JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 112, B06311, doi:10.1029/2006JB004381, 2007

Imaging of crustal heterogeneous structures using a slowness-weighted back-projection with effects of scattering modes: 1. Theory

T. Taira[1,2] and K. Yomogida[1]

Received 6 March 2006; revised 11 November 2006; accepted 12 February 2007; published 27 June 2007.

[1] This is the first paper in a two-part series on a newly developed imaging approach for small-scale heterogeneities (<1 km) in the crust with effects of scattering modes. To obtain a reliable crustal heterogeneous structure, we follow the six major steps: (1) removing overall complex propagation effects, including anelastic attenuation, by using a statistical technique with the use of the Akaike Information Criterion (AIC), (2) obtaining high-resolution frequency-wave number (*f-k*) power spectra and slowness vectors of spectral peaks in the time-frequency domain, based on a stationary autoregressive model, (3) estimating polarization vectors of the scattered waves identified in step 2 with a stationary multivariate autoregressive model, (4) determining scattering modes (i.e., *P* or *S* wave arrival) from the angle between the slowness and polarization vectors obtained in steps 2 and 3, respectively, (5) correcting effects of seismic-source radiation and surface geology by a coda-normalization approach, and finally (6) mapping the *f-k* power spectra into small blocks in a model space as scattering coefficients, using a slowness-weighted back-projection. We can incorporate scattering modes as well as propagation effects such as anelastic attenuation factors in the background medium, with the AIC based amplitude recovery technique. The resolution in *f-k* spectrograms and the accuracy of polarization estimates are significantly improved through the present approach, so that not only more scattered phases are clearly identified but also their spatial three-dimensional locations are pinpointed more precisely and stably than previous approaches in imaging based on scattering theory.

Citation: Taira, T., and K. Yomogida (2007), Imaging of crustal heterogeneous structures using a slowness-weighted back-projection with effects of scattering modes: 1. Theory, *J. Geophys. Res.*, *112*, B06311, doi:10.1029/2006JB004381.

1. Introduction

[2] The Earth has been found to be heterogeneities in a broad range of scales [*Wu and Aki*, 1988; *Sato and Fehler*, 1998]. Seismic coda waves (hereafter called simply "coda waves") are considered to be scattered waves due to small-scale heterogeneities distributed randomly and uniformly [*Aki*, 1969; *Aki and Chouet*, 1975].

[3] Over the past several decades, a considerable number of studies have been made to develop imaging approaches for small-scale heterogeneities, using by teleseismic [e.g., *Gupta et al.*, 1990; *Revenaugh*, 1995; *Scherbaum et al.*, 1997; *Bostock et al.*, 2001], regional [e.g., *Nishigami*, 1991; *Chávez-Pérez and Louie*, 1998; *Asano and Hasegawa*, 2004; *Taira and Yomogida*, 2004], and local records [e.g., *Spudich and Bostwick*, 1987; *Dodge and Beroza*, 1997] for earthquakes (passive-sources) since the pioneering work of *Aki* [1969]. Several novel migration methods have also been introduced by using active-sources in exploration seismology [e.g., *Matsumoto et al.*, 1998; *Huang et al.*, 1999a, 1999b], as summarized by *Fehler and Huang* [2002] and *Wu* [2003].

[4] Several seismic array analyses of coda waves have recently revealed that *P*- and *S*-coda waves contain coherent scattered waves from specific seismic scatterers distributed nonuniformly. *Wagner* [1997] analyzed seismograms for regional earthquakes recorded by the broadband seismic array located at the Piñion Flats Observatory, southern California, in order to reveal the nature of *P*-, *S*- and *Lg*-coda waves. He concluded that the observed *P*-, *S*-, and *Lg*-coda waves in the region are mainly composed of coherent forward scattered waves.

[5] Once we identify coherent arrival phases in coda waves, a simple linear inversion scheme can be generally derived between seismic data and strengths as well as locations of seismic scatterers distributed in a medium, based on the single isotropic scattering model [*Sato*, 1977]. These strengths may be called scattering coeffi-

[1]Division of Earth and Planetary Sciences, Graduate School of Science, Hokkaido University, Sapporo, Japan.
[2]Now at Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, D. C., USA.

0148-0227/07/2006JB004381$09.00

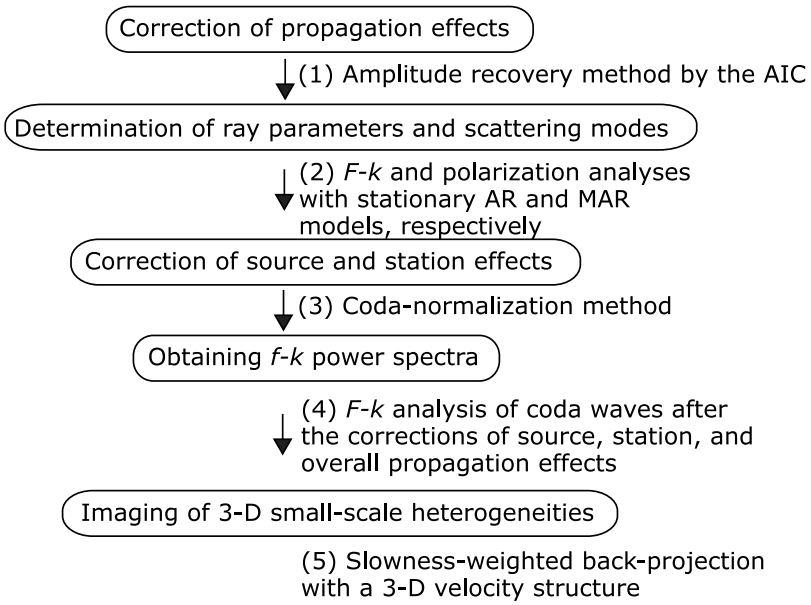

Figure 1. Flowchart of data processing steps in this study. To correct source and station effects, including the wave-type (i.e., *P* or *S* wave) of each coherent scattered wave in coda waves, we need to evaluate scattering modes in the previous step. We conduct the *f-k* analysis twice so that the former in step 2 and latter in step 4 can determine scattering modes and *f-k* power spectra, respectively.

