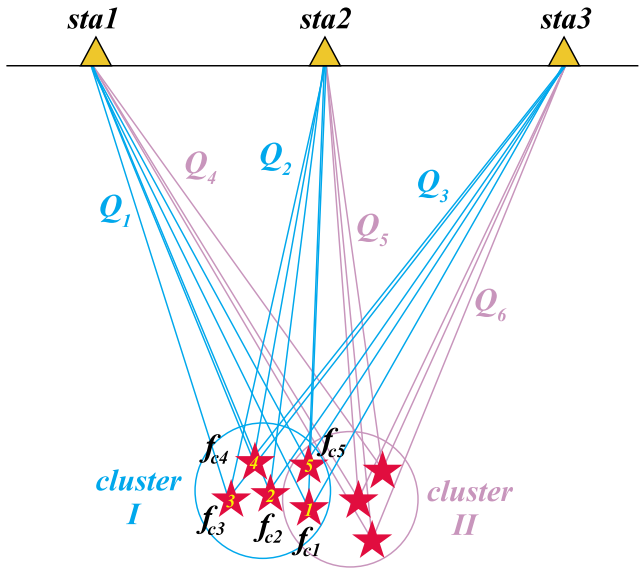

Figure 4. Synopsis of the CEM. Nearby events (events 1–5) are grouped into a cluster and the raypaths from this cluster to the same station share the same Q while each event in this cluster is assigned only one f_c. The single f_c constraint applies only to events in the same cluster. Clusters may overlap and share events, e.g., cluster I and II share events 1 and 5, each of which may have different values determined for different clusters. These cross-cluster events provide data to examine how robustly each f_c is determined.

[10] *Stachnik et al.* [2004, hereinafter ST] formulated a one-event, many station system in which f_c is allowed for only a single value. Equation (1) can be rearranged into the form

$$\ln(A(f)) + \ln\left(\frac{1}{1+\left(\frac{f}{f_c}\right)^2}\right) = -\pi f t^* + \ln(\Omega_0). \qquad (5)$$

[11] For each fixed f_c, equation (5) is linear for t^* and ln (Ω_0), enabling a linearized inversion for Q and Ω_0 for all paths. The best f_c can be found to give the "global" minimum misfit for all trial f_c's. The linearized equation (5) has also been a common tool in the investigations whose major concern is earthquake properties rather than Q structures [e.g., *Anderson and Humphrey*, 1991; *Ide et al.*, 2003]. We performed a synthetic test simulating a case in which one event with f_c = 3 Hz is recorded by 10 stations giving rise to 10 values of t^*. Solving equation (5) for a range of trial f_c's locates the minimum misfit at f_c = 3 Hz successfully (Figure 3a). However, this minimum is by nature non-robust, and in the presence of random errors, can be ambiguous or even absent as the misfit curve turns to an L-shape (Figure 3b). We applied the ST method to the real data analyzed above (Figure 2), and obtained a loose lower bound for f_c and an open solution at higher frequencies (Figure 3c).

3. Methodologies and Data Analysis

3.1. Determining the t^* Values and Source Parameters

[12] Judging from the exercises in Figure 3, merely imposing single f_c does not guarantee a robust solution. In this section, we describe the approach used in the cluster-event method (CEM) which we show achieves better constraints on both source and path parameters for events in a seismically active region.

[13] Figure 4 summarizes the basic idea of the CEM. We regard the nearby events (events 1–5) as a cluster, and assume that all the raypaths from a cluster to the same station are close enough to share an identical Q while each event in the cluster, usually recorded by several stations, is allowed only for one single f_c. Assuming K events and N stations, the degrees of freedom for CEM, ST, and individual fitting are $K + N$, $(K \times N) + K$, and $(K \times N) \times 2$, respectively. In the example in Figure 4, there are only 8 parameters (3 Q's and 5 f_c's) for cluster I to be determined to fit the spectral falloff (other than the constant Ω_0). The ST approach requires 20 parameters, while fitting individual spectra at all stations involves 30 unknowns (a pair of Q and f_c for 15 paths). The CEM gains robustness by efficiently reducing the degrees of freedom of the inverse problem while obey the basic physic of earthquake sources under the point source assumption. The CEM comes not without a cost, which is a sacrifice in spatial resolution of Q when bundling many otherwise independent, path-average Q's into one. This cost, however, becomes transparent as the spatial size of the cluster decreases and the raypaths converge. Meanwhile, as demonstrated above, attaching significance to each along-path Q seems impractical or even misleading.

[14] Following equation (1), we formulate a misfit function in L2 norm

$$\Theta = \sum_{k=1}^{M} \left\| \left[D_{kj}(f) - \frac{\Omega_{0kj}\exp(-\pi f t_j^*)}{1+\left(\frac{f}{f_{c_k}}\right)^2} \right] \cdot \varphi_{kj} \right\|,$$
$$\varphi_{kj} = \frac{\Delta f_{kj}}{\sum_{k=1}^{M}\sum_{j=1}^{N}\Delta f_{kj}}, \qquad (6)$$

where D_{kj} is the observed spectrum for kth event within a cluster recorded by jth station. To weight more strongly the contribution from the spectrum with a longer frequency range $D(f)$, a weighting factor φ_{kj} which is proportional to $D(f)$ is applied to each spectrum in the calculation of the total cluster misfit. Equation (6) is cluster-based, so cluster index is omitted. Note that Q in t^* along the cluster-station path is the attenuation parameter to be solved for. Θ measures the deviation of the theoretical spectrum from the observed spectrum for each cluster. An optimization algorithm of simulated annealing (SA) [*Kirkpatrick et al.*, 1983], a generic probabilistic metaheuristic for the global optimization problem, is employed to seek global minimum of the misfit function in a multidimensional parameter space. A grid search can achieve the same goal but much inefficiently. For more detailed descriptions of the SA method see *Aarts and Korst* [1989].

3.2. Synthetic Tests

[15] To compare the solutions obtained from CEM and the other approaches, we carried out a series of synthetic tests with 3 stations and 3 events. We present an example with t^* = 0.03, 0.05 and 0.08, and f_c = 2, 4 and 6 Hz in Figure 5.