cients or scattering potentials, depending on assumed scattering models. Unlike relatively low frequency seismic waves (<1 Hz) used for now popular seismic travel time tomographic inversion studies, we must carefully take into consideration complex behaviors of both effects of seismic-source radiation (hereafter called "source effect") and surface geology (hereafter called "station effect") on coda waves for imaging a spatial distribution of seismic scatterers in the above scheme. To minimize these effects, most of the previous studies on imaging of small-scale heterogeneities normalized the amplitudes of the observed coda waves by a reference amplitude in the very beginning of their analyses. This "normalization" leads them to obtain the distribution of scattering coefficients or potentials as not "absolute" but "relative" values.

[6] In this study, we try to map the absolute power spectrum of each coda wave into a scattering coefficient after corrections of the source, station, and propagation (e.g., anelastic attenuation) effects, followed by applications of a coda-normalization approach and an amplitude recovery with the use of the Akaike Information Criterion (AIC) based on a stationary autoregressive (AR) model. The idea of both the AIC and a stationary AR model is also applied to the array processing of waveform data, so that the spatial resolution of three-dimensional distribution of scattering coefficients is much improved, compared with the previous studies. This idea also enables us to discuss their frequency-dependent characteristics in a quantitative manner.

[7] In addition to the above corrections, we conduct the frequency-wave number (*f-k*) and polarization analyses with a stationary AR model of three-component (3-C) seismic array data sets, in order to determine scattering modes of scattered phases, based on their slowness and polarization vectors. We shall particularly emphasize the importance of such a signal processing because of the enhancement of the detectability of relatively weak scattered waves. Although 3-C arrays require much more cost in field and a large amount of data processings, we shall show the importance of analysis techniques for a 3-C array in the imaging studies of coda waves.

[8] In the present paper, called Paper I, we shall explain our new imaging approach for three-dimensional distributions of scattering coefficients, using high-frequency (>1 Hz) coda waves. As an example, we shall employ the seismic data sets from the dense array observations in the Nagamachi-Rifu fault area, northeastern Japan [e.g., *Hasegawa et al.*, 2001]. The concrete results of these data sets with the approach in this paper will be given in the other paper, called Paper II [*Taira et al.*, 2007], including their seismological as well as tectonic implications.

[9] Let us outline our imaging approach here. Figure 1 shows the flowchart of data processing steps of the present approach. First, we remove overall propagation effects, based on amplitude recovery with the AIC. Next, we determine ray parameters and scattering modes (i.e., *P-P* or *P-S*) of coda waves, using *f-k* and polarization analyses with stationary AR and MAR models, respectively. Thirdly, source and station effects, including frequency dependency and scattering modes, are corrected on the basis of a coda-normalization approach. Before estimating scattering coefficients, we once again conduct the *f-k* analysis of coda waves, after the corrections of source, station, and overall propagation effects, in order to obtain *f-k* power spectra related only to scattering coefficients. Finally, we map these *f-k* power spectra into scattering coefficients, and obtain the three-dimensional distribution of scattering coefficients for each scattering mode at each frequency. As summarized in Figure 1, the present approach requires five

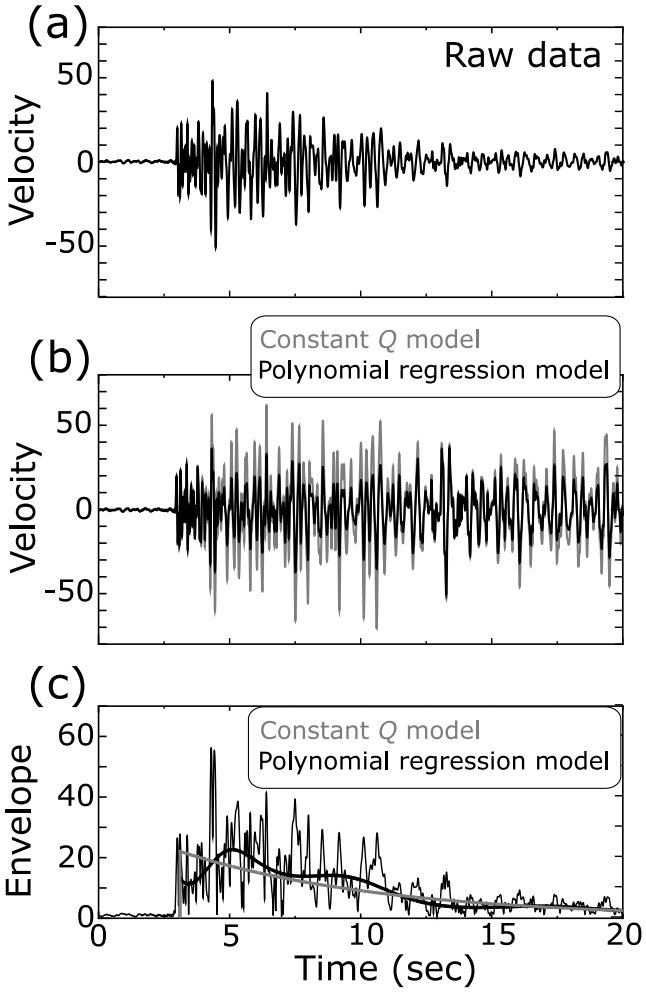

Figure 2. An example of results of the amplitude recovery with the AIC. (a) A raw seismogram and (b) amplitude recovered seismograms by using a constant Q model (gray) and an optimal polynomial regression model (black). The constant Q model corresponds to a first-order polynomial regression model. The order of the optimal polynomial regression model is determined so as to minimize AIC values. In this example, the order takes 12. (c) The envelope of the raw seismogram and the decay curves determined on the basis of the constant Q model (gray) and the optimal polynomial regression model (black).

phases. We shall explain the approach in the subsequent sections: (1) amplitude recovery, (2) correction of source and station effects, (3) determination of ray parameters and scattering modes, and (4) mapping of small-scale heterogeneities.

2. Amplitude Recovery With the AIC

[10] Although we attempt to image seismic scatterers from the spatial and temporal variations of the observed coda wave amplitudes, the amplitudes with frequency higher than 1 Hz are complexly affected by geometrical spreading of the corresponding wavefront and energy loss due to anelastic processes (or internal friction) along individual seismic raypaths. The precise correction of these effects is essential in a reliable estimation of scattering coefficients. The correction of propagation effects (i.e., amplitude recovery) has been usually made by the least squares fit of a given representation of attenuation in previous studies of both exploration seismology [e.g., *Yilmaz*, 1987] and earthquake seismology, that is, assuming the common value of coda Q [e.g., *Taira and Yomogida*, 2004]. This assumption has been proved to be extremely valid in late part of coda waves (e.g., twice the S-wave travel time) since the discovery of *Aki* [1969], and their incoherency is very high.

[11] We are, however, particularly interested in distributions of scatterers around fault zones (see Paper II) where the constant Q model could not be valid. For example, *Korneev et al.* [2003] estimated a detail 2-D attenuation structure along the Parkfield segment of San Andreas Fault, California, by using fault-zone guided waves, showing that a low-attenuation zone is localized within approximately 5 km southeast from the epicenter of the 1966 Parkfield earthquake and up to 5 km depth from the surface. Figure 2a shows an example of the observed seismograms for an explosion source used in our application (Paper II), recorded near the Nagamachi-Rifu fault, northeastern Japan. Several isolated distinct later phases are noticed with amplitude often larger than the direct P-wave. Although strong seismic reflectors could explain such later phases, an individual attenuation along each raypath can also affect these amplitudes.

[12] In this study, we introduce a statistical amplitude recovery technique to minimize complex propagation effects on observed seismograms, by adopting the AIC in the similar manner to *Matsuoka et al.* [1986].

[13] We first obtain an envelope $E(t)$ of each seismogram, using the Hilbert transform, with the reference starting time of a seismogram t_p,

$$E(t) = \sqrt{f(t)^2 + g(t)^2} : t > t_p, \tag{1}$$

where $f(t)$ indicates the seismogram for the lapse time t after processed by an appropriate bandpass filter and corrected a geometrical spreading factor. t_p is the travel time of the direct P-wave. Here $g(t)$ is advanced in a phase by $\pi/2$ with respect to $f(t)$, and $g(t)$ is expressed by

$$g(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{f(\tau)}{t-\tau} d\tau, \tag{2}$$

where P represents the process of taking Cauchy's principal value.

[14] We next introduce a polynomial regression model to estimate propagation effects on seismograms and then minimize them. Instead of a constant Q model, we assume the envelope of the observed seismograms $E(t)$ is expressed by a linear combination of M-th-order polynomial, and thus the time series of an envelope $E(t_i)$ in a discrete form with t_i can be expressed by

$$E(t_i) = \exp\left(r(0) + r(1)t_i + r(2)t_i^2 + \cdots + r(M)t_i^M + \varepsilon_i\right). \tag{3}$$

[15] Taking the logarithm of equation (3), we obtain

$$\log(E(t_i)) = r(0) + r(1)t_i + r(2)t_i^2 + \cdots + r(M)t_i^M + \varepsilon_i \quad (4)$$

$$= \sum_{m=0}^{M} r(m)t_i^m + \varepsilon_i \qquad (i = 1, 2, \cdots, N), \quad (5)$$

where $r(m)$ (m = 0, 1, $\cdots$, M) and ε_i indicate the regression coefficients and Gaussian white noise at time t_i, respectively. N and M are the number of data (i.e., the length of the analyzed seismogram) and the order of the regression coefficients. To estimate the optimal decay curve, the maximum of M, $M_{\max}$, should be $M_{\max} < (2 \sim 3)\sqrt{N}$, as shown by *Akaike and Kitagawa* [1994] for a general discussion. We applied $M_{\max} = 2.5\sqrt{N}$ to seismograms from the Nagamachi-Rifu fault data to be used in Paper II.

[16] To select an optimal value of M, we use the AIC in terms of a statistical technique [*Akaike*, 1969]. The residual variance and the AIC value for the regression model of the M-th-order equation (5) are defined by

$$\sigma^2(M) = \frac{1}{N}\sum_{i=1}^{N}\left(\log(E(t_i)) - r(0) - r(1)t_i - \cdots - r(M)t_i^M\right)^2 \quad (6)$$

$$\mathrm{AIC}(M) = N\log\left(2\pi\sigma^2(M)\right) + N + 2(M+2). \quad (7)$$

[17] We determine M so as to minimize AIC(M). This choice of M is based on the criterion to balance the trade-off between the fitting of the decay curve to the data and the complexity of the model. The gain value at t is determined as the inverse of the ratio of the decay curve value at t to one at the direct P-wave arrival in each seismogram.