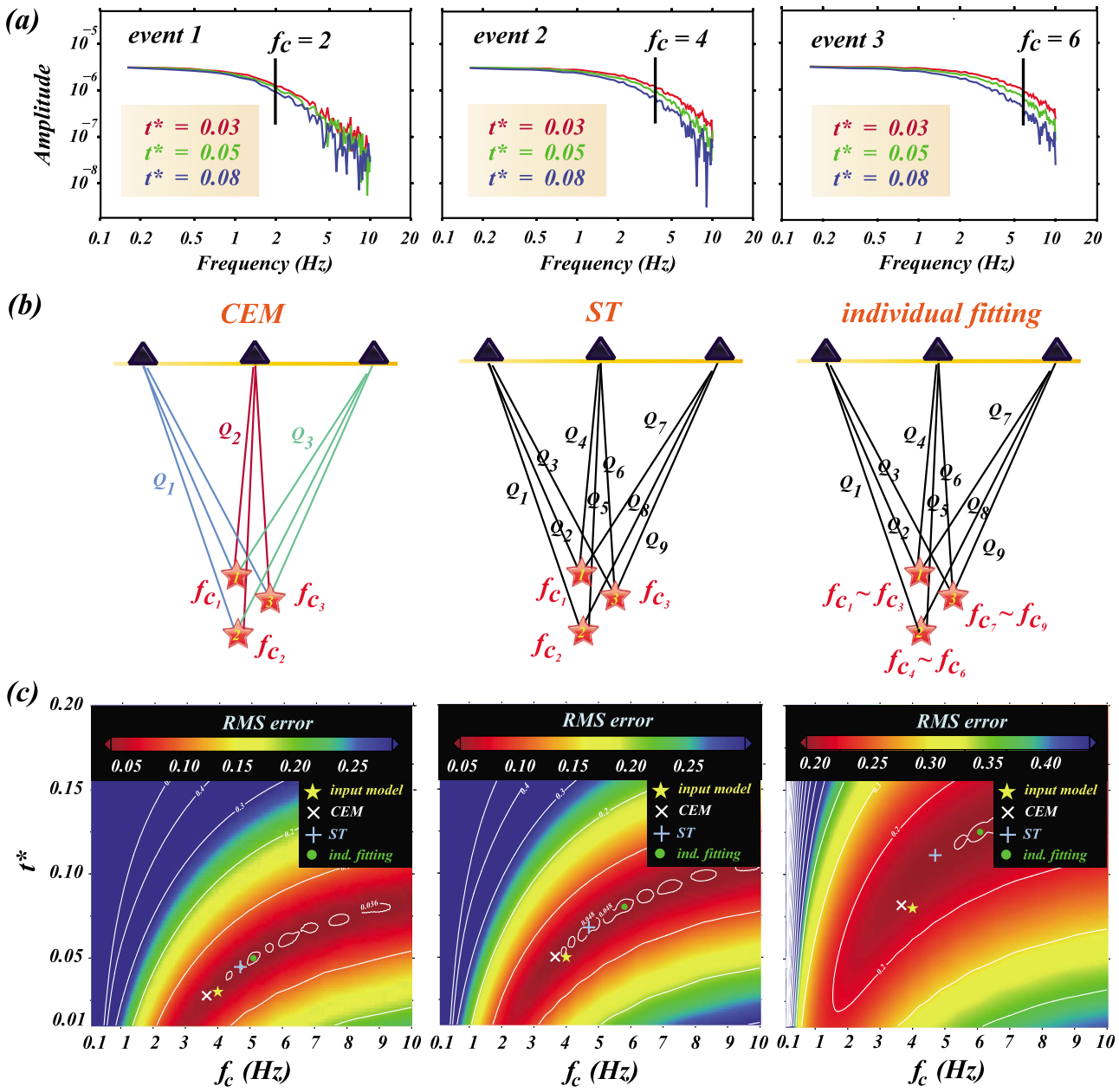

Figure 5. Test of CEM and the two other approaches. (a) Synthetic data with t^* and f_c annotated. (b) Configurations of the 3 approaches with increasing degrees of freedom: CEM, ST method, and individual fitting. (c) The misfit maps for synthetic models (0.03, 4), (0.05, 4), and (0.08, 4), respectively, from left to right. The input models are denoted by yellow stars. The solutions from CEM, ST, and individual fitting are denoted by white x, blue cross, and green dot, respectively. The CEM recovers the input t^* and f_c satisfactorily whereas the other 2 approaches fumble in finding the correct solutions.

The synthetic data are theoretical spectra with random errors (Figure 5a), and the synopsis of the three approaches under test are (Figure 5b) (1) The CEM with 6 parameters (3 Q and 3 f_c) for spectral falloff; (2) The ST approach with 12 parameters (9 for Q and 3 for f_c); and (3) Fitting individual spectra to solve for 18 parameters (9 pairs of Q and f_c).

[16] Figure 5c shows that the CEM returns a solution that is closest to the input model, demonstrating a strong resolvability despite data shrouded in noise. In contrast, the other 2 approaches yield parameters deviating from the input, especially for the highly over-parameterized individual spectral fitting. This occurs as a general phenomenon in a series of tests with wide choices of input combinations. In some cases, when different sets of t^* and f_c give similar shape of spectral decay, it is the noise that determines the numerical best fit. A similar conclusion was reached by *Sonley and Abercrombie* [2006] that a major tradeoff occurs

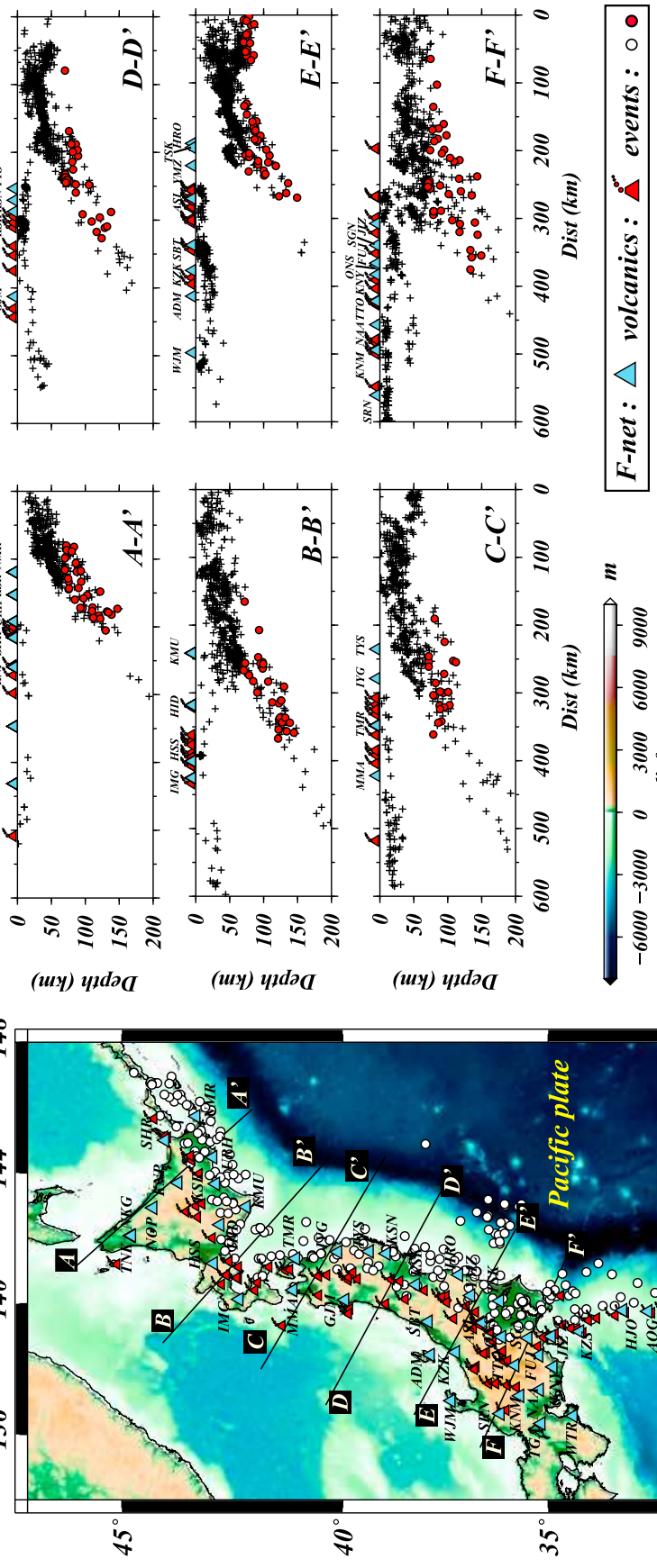

Figure 6. The CEM is applied to events in Japan subduction zone. (left) Events (white circle) and F-net stations (blue triangle) used, active volcanoes (red triangle), and locations of vertical cross-sections (black line). (right) Cross-sections show the subducting Pacific slab (background events in black crosses), events used (red circle) with focal depths 70–150 km, and volcanoes and stations on the surface. Each cross-section is 200 km wide.