[18] A first-order polynomial (i.e., M = 1) can correspond to a constant Q model, similar to the classical result of coda [e.g., *Aki*, 1969]. We therefore expect that the order of the polynomial regression model is close to 1 if the shape of seismogram envelopes shows a simple exponential decay. The single scattering model with a stationary random medium predicts a simple exponential decay, that is, only $r(0)$ and $r(1)$ of equation (3), as widely known since *Aki and Chouet* [1975]. Although not based on any concrete physical models, it is natural to extend the formulation of the single scattering model into polynomials in exponential for complex random media with heterogeneous intrinsic Q distribution. On the other hand, one of the advantages of the present amplitude recovery technique does not involve the assumption of the common propagation effect to all the earthquake-station pairs, as having been used widely in previous studies on small-scale heterogeneities [e.g., *Sato and Fehler*, 1998].

[19] We performed the amplitude recovery technique to the observed seismogram in Figure 2a. Note that we used the raw seismogram to simply show how the amplitude recovery technique works. Figures 2b and 2c show the amplitude recovered seismogram by applying a polynomial regression model with the AIC and the estimated decay curve. We also computed the decay curve based on a constant Q model, and performed the amplitude recovery by using this decay curve. As mentioned above, we see the large-amplitude later phases on the original seismogram (Figure 2a). The decay curve based on the constant Q model could be too simple for this seismogram, and then could cause an overestimation of the amplitudes of the later phases. On the other hand, the amplitude recovery with a polynomial regression model tends to suppress the amplitudes of later phases, so that scattering intensity in our final result, especially scatterers in a deep region, might be underestimated.

3. Correction of Source and Station Effects

[20] We next consider how to correct source and station effects with the observed high-frequency seismograms. Although most of previous studies adopted the normalization of all seismograms as the first step of data processings, we presently attempt to image scattering coefficients after correction of source and station effects as well as propagation effect. To isolate source, station, and propagation effects of high-frequency seismic waves, direct P- and S-waves have been used in many studies because we can express their effects straightforwardly in theory [e.g., *Iwata and Irikura*, 1988]. In practice, however, abundant and uniformly sampled data in space are required to obtain their reliable estimations. Coda wave, on the other hand, provides a reliable way to estimate them even if limited amounts of data are available, due to its averaging effect over a given model space [e.g., *Aki*, 1969]. For example, *Tsujiura* [1978] showed that station effects with frequency dependency determined by the coda-normalization method are more stable than those determined by the direct S-wave. *Phillips and Aki* [1986] estimated the distribution of station effects, using coda waves for local earthquakes in central California. They suggested that station effects reflect surface geology, such as sediment and granite, at least in a low-frequency range (<4 Hz). Many directions of progress have been made in this research field since then, as summarized by *Sato and Fehler* [1998].

[21] Coda waves may be classified into two parts on the basis of the travel time of each scattered phase so that both source and station effects can be corrected quite accurately by using a coda-normalization approach. The early coda part is assumed to contain a specific scattered phase from each seismic scatterer distributed nonuniformly [e.g., *Spudich and Bostwick*, 1987; *Gupta et al.*, 1990] and to be imaged in the following back-projection procedure. On the other hand, the later coda part consists of scattered phases from seismic scatterers distributed over a large extent of the studied area and represents the average characters of the scatterers [e.g., *Aki and Chouet*, 1975; *Sato and Fehler*, 1998].

[22] Later coda parts of local earthquakes are generally defined as the waves arriving after twice of the direct S-wave travel times [*Rautian and Khalturin*, 1978]. *Aki* [1969] showed that the temporal decay of the later coda parts for aftershocks of the 1966 Parkfield earthquake are nearly identical to each other, suggesting that the later coda parts are scattered waves from scatterers distributed over a large volume surrounding station and earthquake locations, rather than distinct seismic phases from seismic velocity discontinuities. In addition, many seismological observations

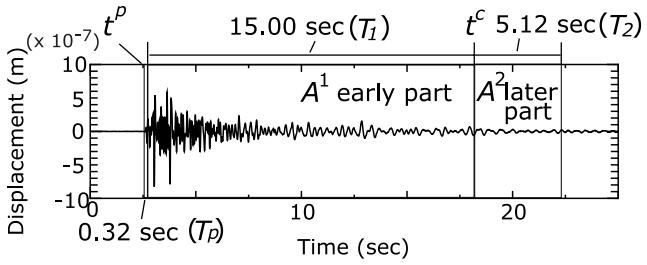

Figure 3. Two time windows (T_1 and T_2) and the reference time (t^c) of the coda-normalization approach are shown for an example of the observed seismogram. The early coda part starts from $t^p + T_P$ where t^p and T_P are the travel time and the time window of the direct *P*-wave, respectively. The time window of the early part of coda T_1 is 15 s while one of the later part T_2 is 5.12 s in Paper II.

strongly suggest that the temporal decay rate of the later coda part is nearly independent of source and station locations and source radiation patterns for a given research area, and it is dependent only on the frequency considered [e.g., *Phillips and Aki*, 1986].

[23] In this section, we shall explain how source and station effects can be taken into account, and how to parameterize the time windows for the early and later coda parts as well as the direct *P*-wave. Specifically, we here explain the parameterizations for seismograms from explosion sources that were used for our application in Paper II. The majority of seismic radiation from explosion sources is *P* wave although there appear to exist some amounts of direct-like *S* waves generated mainly by the reflection at the ground surface. In addition, its radiation pattern can be treated as to be isotropic. The effect of source radiation patterns of explosion sources should be small on the seismograms that we used in Paper II.

[24] We define the onset time of the analyzed coda wave, measured from the direct *P*-wave travel times (t_p). For our application in Paper II, the time window of the direct *P*-wave (T_p) is set to be 0.32 s while the time windows of the early (T_1) and later (T_2) coda parts are chosen to be 15 s and 5.12 s, respectively (Figure 3). T_p is chosen on the basis of the dominant frequency of the direct *P*-waves so that any portions of the direct *P*-waves could not be included in the early coda part. On the other hand, the time windows of the early and later coda parts are determined on the basis of the observed amplitudes of coda waves in a later coda part. As mentioned above, we assume that the later coda part does not include any distinct seismic phases such as phases converted from seismic velocity discontinuities. With the actual data used in Paper II, there are indeed no clear distinct seismic phases in the later coda parts, and we did not see systematic variations in amplitudes of these parts. For seismograms of local events, the starting time of the later coda part for normalization can be about twice the travel time of the direct *S*-wave [*Aki*, 1969; *Sato and Fehler*, 1998], and the time window of the later coda part can be chosen on the basis of signal-to-noise ratios.