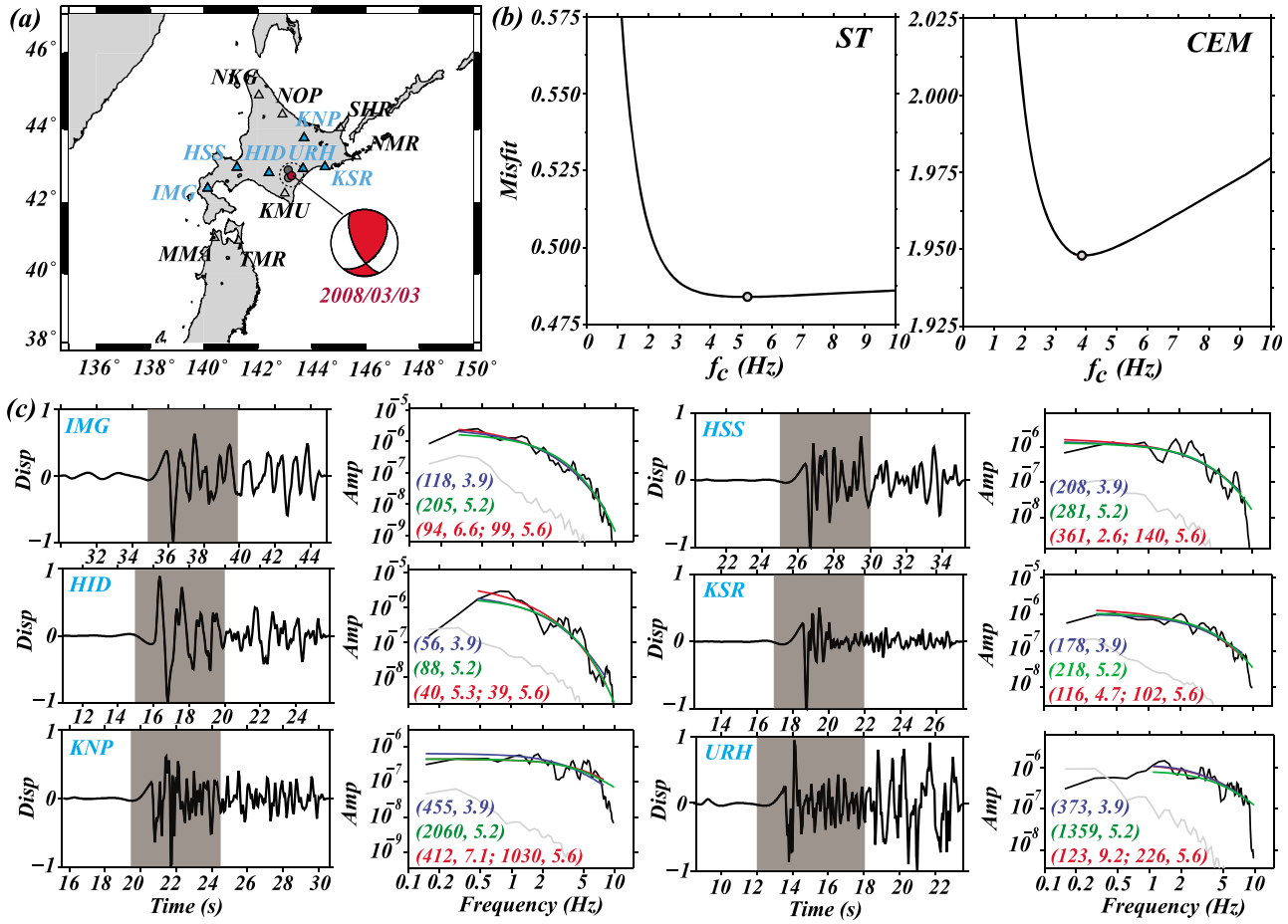

Figure 7. Example of CEM measurements and solutions. (a) Map of the 3-event cluster (circled by dotted line) and the event illustrated as the example (red circle with focal mechanism; date 2008/03/03 and depth 83 km). Twelve stations (triangle) received good P waves and 6 of them (blue triangle) are shown in Figure 7c. (b) Solutions (open circle) from ST method (left), and CEM (right). The CEM curve is constructed by the SA procedure explained in the text. (c) Example waveforms and spectra. P wave windows are shaded. Each spectrum is compared with solutions from CEM (blue curve), *Stachnik et al.*'s [2004] method (green curve), and individual fitting (red curves), with solutions (Q, f_c) annotated on the lower-left corner of each spectrum panel. For individual fitting, the second solution is for the averaged f_c of 5.6.

between t^* and f_c. The CEM suppresses this likelihood with much less degrees of freedom to satisfy all the spectrum data.

4. Data

4.1. Event Selection

[17] The CEM performs best in high seismicity regions where clusters can be easily formed. We choose the Pacific plate portion of the Japan subduction zone to apply the method, and the F-net [*Okada et al.*, 2004] waveforms, which have been archived since 1994, were used (Figure 6). In this study we use events deeper than 70 km to both deemphasize the crustal effects and thin the ray volume from cluster to station. 87% of the paths exceed 150 km in length. These paths travel a long distance in the mantle wedge, allowing us to map the attenuation structure there and simultaneously extract the source properties of intraslab earthquakes. The final data set consists of 243 events, 17,443 spectra, and 123 clusters. About 87, 12, and less than 1% of the paths are from events with M_w 3.5–5, 5–6, and >6, respectively.

4.2. Criteria of Clusters

[18] For each target event, neighboring events within a certain radius from the target event is considered of the same cluster. The smaller the radius, the closer the raypaths are to each other for one station and the better the shared Q assumption is satisfied, but the "density" of the cluster is lower. Because 87% of the path lengths are 150 km or longer, we adopt a radius of 30 km in this study to ensure that a predominance of data have a path length 5 times longer than the radius of the cluster. This is a compromise between the number of clusters and the validity of the similar-Q assumption. We tested the radii of 10 and 20 km and found

the results essentially the same as for the 30 km except many events are left isolated and abandoned. The region was scanned for each event with the 30 km radius to establish as many clusters as possible to maximize the coverage of the cluster-based raypaths. The density of cluster is 3–11 events with a majority exceeding 5. In this scheme, one event can belong to several overlapping clusters as illustrated in Figure 4. However, although the "one f_c" requirement is applied to each event in one cluster, the requirement is not enforced across clusters, leaving the cross-cluster events susceptible to multiple values of f_c determined from different clusters. This, although can be improved in the future, provides a means to test the statistical robustness of the f_c determined with CEM, which will be addressed in section 5.2.

4.3. Waveform and Spectra

[19] The vertical components of the velocity recordings were deconvolved to remove the instrument response. Only those *P* waves without dominating coda energy were used to minimize the interference from scattering between source and receiver. While coda waves average out the heterogeneity of the medium efficiently, we focus on body wave phases for which the physics of the intrinsic attenuation has long been established. Examples of waveforms and corresponding spectra are shown in Figure 7. *P* wave onsets were manually picked for traveltimes. The displacement spectrum is calculated from a 5 s window of waveform starting 0.5 s before the *P* arrival time with 0.5 s long cosine tapers on both ends. The noise spectrum is calculated from a window of the same length 1 s preceding the *P* wave window. The 5 s window limits the minim frequency resolvable at about 0.5 Hz. In our data set, the minimum f_c is 0.7 Hz for an M_w 5.5 event. We found that longer windows, e.g., 5–10 s, produce spectra that do not yield distinguishable results after inversion. This is consistent with *Sonley and Abercrombie*'s [2006] note that amplitude spectra are insensitive to the window length as long as it exceeds the expected pulse width of the target phase. The spectra were smoothed using a Hanning window. To avoid aliasing, the upper limit of the frequency band is set at 80% of the Nyquist frequency of the F-net data, or 8 Hz (Figure 7). Inversion is performed only for the frequency band with a signal-to-noise ratio greater than 5.

5. Results

[20] Figure 7b illustrates a typical constraint on f_c by CEM. While the ST inversion usually provides a loose lower limit of f_c, the CEM delivers a V-shape misfit curve to better constrain this parameter. In practice, the highly nonlinear inversion converges to a curve like that in Figure 7b in a few steps, but we repeated the SA process several times with different initial guesses to secure a global minimum. We also applied the ST method to the same data set, and the comparison in f_c between the CEM and the ST is shown in Appendix A, section A1.

5.1. Q Structure

[21] The path-average Q measurements for intraslab events mostly map the mantle wedge of the Japan subduction zone, which potentially constitutes a data set for follow-up investigations such as tomographic imaging. For the purpose of this study we describe the overall pattern of Q distribution. The first order feature is low Q values, i.e., <300, for paths sampling the mantle mostly beneath the arc and back-arc, in contrast to the high Q values of greater than 600 for paths traversing the fore-arc mantle (Figure 8). This pattern agrees with previous observations in Japan [e.g., *Tsumura et al.*, 2000] and other subduction zones [e.g., *Roth et al.*, 1999; *Schurr et al.*, 2003; *Stachnik et al.*, 2004; *Rychert et al.*, 2008]. Low Q values have been attributed to the high temperature and probably the presence of melt retained under the arc and back-arc. The high Q values in the fore-arc are thought to characterize the low temperature "nose" of the wedge corner initially induced by decoupling between the wedge corner and the subduction. Differences in hydroxyl content may also cause variations in attenuation (e.g., Y. T. Ko et al., The southwestern edge of the Ryukyu subduction zone: A high Q mantle wedge, submitted to *Earth and Planetary Science Letters*, 2011), but viable interpretations will rely upon an integration of evidence from attenuation and velocity observations, arc geochemistry, deformation experiments, and dynamic modeling [e.g., *Jackson et al.*, 2002; *Karato*, 2003; *Chou et al.*, 2009; *Lin et al.*, 2010].

[22] In addition to the first order back-arc–fore-arc variation, along-strike variations appear to exist in the measured Q values. At the two southernmost profiles, which straddle the region of strong change in trench curvature and slab dip, Q in the fore-arc is low relative to at the northern profiles. Dynamic modeling demonstrates that lateral flow may be driven by the pressure gradient induced by trench curvature and slab dip [*Kneller and van Keken*, 2008]. Further south, rocks of adakitic composition were reported that suggest the presence of a high-temperature subduction environment [*Sugimoto et al.*, 2006]. Detailed tomographic imaging will map the attenuation in 3D and shed light on the interpretation of this along-strike feature.

5.2. The Determination of Corner Frequency

[23] As described above, the single-f_c requirement is implemented in the inversion for each cluster, but not across the cluster, i.e., different values of f_c may be obtained for events shared by overlapping clusters. If the intracluster source parameter is resolved robustly, those cross-cluster f_c's should converge. This presents an opportunity to appraise the method. We calculated the statistics of the f_c's determined by CEM and individual spectrum fitting (Figure 9). The number of f_c's from CEM for each event equals the number of clusters containing this event, and the number of f_c's from individual fitting corresponds to the number of stations.

[24] Overall, the scatter in f_c for each event is significantly smaller for CEM than for individual fitting. We calculate the standard deviation of f_c for each event and average of the standard deviations for all events. The average is 0.56 Hz for CEM, compared with 3.84 Hz for individual fitting. Almost 90% of the events under test have standard deviations < 1.3 Hz, indicating a robust determination of f_c within each cluster. For each event, the mean f_c's from the two approaches differ by 0.1 to 5 Hz, with 2/3 of the events

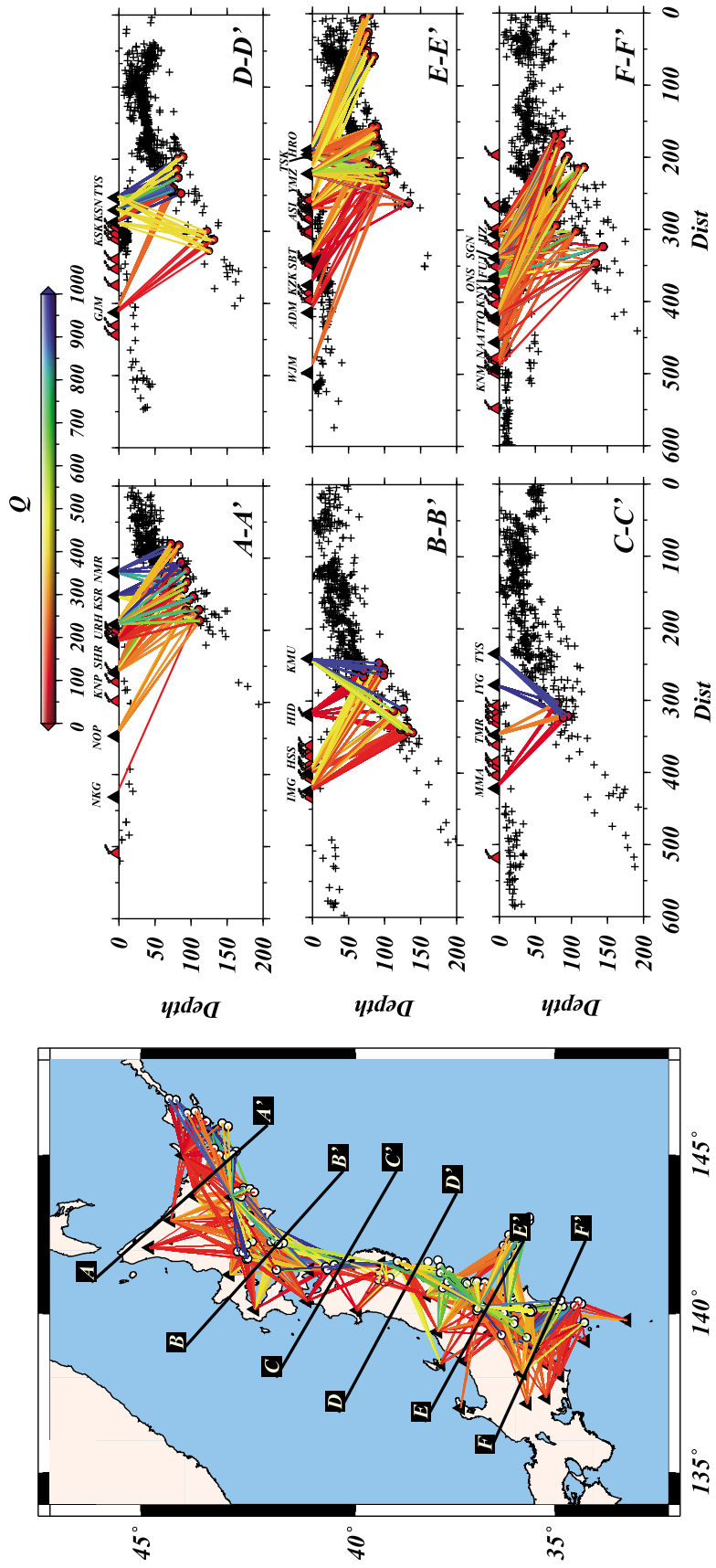

Figure 8. Path-dependent Q (color-coded) derived from CEM (left) on the map view and (right) in cross-sections of Japan subduction zone. All the paths with $Q > 600$ appear to sample the fore-arc mantle, while those with $Q < 300$ travel to the arc and back-arc.

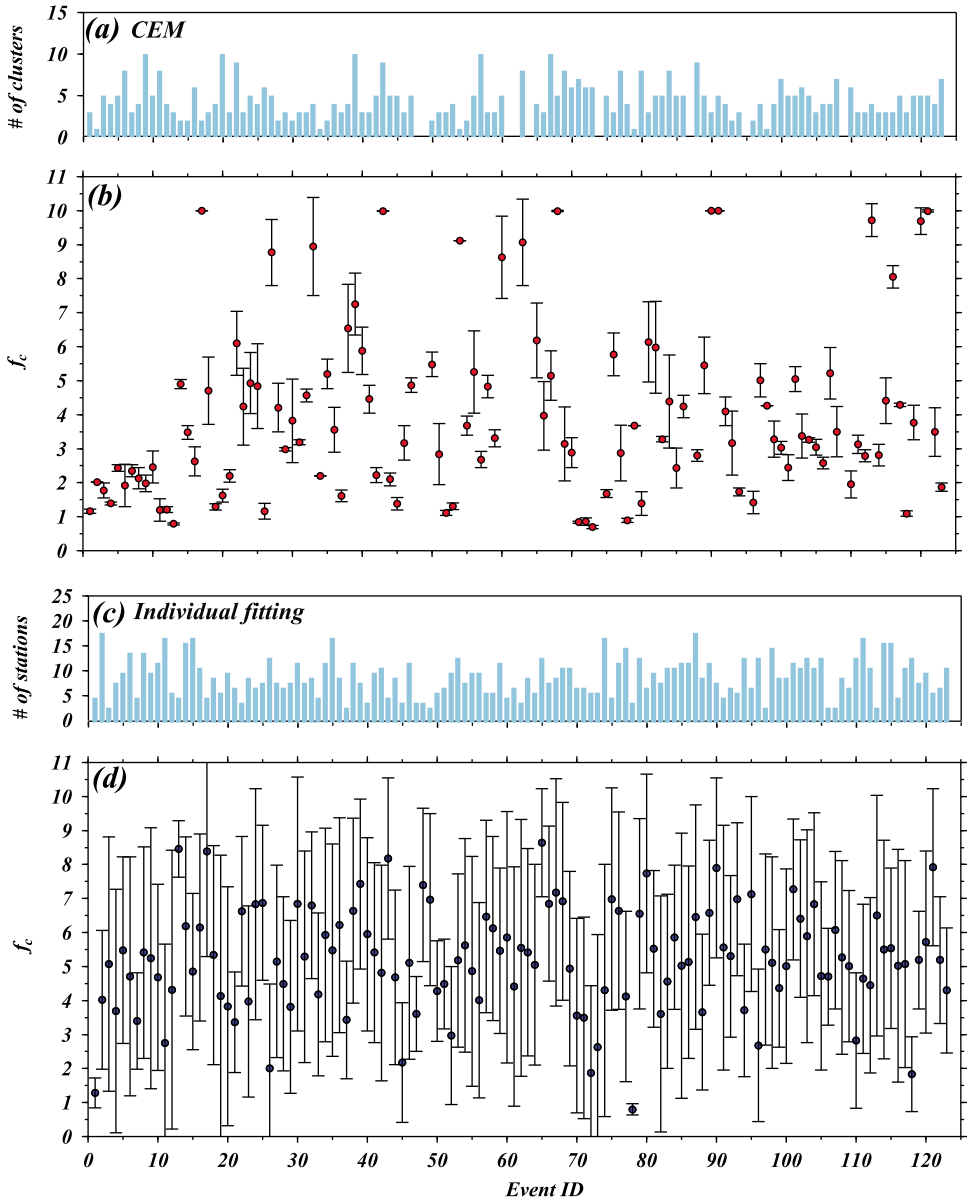

Figure 9. Comparison of f_c's derived from (top) CEM and (bottom) individual spectrum fitting. (a) Number of f_c's corresponding to number of clusters sharing one event (see text). (b) Red dots are the means of the f_c's obtained from CEM and the vertical bars are one standard deviations. (c and d) Same as Figures 9a and 9b, respectively, except for individual fitting. The small standard deviations of f_c for CEM, averaged to 0.56 Hz for all events, confirm a robust determination of the source parameter and consequently Q.

reaching 3 Hz or greater. This deviation is large enough to bias the interpretation of Q.

6. Discussion

6.1. Assumptions and Errors

[25] In this study, we ignore site effects, which are used to parameterize near-station geology or topography and may be isolated from the path effect. However, extracting site effects require both extra parameterizations and the spectra sensitive to surface structures. Recent studies using high density array data and crustal events indicate that the frequency-dependent site amplification increases with frequencies and is sensitive to near-station structures [e.g., *Edwards et al.*, 2008; *Edwards and Rietbrock*, 2009]. We assume that for the path length mostly greater than 150 km in this study ignoring site effects does not bias the determination of the path effect.

[26] Errors may be introduced by the assumption of point source invoked to ignore source directivity. The directivity effect has been assumed to be averaged out either with a wide azimuthal coverage of stations in teleseismic studies [e.g., *Allmann and Shearer*, 2009] or through the diversity of focal mechanism in mantle wedge models [*Stachnik et al.*, 2004]. This merely reflects the difficulty in identifying the fault plane of all the events used. In Appendix A, section A2

we conduct a first order evaluation of how directivity effects could bias the measurement of f_c in this study. We show that for an intraslab event with stations widely dispersed above it, the directivity effect is small for both the horizontal fault plane and the vertical fault plane with a downward rupture, but culminates for the vertical fault plane that ruptures toward the surface. The maximum directivity effect reaches 0.4 and 0.9 Hz for events whose sizes are 5 km ($\sim M_w$ 5) [*Geller*, 1976] and 2 km, respectively. These are not small numbers compared with the average of 0.56 Hz error in f_c derived from cross-cluster events in section 5.2. However, one notable feature is that these errors are much smaller than the errors stemming from the tradeoff when f_c is freely determined along with path, i.e., 3.84 Hz as reported in section 5.2, implying that fixing f_c in the inversion comes with a smaller cost.