[25] The time- and frequency-dependent spectral amplitude of the coda wave after corrections of the overall propagation effects in the previous section is denoted by $\overline{A}^{ml}_{ijk}(\omega; t)$, where l represents the wave type (i.e., $l = 1$ for *P*- and $l = 2$ for *S*-wave) and m indicates the portion of coda waves (i.e., $m = 1$ for the early coda part and $m = 2$ for the later), observed at the j-th station for the i-th source in the k-th component for the lapse time t of the angular frequency ω.

[26] The spectral amplitude of the early coda part can be expressed by

$$\overline{A}^{1l}_{ijk}(\omega; t_1) = R_{\theta\varphi k} \cdot S_{ik}(\omega) \cdot G^l_{jk}(\omega, \varphi) \cdot I_{jk}(\omega) \cdot P^l_{ijk}(\omega; t_1) : t_p + T_p < t_1 < t_c, \tag{8}$$

and for the later

$$\overline{A}^{2l}_{ijk}(\omega; t_2) = S_{ik}(\omega) \cdot G^l_{jk}(\omega) \cdot I_{jk}(\omega) \cdot P^l_{ijk}(\omega; t_2) \quad : t_c < t_2 < t_c + T_2, \tag{9}$$

where $S_{ik}(\omega)$ is the source term of the i-th source and $G^l_{jk}(\omega)$ is the station term of the j-th station. $I_{jk}(\omega)$ is the instrumental transfer factor for the j-th station and $P^l_{ijk}(\omega; t)$ indicates the scattering term, which is related to the intensity of scattering coefficients sampled by the corresponding part of the coda wave. $R_{\theta\varphi k}$ is the source radiation pattern whose azimuth θ and incident angle φ are specified for a given combination of the i-th source and the j-th station. t_c is the starting time of the later coda part in the form of $t_c = t_p + T_P + T_1$. Note that propagation effects in both the two parts have been already corrected by a statistical amplitude recovery technique, based on the AIC explained in section 2.

[27] To suppress the source and station terms, that is, S_{ik} and G^l_{jk}, we take the spectral amplitude ratio of $\overline{A}^{1l}_{ijk}(\omega; t_1)$ to $\overline{A}^{2l}_{ijk}(\omega; t_2)$,

$$\frac{\overline{A}^{1l}_{ijk}(\omega; t_1)}{\overline{A}^{2l}_{ijk}(\omega; t_2)} = \frac{R_{\theta\varphi k} \cdot S_{ik}(\omega) \cdot G^l_{jk}(\omega, \varphi) \cdot I_{jk}(\omega) \cdot P^l_{ijk}(\omega; t_1)}{S_{ik}(\omega) \cdot G^l_{jk}(\omega) \cdot I_{jk}(\omega) \cdot P^l_{ijk}(\omega; t_2)} \simeq \frac{P^l_{ijk}(\omega; t_1)}{P^l_{ijk}(\omega; t_2)} \equiv \widehat{A}^l_{ijk}(\omega). \tag{10}$$

[28] Besides the common recorded system $I_{jk}(\omega)$, implicit assumptions in equation (10) are supported by the results of previous studies on coda waves, that is, the following simplifications can be grossly applied: the explosion source radiation pattern $R_{\theta\varphi k}$ is isotropic (i.e., not as a function of azimuth θ and incident angle φ) and the station term $G^l_{jk}(\omega, \varphi)$ is grossly independent of incident angle φ [e.g., *Tsujiura*, 1978]. The coda wave $\widehat{u}^l_{ijk}(t)$ after the above correction of source and station effects is calculated by the inverse Fourier transform of $\widehat{A}^l_{ijk}(\omega)$ defined by equation (10),

$$\widehat{u}^l_{ijk}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{A}^l_{ijk}(\omega) e^{-i\omega t} d\omega. \tag{11}$$

Hereafter we shall call $\widehat{u}^l_{ijk}(t)$ a normalized coda wave. Note that we do not consider phase shifts due to station terms

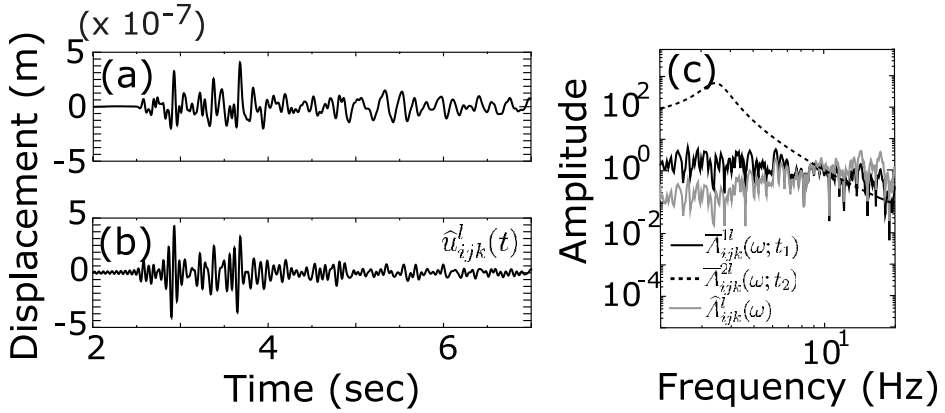

Figure 4. An example of a normalized coda wave, using the coda-normalization approach. (a) An original coda wave and (b) its normalized coda wave. (c) Spectral amplitudes of the original late (black dashed line) and early (black solid line) coda waves, and the normalized coda wave (gray solid line).

because it is difficult to evaluate/remove them, particularly high-frequency (>1 Hz) seismograms.

[29] We can now take the normalized coda wave $\widehat{u}^l_{ijk}(t)$ to be affected only by the non-uniform distribution of seismic scatterers. Figure 4 shows an example of normalized coda waves, using the coda-normalization approach explained above. The dominant frequency of this normalized coda wave is higher than one of the original coda wave because the later part of coda wave is dominated by low-frequency (5 Hz) seismic waves, compared with the early part of coda waves. It should be note that we remove effects of source and station with frequency dependency. In other words, we are now able to focus on seismograms only affected by scattering terms as shown in equation (10).

4. Determination of Ray Parameters and Scattering Modes

4.1. Frequency-Wave Number Analysis With a Stationary AR Model

[30] Through an appropriate analysis technique of seismic array data sets, a seismic signal of relatively small amplitude can be enhanced, with respect to ambient seismic noise in records. This is the primary reason why seismic arrays prove very useful in studies on small-scale heterogeneities of the Earth's interior [e.g., *Neidell and Taner*, 1971; *Aki et al.*, 1976] as well as on high-frequency earthquake rupture processes [e.g., *Spudich and Cranswick*, 1984; *Goldstein and Archuleta*, 1991]. Several sophisticated seismic array techniques have been proposed and applied for the identification of seismic signals, such as the *f-k* [e.g., *Capon*, 1969; *LaCoss et al.*, 1969], cross-correlation [e.g., *Frankel et al.*, 1991; *Del Pezzo et al.*, 1997] and multiple signal classification analyses [e.g., *Schmidt*, 1986; *Goldstein and Archuleta*, 1991; *Almendros et al.*, 2001].

[31] The *f-k* analysis technique with the Fourier transform has been widely utilized to reveal the crustal and upper mantle seismic structure beneath seismic arrays, as summarized by *Rost and Thomas* [2002]. Although this *f-k* analysis technique is one of the most effective seismic array techniques, this technique has limitation in accuracy, for example, there are large sidelobes in the time-frequency domain. We develop a *f-k* analysis technique with a stationary AR model for more accurate detection of seismic signals in coda waves than the conventional approach.

[32] In the case of a two-dimensional seismic array, the *f-k* power spectrum with the designated wave number vector $\mathbf{k}(k_x, k_y)$ and angular frequency ω is defined by

$$\widehat{P}(k_x,k_y,\omega) = \int_{-\infty}^{\infty} \exp(i\omega\tau)d\tau \frac{1}{N^2}\langle b(k_x/\omega,k_y/\omega,t) \cdot b(k_x/\omega,k_y/\omega,t+\tau)\rangle, \quad (12)$$

where the symbol $\langle\rangle$ indicates the average over time t. N is the number of data. $\widehat{P}$ is conventionally called *f-k* power spectrum. $b(k_x/\omega, k_y/\omega, t)$ is the beam output expressed by

$$b(k_x/\omega,k_y/\omega,t) = \frac{1}{L}\sum_{l=1}^{L} d_l(t+t_l), \quad (13)$$

where L indicates the number of stations and $d_l(t)$ is the seismogram at the l-th station. The time delay t_l is calculated by using the distance vector, $\mathbf{r}_l - \mathbf{r}_0$, with a given wave number vector $\mathbf{k}$, where $\mathbf{r}_l$ and $\mathbf{r}_0$ are the coordinate vectors of the l-th station and a reference station, respectively,

$$t_l = \mathbf{k}\cdot(\mathbf{r}_l - \mathbf{r}_0)/\omega. \quad (14)$$

[33] As shown in equation (12), the power spectrum $\widehat{P}(k_x, k_y, \omega)$ of beam outputs b $(k_x/\omega, k_y/\omega, t)$ is obtained by calculating its autocorrelation, followed by performing its the Fourier transform in the standard *f-k* analysis technique. In other words, the *f-k* analysis technique transforms observed time domain array seismograms to frequency domain periodograms. We then need to smooth each autocorrelation with a certain time domain window or take the squared magnitude of the periodograms obtained by the Fourier transform. Such a smoothing filtering corresponds to an autocorrelation decaying in both temporal directions. The use of a smoothing filter or finite time-window is seldom ideal in the present analysis because it distorts signals, and there are some limitations for the resolution in the frequency domain.

[34] Several new methods have been devised for obtaining spectra of more enhanced resolution than conventional filters. One of the most promising methods was devised by *Burg* [1967], which is called the maximum entropy method (MEM). The biggest advantage of MEM is its high resolution for a time series of a short length sampled at an equal interval. The MEM is principally identical to the AR spectral estimator devised by *Akaike* [1969]. A stationary AR model, which was discovered by *Yule* [1972], can be described as a stochastic process in which the current value, $u(t_n)$, of a given time series is obtained by the convolution of a wavelet with past M values of the signal, $u(t_{n-m})$ ($m = 1, 2, \cdots, M$), and $\varepsilon(t_n)$, which is usually a sample of random (white) noise.

[35] In this study, the time series in the MEM corresponds to the observed seismogram and should not be stationary. If we properly divide the observed seismograms into short intervals, however, the one at each short interval can be regarded as a locally stationary time series [*Kitagawa and Akaike*, 1978; *Takanami*, 1991].