[27] Another first order feature is that the directivity effect is proportional to f_c, and thus it cannot be ruled out through event selection [*Wang and Rubin*, 2011]. However, the actual error must be smaller than that presented here, because (1) narrow pulses are susceptible to attenuation more strongly than wide pulses, naturally suppressing the most severe bias from upward rupture; and (2) the enhanced robustness of the CEM, along with the multievents in each cluster, likely average out large errors associated with individual events. Nonetheless, directivity effects cannot be completely eliminated and may ultimately contribute to the errors in both f_c and Q.

[28] The assumption of $\alpha = 0.27$ in equation (4) could also introduce errors. To evaluate this assumption, we repeated the synthetic experiments in Figure 5 with $\alpha = 0$ for the input synthetic spectra and inverted them with $\alpha = 0.1$ and 0.27. We found that the maximum error in f_c is <0.5 Hz. As described in section 2, it is difficult to invert for α as an additional independent parameter. We adopt the value widely used in studies of mantle wedge [e.g., *Stachnik et al.* 2004; *Rychert et al.*, 2008; *Pozgay et al.*, 2009]. The difference in Q for different α is systematic and the relative pattern remains consistent.

6.2. Estimation of Scalar Moment

[29] Other than the spectral falloff which is characterized by (t^*, f_c), the base-line level of the low-frequency spectral amplitude, or Ω_0, is another source parameter to be determined. The seismic moment is related to Ω_0 via equation (2) [*Brune*, 1970]. In this work, v_p and ρ were assumed to be 8 km/s and 3.3 g/cm^3, respectively, for events at 100 km depth based on the Preliminary Reference Earth Model (PREM) [*Dziewonski and Anderson*, 1981].

[30] We calculate seismic moments for all stations and averaged them for each event. This estimated average seismic moment, $\overline{M_0}$, is compared with the cataloged M_0 provided by NIED CMT solutions. The two sets of seismic moment are highly correlated with a correlation coefficient of 0.98 and the best fit slop of 0.95. The small discrepancy between $\overline{M_0}$ and M_0 may be attributed to several factors. The catalog M_0 was determined with surface waves in bandwidths 20–100 s, while the bandwidth in our spectrum analysis is predominantly 0.5–8 Hz. In addition, equation (2) employs simplified assumptions in radiation pattern and geometrical spreading.

6.3. Scaling Law Between Corner Frequency and Seismic Moment

[31] For each f_c determined by CEM, we computed the radius of the circular source r_c according to *Madariaga* [1976],

$$r_c = 0.32\frac{v_s}{f_c}, \tag{7}$$

where v_s is the near-source shear wave velocity and 0.32 is the coefficient for P wave. The stress drop $\Delta\sigma$ can be estimated from the seismic moment and the source radius [*Eshelby*, 1957]

$$\Delta\sigma = \frac{7M_0}{16r_c^3}. \tag{8}$$

Using v_s = 4.5 km/s as the near-source velocity for ~100 km depth events, the estimated stress drops spread mostly over 1–100 MPa with a median at 22.5 MPa (Figure 10a).

[32] Substituting (7) into (8) to eliminate r_c, the relation between corner frequency and seismic moment reads

$$f_c = 0.42v_s\left[\frac{\Delta\sigma}{M_0}\right]^{1/q.} \tag{9}$$

[33] We fit equation (9) to the measured f_c in this study and the cataloged M_0 and obtained a best fit $\Delta\sigma$ of 21.9 ± 6.9 MPa (one standard error) with a variance reduction of 86%. In equation (9) the exponent $1/q$ with q explicitly set at 3 is consistent with a self-similarity scenario. If q is allowed to change with stress drops bracketed between 1 and 100 MPa, the best fit q is 3.1 with a variance reduction of 85%. As the deviation from 3 is statistically insignificant, we infer that the data are compatible with the theory, implying a self-similarity for intraslab earthquakes at ~100 km depths over M_w 3.5–6 or 4 orders of magnitude in seismic moment.

[34] A number of studies have verified self-similarity of earthquake across a certain range of magnitude [e.g., *Ide et al.*, 2003; *Prieto et al.*, 2004]. Others question it as a property for all earthquake sizes [e.g., *Tajima and Tajima*, 2007; *Mayeda et al.*, 2007]. *Allmann and Shearer* [2009] compiled a data set consisting of most measurements up-to-date (referred to as "global" data set) and suggested the presence of self-similarity over 13 orders of magnitude in seismic moment. We modified this data set by converting some of the f_c's that were originally measured from S-wave spectra to the P-wave f_c's by multiplying a factor of 1.5 [e.g., *Madariaga*, 1976; *Pozgay et al.*, 2009]. In this global data set, events are shallower than 60 km depths and skewed toward shallow crust above 20 km. For v_s = 3.9 km/s representing the upper crust [*Allmann and Shearer*, 2009], the median stress drop is 3.3 MPa. We also fit equation (10) to these data and obtained a best fit $\Delta\sigma$ = 1.4 ± 1.1 MPa and a variance reduction of 81%. Our data merge concordantly with this data set except with higher corner frequencies and higher stress drops (Figure 10b).

[35] Because biases from different methodologies exist, a cross-study comparison at present should be perceived as a proof of concept exercise before a comprehensive, equally based comparison is conducted in the future. We compare the best fit $\Delta\sigma$ constrained by the similarity relationship

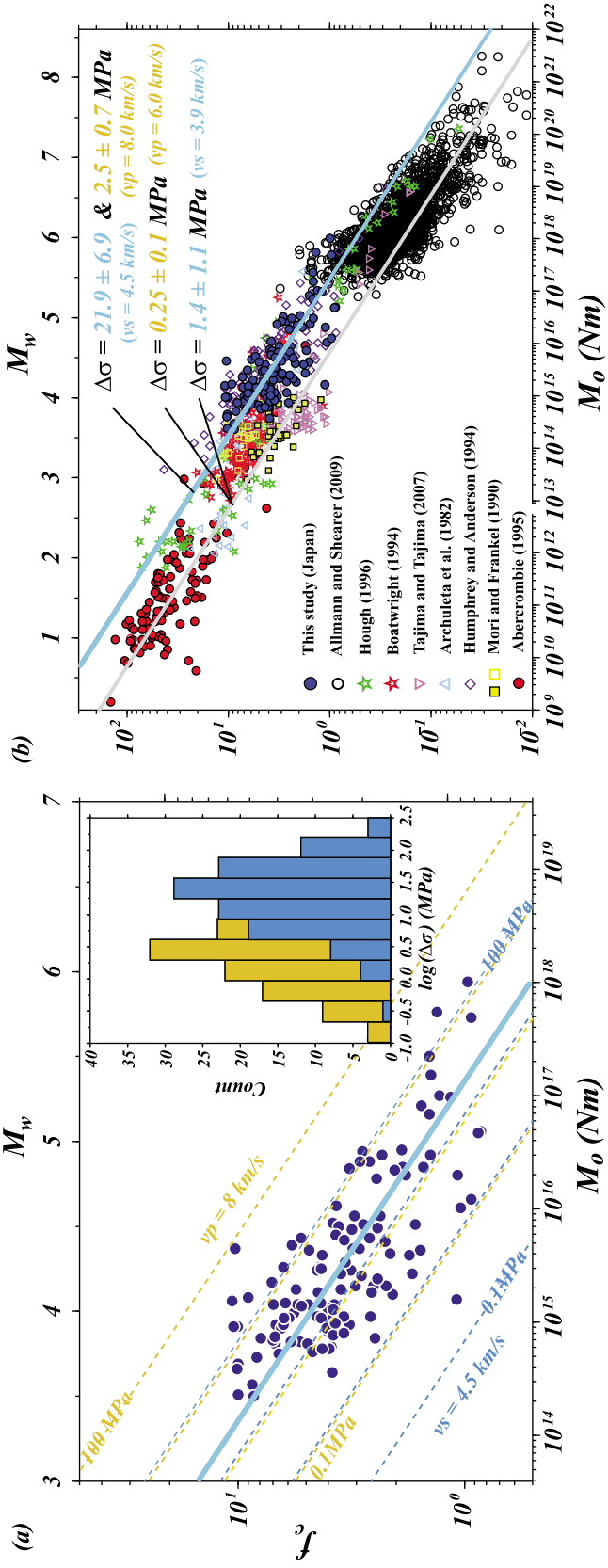

Figure 10. (a) Measured f_c of this study against M_0 (blue dots) fitted by the theoretical relationship (blue thick line) with the best fit $\Delta\sigma$ of 21.9 $\pm$ 6.9 MPa based on *Madariaga*'s [1976] source model. The best fit HW model with $\Delta\sigma$ of 2.5 $\pm$ 0.8 MPa occupies nearly the same line. Background dotted lines represent $\Delta\sigma$ of 0.1, 1, 10, and 100 MPa for *Madariaga* [1976] (blue) and HW (golden). Inset is the histogram for $\Delta\sigma$ with the same color code. (b) Merging of our data with a global data set (sources shown in the lower-left) over much wider data ranges. The blue line is the same as in Figure 10a, and the gray line correspond to best fit theoretical relationships to the global data with $\Delta\sigma$ of 1.4 $\pm$ 1.1 MPa for *Madariaga* [1976] and 0.25 $\pm$ 0.1 MPa for HW. The $\Delta\sigma$ annotating the lines follow the same color code. In the global data set, the higher values of *Mori and Frankel* [1990] (yellow open square) are estimated from the HW model (see text). The difference in $\Delta\sigma$ between this study and the global data set is much larger for *Madariaga*'s [1976] model than for HW although it is statistically significant for both.

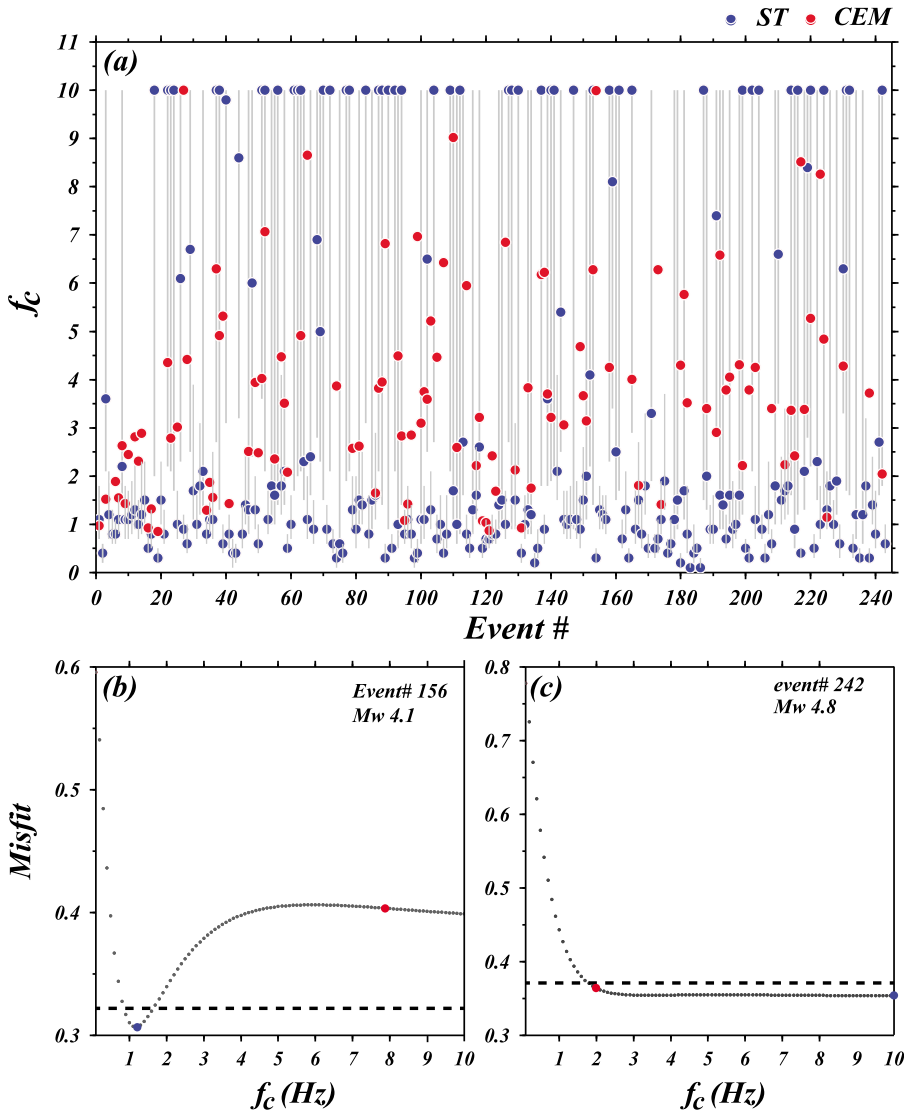

Figure A1. (a) Comparison in f_c between the CEM (red dot) and the ST method (blue dot) for all events used in this study. The shaded vertical lines denote the uncertainty range for the f_c determined by the ST method. (b and c) Examples of misfit curves for the ST method. In Figure A1b, for an *Mw* 4.1 event, the best fitting f_c is tied at 1.15 Hz, in contrast to the f_c of 7.9 Hz by the CEM. Figure A1c shows the typical misfit curve with the minimum at the upper limit of the frequency range. The CEM yields an f_c of about 2 Hz for this *Mw* 4.8 event.

equation (9) rather than the statistics directly from $\Delta\sigma$ estimates through equations (7) and (8), although the conclusions would be similar. We tested the null hypothesis that the two stress drop estimates, 21.9 $\pm$ 6.9 MPa and 1.4 $\pm$ 1.1 MPa, are the same, and found that the null hypothesis can be rejected at the 99% confidence level by the *t*-test. The alternative hypothesis that the two estimates are different implies that the faults of the intraslab earthquakes analyzed in this study break at higher stress level than that in the shallow crust. Although not for intraslab events, the depth-dependence of $\Delta\sigma$ for crustal events has been documented. *Hardebeck and Aron* [2009] reported a dramatic increase of $\Delta\sigma$ from 5 MPa near the surface to 50 MPa in crust deeper than 13 km in the vicinity of the Hayward fault, California. *Oth et al.* [2010] showed median $\Delta\sigma$ of 1.1 and 9.2 MPa for events shallower and deeper than 30 km, respectively, from an analysis of borehole acceleration spectra in Japan. Statistics were not explicitly given in these studies. The latter used the source formulation of *Hanks and Wyss* [1972] that entails smaller stress drops as described below.

[36] Another source dimension formulation commonly employed was proposed by *Hanks and Wyss* [1972, hereinafter HW]:

$$r_c = \frac{2.34 v_p}{2\pi f_c}, \tag{10}$$

which is an extension from *Brune*'s [1970, 1971] original rupture model for *S* wave. The larger product of near-source velocity v_p and the constant (0.37) in equation (10) than in equation (7) renders a larger source dimension and consequently a smaller $\Delta\sigma$ than for *Madariaga* [1976]. The

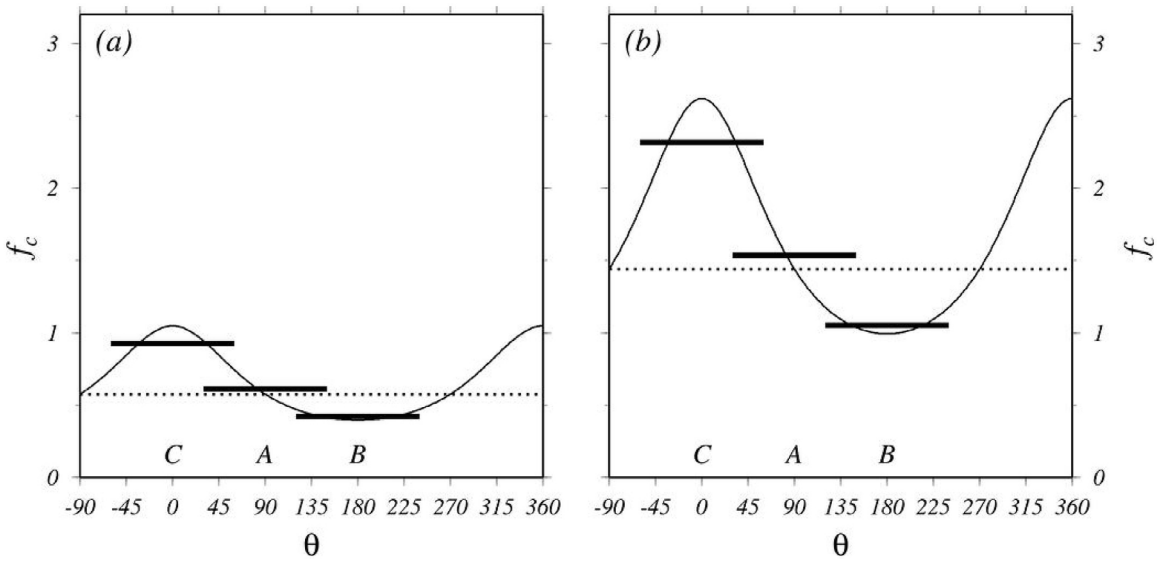

Figure A2. Plots of f_c versus takeoff angle θ (from rupture direction). The lengths of the fault are (a) 5 and (b) 2 km. Thin curves are apparent f_c due to directivity. Dotted lines are the "true" f_c without directivity observed at $\theta = 90^o$. Thick lines are average f_c's over the respective θ ranges for cases A, B, and C. The difference in f_c between $\theta = 0^o$ and the dotted line (for $\theta = 90^o$) characterizes the maximum directivity effect. Note the directivity effect is proportional to f_c or $1/L$ as well.

corresponding $f_c - M_0$ theoretical relationship takes different constants:

$$f_c = 0.49 v_p \left[\frac{\Delta\sigma}{M_0}\right]^{1/3.} . \quad (11)$$

[37] Fitting equation (11) to our data yields a $\Delta\sigma$ of 2.5 ± 0.7 MPa for v_p = 8 km/s. To apply the HW model to the global data set, we convert the r_c's reported by *Mori and Frankel* [1990] to f_c's using equation (10). The fitting to the global data gives $\Delta\sigma = 0.25 \pm 0.1$ MPa with $v_p = 6$ km/s. Both fitting reduce the data variance by about 85%. The gap in $\Delta\sigma$ is now substantially narrowed from that for *Madariaga* [1976] and the hypothesis that the two estimates are equal cannot be rejected at any high confidence level.