[36] To estimate *f-k* power spectra in the time-frequency domain of high resolution, we divide the entire beam output into many short beam outputs $b(k_x/\omega, k_y/\omega, t_n)$ at discrete times of $n = 1, 2, \cdots, N$ with a constant interval δt from equation (13). We assume that each short beam output $b(k_x/\omega, k_y/\omega, t_n)$ is expressed as the following linear combination of the previous values (up to the M-th previous increment of time), based on a stationary AR model:

$$b(k_x/\omega, k_y/\omega, t_n) = \sum_{m=1}^{M} a(m) b(k_x/\omega, k_y/\omega, t_{n-m}) + \varepsilon(t_n), \quad (15)$$

where $a(m)$ are the AR coefficients of an order m, and $\varepsilon(t_n)$ represents the Gaussian white noise with zero mean and the variance σ^2. The AR coefficients $a(m)$ are estimated by minimizing the variance of the distribution of noise $\varepsilon(t_n)$, which is defined by

$$\sigma^2(M) = \frac{1}{N} \sum_{n=1}^{N} \left(b(k_x/\omega, k_y/\omega, t_n) - \sum_{m=1}^{M} a(m) b(k_x/\omega, k_y/\omega, t_{n-m}) \right)^2. \quad (16)$$

The highest probability of realization is given by minimizing the value of $\sigma^2(M)$.

[37] In the process of estimating AR coefficients $a(m)$ for real data, we essentially have ambiguity in the choice of the order of an stationary AR model, M. Increasing the value of M provides a narrower distribution of $\varepsilon(t_n)$, resulting in higher realization probability. This, however, leads imprecise estimation of AR coefficients because any excessive reduction of $\sigma^2(M)$ overestimates the probability of realization. A statistically sound trade-off between increasing M and decreasing $\sigma^2(M)$ is obtained by measuring the following AIC value [e.g., *Akaike*, 1973]:

$$\mathrm{AIC}(M) = N \log\left(2\pi\sigma^2(M)\right) + N + 2(M+1). \quad (17)$$

A smaller value of AIC indicates a better model for applying the locally stationary AR model, balancing the choice of M and $\sigma^2(M)$. This value gives an objective criterion to understand the trade-off between these two parameters.

[38] After obtaining the coefficients $a(m)$ ($m = 1, 2, \cdots, M$) and the variance $\sigma^2(M)$, we estimate the *f-k* power spectra $\widehat{P}(k_x, k_y, \omega)$ of $b(k_x/\omega, k_y/\omega, t_n)$. A stationary AR model can express a linear-system, that is, input data and output data are time series (or beam output $b(k_x/\omega, k_y/\omega, t_n)$) and $\varepsilon(t_n)$, respectively [*Burg*, 1967]. The transfer function between input and output can be expressed by AR coefficients $a(m)$. In the case of a linear system, the power spectral density of beam output (i.e., input data) in the frequency domain is defined by

$$\widehat{P}(k_x, k_y, \omega) = \frac{\sigma^2(M)}{\left|1 - \sum_{m=1}^{M} a(m) \exp^{(-i\omega m \Delta t)}\right|^2}. \quad (18)$$

[39] The dominant frequency of scattered phases can be proportional to the size of the scatterers [e.g., *Wu*, 1982; *Sato*, 1984]. *Aki and Chouet* [1975] showed that the temporal decay of coda waves is strongly frequency dependent, by analyzing the seismograms of local earthquakes recorded in the Kanto area, Japan and by the NORSAR array, Norway. *Yomogida et al.* [1997] showed the dominant frequency of scattering to be the strongest if the wavelength of seismic waves is comparable to the size of the scatterers, based on the numerical analysis of the properties of scatterers. It is, therefore, important not only to detect the scattered phases from observed seismograms but also to evaluate their dominant frequencies.

[40] We here performed a simple numerical experiment, in order to examine the resolution of time-frequency spectra with a stationary AR model. We considered six scattered phases of a Ricker wavelet type with the three different dominant frequencies: 3 Hz, 6 Hz, and 12 Hz (Figure 5a). We assumed a sampling interval of 0.01 s and a time window of 2.5 s. We compared the spectrum estimated on the basis of a stationary AR model (Figure 5b) with the smoothed spectrum estimated on the basis of a discrete Fourier transform (DFT) (Figure 5c). The length of the time window for calculating spectra was 0.32 s.

[41] The stationary AR model method with a short time window of 0.32 s can reveal detail features of spectra in the time-frequency domain of high resolution. In contrast, the result with a 32-s (point) DFT method shows relatively poor properties with noticeable spectral leakage. This advantage of a stationary AR model method illustrated in Figure 5 will enhance the accuracy in measurement of coherent scattered phases in the early coda part defined in the previous section and to be mapped in a later section.

[42] To assess the stability of the *f-k* analysis technique explained above, we performed a reliability test, by using observed seismograms recorded by seismic arrays deployed near the Nagamachi-Rifu fault, northeastern Japan. These seismic data were recorded by seismic refraction/reflection experiments and were used to estimate a distribution of scatterers in Paper II. Since the shotpoints and their origin times of the explosion sources in the seismic experiments were determined more accurately than those of earthquakes, we can calculate the theoretical 2-D slowness vectors precisely, on the basis of the geometries of the shots and the arrays and a 3-D *P*-wave velocity structure.