[38] The HW model reconciles the apparent depth-dependent f_c by the more depth-sensitive v_p and a larger constant, together suppressing the variation in $\Delta\sigma$. This effect has not caught much attention because so far most investigations have targeted crustal events for which a constant near-source velocity usually suffices. Both *Madariaga*'s [1976] and the HW models are appealing in source studies, and distinguishing which quantifies the fault strength more appropriately is beyond the current study. It is implied that a conclusion on whether intermediate-depth, intraslab earthquakes are "stronger" than shallow crustal events awaits not only more sampling of deep events and more vigorous analysis of measurement errors but also an appraisal of source models.

7. Conclusions

[39] The cluster-event method uses an improved inversion scheme to simultaneously determine the corner frequency of the source and the attenuation of the medium. We demonstrate that distinguishing between these two parameters is difficult in the presence of noise unless the number of unknowns is sufficiently reduced. The CEM bundles nearby events into cluster and solve for Q on a cluster-station basis, while imposing single-f_c to each event of the cluster, largely reducing the degrees of freedom of the inversion problem. The cross-cluster comparison in f_c proves that the CEM yields a robust constraint on f_c, which assures equivalently a reliable determination of Q. Applying CEM to the Japan subduction zone depicts a characteristic variation in Q from mostly lower than 300 in the arc and back-arc to >600 in the fore-arc. This will form a basis for detailed mapping of subduction zone structures. The resolved f_c's fit well the relationship with seismic moment in the form of $M_0 \propto f_c^{-3}$, suggesting self-similarity of intermediate-depth subduction zone events. We report stress drop estimated from f_c for our data and a data set composed of prior measurements. We found that the stress drops of this study are significantly higher than previously estimated for crustal events based on *Madariaga*'s [1976] formulation, while the difference is much suppressed if *Hanks and Wyss*'s [1972] model is used. We consider the results shown here as motivation for studies of earthquake sources at greater depths of subduction zones and in different tectonic settings.

Appendix A

A1. ST Method

[40] We applied the ST method to the same data set and show the resultant f_c in Figure A1. There are some f_c's that are constrained at a local valley of the misfit curve, although the solution could be significantly different from that of

CEM. However, because the constraint is half-sided by nature, the minimum misfit usually occurs at the upper limit of the frequency range used.

A2. Directivity

[41] We estimate the takeoff-angle dependence of f_c for events with rupture lengths (L) of 5 and 2 km. L = 5 km roughly corresponds to M_w of 5 [*Geller*, 1976]. The apparent rupture duration τ_a is determined by

$$\tau_a = \frac{L}{v_r} - \left(\frac{L\cos\theta}{v_p}\right) \tag{A1}$$

where v_p is set at 8 km/s, v_r is the rupture velocity assumed to be $0.8v_s$ with v_s 4.5 km/s, and θ is the angle between the rupture direction and the takeoff vector. For simplicity, we assume a 2 dimensional problem with the source and the stations in a vertical section to which the fault plane is perpendicular and that all the stations are above the source with raypaths approximated by straight lines. Half of the τ_a multiplied by v_r is assumed as the radius of the circular fault [*Boatwright*, 1980; *Mori and Frankel*, 1990],

$$r_c = \frac{L}{2}\left(1 - \frac{v_r\cos\theta}{v_p}\right) \tag{A2}$$

which is then converted to f_c using equation (7). (Equation A2) implies that directivity effect gives rise to variable apparent fault size and f_c. Figure A2 shows f_c as a function of θ. We assume that the stations form a 120° wide observation aperture above the source and that the average f_c over this aperture is the CEM solution. We tested 3 types of fault geometry. A: Horizontal fault plane rupturing toward right. In this case, the stations are located in the θ range of (30–150°). B: Vertical fault plane rupturing downward with the station θ range of (120–240°). C: Vertical fault plane rupturing upward with the station θ range of (300–60°). The true f_c of the source can be obtained at θ = 90° where directivity is absent. For L = 5 km the deviations of the measured f_c from the true f_c are 0.04, −0.15, and 0.35 Hz for case A, B, and C, respectively. For L = 2 km, the deviations for the 3 cases are 0.1, −0.4, and 0.9 Hz, respectively. The fact that the directivity effect in f_c is magnified at smaller events can be illustrated by comparing cases at θ = 0 and 90°, i.e., $f_c(\theta = 0) \propto (2/L)/(1 - v_r/v_p)$, and $f_c(\theta = 90) \propto (2/L)$. The difference between them is the maximum directivity effect

$$\delta f_c \propto \frac{2}{L}\left(\frac{v_r}{v_p} - O^2\right) \tag{A3}$$

where O^2 represents a second order term much smaller than (v_r/v_p). It is noted that not only f_c itself but also the magnitude of the directivity effect in f_c are inversely proportional to the fault size.

[42] **Acknowledgments.** The authors thank Rachel E. Abercrombie and an anonymous reviewer for their critical questions and constructive comments, which much improved the revised manuscript. The broadband seismic records were provided by F-net operated by National Research Institute for Earth Science. The work was supported by National Science Council of Taiwan under grant NSC 100-2627-E-001-001.

References

Aarts, E., and J. Korst (1989), *Simulated Annealing and Boltzmann Machine: A Stochastic Approach to Combinatorial Optimization and Neural Computing*, John Wiley, Chichester, U. K.

Abercrombie, R. E. (1995), Earthquake source scaling relationships from −1 to 5 M_L using seismograms recorded at 2.5-km depth, *J. Geophys. Res.*, *100*, 24,015–24,036, doi:10.1029/95JB02397.

Abercrombie, R. E., and J. R. Rice (2005), Can observations of earthquake scaling constrain slip weakening?, *Geophys. J. Int.*, *162*, 406–424, doi:10.1111/j.1365-246X.2005.02579.x.

Aki, K., and P. G. Richards (1980), *Quantitative Seismology*, 932 pp., W. H. Freeman, New York.

Allmann, B. P., and P. M. Shearer (2009), Global variations of stress drop for moderate to large earthquakes, *J. Geophys. Res.*, *114*, B01310, doi:10.1029/2008JB005821.

Anderson, J. G., and S. E. Hough (1984), A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, *Bull. Seismol. Soc. Am.*, *74*(5), 1969–1993.

Anderson, J. G., and J. R. Humphrey Jr. (1991), A least squares method for objective determination of earthquake source parameters, *Seismol. Res. Lett.*, *62*, 201–209.

Boatwright, J. (1980), A spectral theory for circular seismic sources; simple estimates of source dimension, dynamic stress drop, and radiated seismic energy, *Bull. Seismol. Soc. Am.*, *70*, 1–27.

Brune, J. N. (1970), Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.*, *75*, 4997–5009, doi:10.1029/JB075i026p04997.

Brune, J. N. (1971), Correction to "Tectonic stress and the spectra of seismic shear waves from earthquakes," *J. Geophys. Res.*, *76*, 5002.

Chou, H.-C., B.-Y. Kuo, L.-Y. Chiao, D. Zhao, and S.-H. Hung (2009), Tomography of the westernmost Ryukyu subduction zone and the serpentinization of the fore-arc mantle, *J. Geophys. Res.*, *114*, B12301, doi:10.1029/2008JB006192.

Dziewonski, A. M., and D. L. Anderson (1981), Preliminary reference Earth model, *Phys. Earth Planet. Inter.*, *25*, 297–356, doi:10.1016/0031-9201(81)90046-7.

Edwards, B., and A. Rietbrock (2009), A comparative study on attenuation and source-scaling relations in the Kantō, Tokai, and Chubu region of Japan, using data from Hi-Net and KiK-Net, *Bull. Seismol. Soc. Am.*, *99*, 2435–2460, doi:10.1785/0120080292.

Edwards, B., A. Rietbrock, J. J. Bommmer, and B. Baptie (2008), The acquisition of source, path, and site effects from microearthquake recordings using Q tomography: Application to the United Kingdom, *Bull. Seismol. Soc. Am.*, *98*, 1915–1935, doi:10.1785/0120070127.

Eshelby, J. D. (1957), The determination of the elastic field of an ellipsoidal inclusion, and related problems, *Proc. R. Soc. London, Ser. A*, *241*, 376–396, doi:10.1098/rspa.1957.0133.

Geller, R. J. (1976), Scaling relations for earthquake source parameters and magnitudes, *Bull. Seismol. Soc. Am.*, *66*, 1501–1523.

Hanks, T. C., and M. Wyss (1972), The use of body-wave spectra in the determination of seismic-source parameters, *Bull. Seismol. Soc. Am.*, *62*, 561–589.

Hardebeck, J. L., and A. Aron (2009), Earthquake stress drops and inferred fault strength on the Hayward fault, east San Francisco Bay, California, *Bull. Seismol. Soc. Am.*, *99*, 1801–1814, doi:10.1785/0120080242.

Ide, S., G. C. Beroza, S. G. Prejean, and W. L. Ellsworth (2003), Apparent break in earthquake scaling due to path and site effects on deep borehole recordings, *J. Geophys. Res.*, *108*(B5), 2271, doi:10.1029/2001JB001617.

Jackson, I., J. D. Fitz Gerald, U. H. Faul, and B. H. Tan (2002), Grain-size-sensitive seismic wave attenuation in polycrystalline olivine, *J. Geophys. Res.*, *107*(B12), 2360, doi:10.1029/2001JB001225.

Jackson, I., U. H. Faul, J. D. Fitz Gerald, and S. J. S. Morris (2006), Contrasting viscoelastic behavior of melt-free and melt-bearing olivine: Implications for the nature of grain-boundary sliding, *Mater. Sci. Eng. A*, *442*, 170–174, doi:10.1016/j.msea.2006.01.136.

Karato, S. (2003), Mapping water content in the upper mantle, in *Inside the Subduction Factory*, *Geophys. Monogr. Ser*, vol. 138, edited by J. Eiler, pp. 135–152, AGU, Washington, D. C., doi:10.1029/138GM08.

Kirkpatrick, S., C. D. Gelatt Jr., and M. P. Vecchi (1983), Optimization by simulated annealing, *Science*, *220*(4598), 671–680, doi:10.1126/science.220.4598.671.

Kneller, E. A., and P. E. van Keken (2008), Effect of three-dimensional slab geometry on deformation in the mantle wedge: Implications for shear wave anisotropy, *Geochem. Geophys. Geosyst.*, *9*, Q01003, doi:10.1029/2007GC001677.

Lin, S.-C., B.-Y. Kuo, and S.-L. Chung (2010), Thermomechanical models for the dynamics and melting processes in the Mariana subduction system, *J. Geophys. Res.*, *115*, B12403, doi:10.1029/2010JB007658.

Madariaga, R. (1976), Dynamics of an expanding circular fault, *Bull. Seismol. Soc. Am.*, *66*, 639–666.

Mayeda, K., L. Malagnini, and W. R. Walter (2007), A new spectral ratio method using narrow band coda envelopes: Evidence for non-self-similarity in the Hector Mine sequence, *Geophys. Res. Lett.*, *34*, L11303, doi:10.1029/2007GL030041.

Mori, J., and A. Frankel (1990), Source parameters for small events associated with the 1986 North Palm Springs, California, earthquake determined using empirical Green functions, *Bull. Seismol. Soc. Am.*, *80*, 278–295.

Mueller, C. S. (1985), Source pulse enhancement by deconvolution of an empirical Green's function, *Geophys. Res. Lett.*, *12*, 33–36, doi:10.1029/GL012i001p00033.

Okada, Y., K. Kasahara, S. Hori, K. Obara, S. Sekiguchi, H. Fujiwara, and A. Yamamoto (2004), Recent progress of seismic observation networks in Japan Hi-net, F-net, K-NET and KiK-net, *Earth Planets Space*, *56*, xv–xxviii.

Oth, A., D. Bindi, S. Parolai, and D. Di Giacomo (2010), Earthquake scaling characteristics and the scale-(in)dependence of seismic energy-to-moment ratio: Insights from KiK-net data in Japan, *Geophys. Res. Lett.*, *37*, L19304, doi:10.1029/2010GL044572.

Pozgay, S. H., D. A. Wiens, J. A. Conder, H. Shiobara, and H. Sugioka (2009), Seismic attenuation tomography of the Mariana subduction system: Implications for thermal structure, volatile distribution, and slow spreading dynamics, *Geochem. Geophys. Geosyst.*, *10*, Q04X05, doi:10.1029/2008GC002313.

Prieto, G. A., P. M. Shearer, F. L. Vernon, and D. Kilb (2004), Earthquake source scaling and self-similarity estimation from stacking *P* and *S* spectra, *J. Geophys. Res.*, *109*, B08310, doi:10.1029/2004JB003084.

Roth, E. G., D. A. Wiens, L. M. Dorman, J. Hildebrand, and S. C. Webb (1999), Seismic attenuation tomography of the Tonga-Fiji region using phase pair methods, *J. Geophys. Res.*, *104*, 4795–4809, doi:10.1029/1998JB900052.

Rychert, C. A., K. M. Fischer, G. A. Abers, T. Plank, E. Syracuse, J. M. Protti, V. Gonzalez, and W. Strauch (2008), Strong along-arc variations in attenuation in the mantle wedge beneath Costa Rica and Nicaragua, *Geochem. Geophys. Geosyst.*, *9*, Q10S10, doi:10.1029/2008GC002040.

Schurr, B., G. Asch, A. Rietbrock, R. Trumbull, and C. Haberland (2003), Complex patterns of fluid and melt transport in the central Andean subduction zone revealed by attenuation tomography, *Earth Planet. Sci. Lett.*, *215*, 105–119, doi:10.1016/S0012-821X(03)00441-2.

Shearer, P. M., G. A. Prieto, and E. Hauksson (2006), Comprehensive analysis of earthquake source spectra in Southern California, *J. Geophys. Res.*, *111*, B06303, doi:10.1029/2005JB003979.

Shito, A., S. Karato, and J. Park (2004), Frequency dependence of *Q* in Earth's upper mantle inferred from continuous spectra of body waves, *Geophys. Res. Lett.*, *31*, L12603, doi:10.1029/2004GL019582.

Sonley, E., and R. E. Abercrombie (2006), Effects of methods of attenuation correction on source parameter determination, in *Earthquakes: Radiated Energy and the Physics of Faulting*, *Geophys. Monogr. Ser*, vol. 170, edited by R. E. Abercrombie et al., pp. 91–97, AGU, Washington, D. C., doi:10.1029/170GM11.

Stachnik, J. C., G. A. Abers, and D. H. Christensen (2004), Seismic attenuation and mantle wedge temperatures in the Alaska subduction zone, *J. Geophys. Res.*, *109*, B10304, doi:10.1029/2004JB003018.

Sugimoto, T., T. Shibata, M. Yoshikawa, and K. Takemura (2006), Sr-Nd-Pb isotopic and major and trace element compositions of the Yufu-Tsurumi volcanic rocks: Implications for the magma genesis of the Yufu-Tsurumi volcanoes, northeast Kyushu, Japan, *J. Mineral. Petrol. Sci.*, *101*, 270–275, doi:10.2465/jmps.101.270.

Tajima, R., and F. Tajima (2007), Seismic scaling relations and aftershock activity from the sequences of the 2004 mid Niigata and the 2005 west off Fukuoka earthquakes (M_w 6.6) in Japan, *J. Geophys. Res.*, *112*, B10302, doi:10.1029/2007JB004941.

Tsumura, N., S. Matsumoto, S. Horiuchi, and A. Hasegawa (2000), Three-dimensional attenuation structure beneath the northeastern Japan arc estimated from spectra of small earthquakes, *Tectonophysics*, *319*, 241–260, doi:10.1016/S0040-1951(99)00297-8.

Viegas, G., R. E. Abercrombie, and W.-Y. Kim (2010), The 2002 M5 Au Sable Forks, NY, earthquake sequence: Source scaling relationships and energy budget, *J. Geophys. Res.*, *115*, B07310, doi:10.1029/2009JB006799.

Wang, E., and A. M. Rubin (2011), Rupture directivity of microearthquakes on the San Andreas Fault from spectral ratio inversion, *Geophys. J. Int.*, *186*, 852–866, doi:10.1111/j.1365-246X.2011.05087.x.

Wang, Y.-J., K.-F. Ma, F. Mouthereau, and D. Eberhart-Phillips (2010), Three-dimensional *Qp*- and *Qs*-tomography beneath Taiwan orogenic belt: Implications for tectonic and thermal structure, *Geophys. J. Int.*, *180*, 891–910, doi:10.1111/j.1365-246X.2009.04459.x.

S.-H. Hung, Department of Geosciences, National Taiwan University, 1 Roosevelt Rd., Sec. 4, Taipei 106, Taiwan.

Y.-T. Ko and B.-Y. Kuo, Institute of Earth Sciences, Academia Sinica, 128 Academia Rd., Sec. 2, Taipei 115, Taiwan. (byk@earth.sinica.edu.tw)