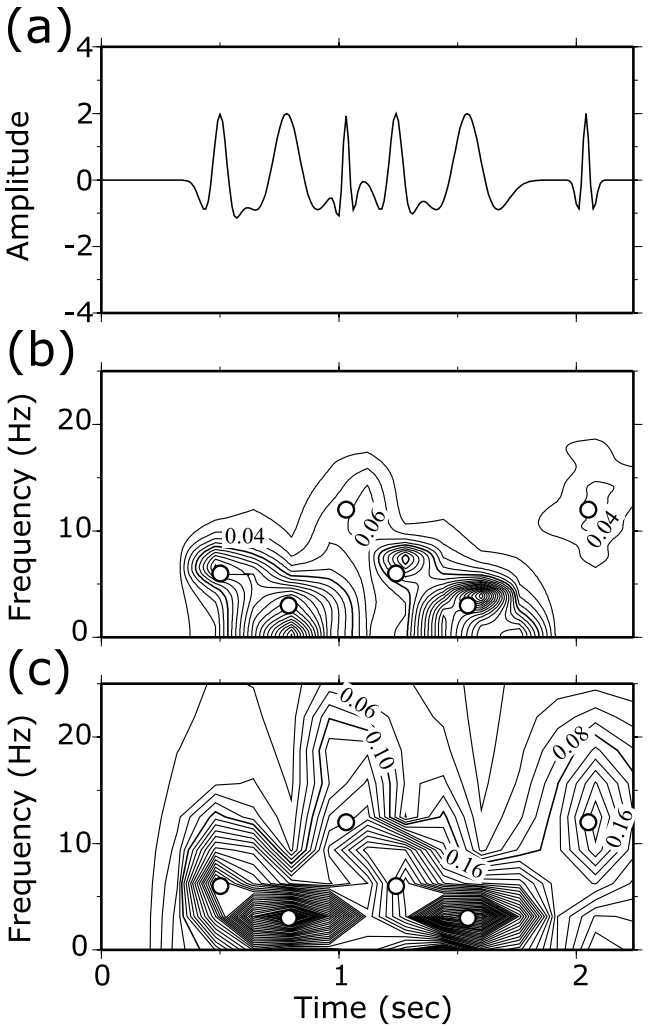

Figure 5. Comparison of spectral estimation using a stationary AR model with a DFT method. (a) A synthetic seismogram containing six Ricker wavelets with three different dominant frequencies (3 Hz, 6 Hz, and 12 Hz). The time-frequency spectra with (b) a stationary AR model and (c) a DFT method. The length of the time window for calculating spectra is 0.32 s. Open circles indicate the spectral peaks of the given Ricker wavelets in the time-frequency domain.

[43] Figure 6 shows the array seismogram recorded at a dense three-component array for an explosion source, and also represents the seismogram recorded at one station. This data set is hereafter called "the actual test data." The geometry of the shotpoint and the array is shown in Figure 7. The coordinate center is the epicenter of the M = 5.2 earthquake (see Paper II). The array contained 60 short-period three-component seismometers with the natural frequency of 4.5 Hz. The average station spacing is approximately 0.05 km (Figure 7). Figure 8 shows the estimated *f-k* power spectrum for the direct *P*-wave of the actual test data. This calculation was performed in the 2–16 Hz band width. The time window length was 0.32 s, which used to our application in Paper II. The slowness domain was a range of −0.512 s/km to 0.512 s/km with the sampling interval of 0.016 s/km. These parameterizations were also the same as those in our application of Paper II. The center of the time window for a central station of the array was the direct *P*-wave arrival (Figure 6b). We found that the estimated *f-k* power spectrum has one peak in the 2-D slowness domain and its maximum point (i.e., 2-D slowness vector) is close to the theoretical one (i.e., the 2-D slowness vector toward to the location of the shot).

[44] We next examined the resolution in the frequency-wave number domain for the array, by using a simple numerical experiment. We generated three Ricker wavelets having the same arrival time at a central station of the array but different wave numbers (k_x and k_y) and dominant frequencies: 3 Hz, 6 Hz, and 12 Hz (Figure 9), and then estimated the *F-k* power spectra in three frequency ranges: 2–4 Hz, 4–8 Hz, and 8–16 Hz, through the *f-k* analysis technique explained above. Gaussian white noise of 50% of the average signal level were added to synthetic waveforms. The time window length and the slowness domain for calculating *f-k* power spectra were the same as those in the reliability test mentioned above. The center of time window for the central station was selected so as to be the theoretical travel time at this station. The array configuration as well as the geometry of the source and the array were also the same as those in the reliability test. Figure 10 shows the estimated *f-k* power spectra for this numerical experiment. The maximum *f-k* power spectrum in each frequency range appear close to its theoretical maximum point. We will discuss an effects of scattered phase durations on the estimation of slowness vectors in the next section.

4.2. Polarization Analysis With a Stationary MAR Model

[45] The information of polarization or particle motion of a certain seismic phase is a useful tool for identifying its type of the seismic wave, such as *P* wave, *S* wave, Rayleigh wave, and Love wave. The degree of elliptical polarization is one basic measure in polarization analysis, which shows the characteristics of seismic phases in a given seismogram [e.g., *Montalbetti and Kanasewich*, 1970; *Samson*, 1977].

[46] The shape of a hodogram (i.e., a projection of three-dimensional particle motions) should have a linear shape for the arrival of scattered phases which are converted to *P* or *S* wave in coda waves. On the contrary, it has a spherical shape when incoherent seismic phases such as random noise are incident. This parameter must help us to identify weak but coherent seismic phases. Several polarization analysis techniques using a covariance matrix have been developed in both earthquake and exploration seismology [e.g., *Kanasewich*, 1973; *Nagano et al.*, 1986; *Vidale*, 1986; *Lilly and Park*, 1995]. *Soma et al.* [2002] proposed the enhanced polarization analysis technique with a wavelet transform, estimating geothermal reservoir structures at the depth range of 1.5 km to 5.5 km at Soultz-sous-Forêts in northeastern France, by using hodogram linearities.

[47] To evaluate the degree of elliptical polarization, including frequency dependency, we improve a standard covariance matrix method [e.g., *Wagner and Owens*, 1996; *Yoshizawa et al.*, 1999], using a stationary MAR model in the similar manner to the *f-k* analysis explained in the previous section. A time series vector $\mathbf{u}(t_n)$ has three-component (u_x, u_y, u_z), 3-C, at discrete times of n =

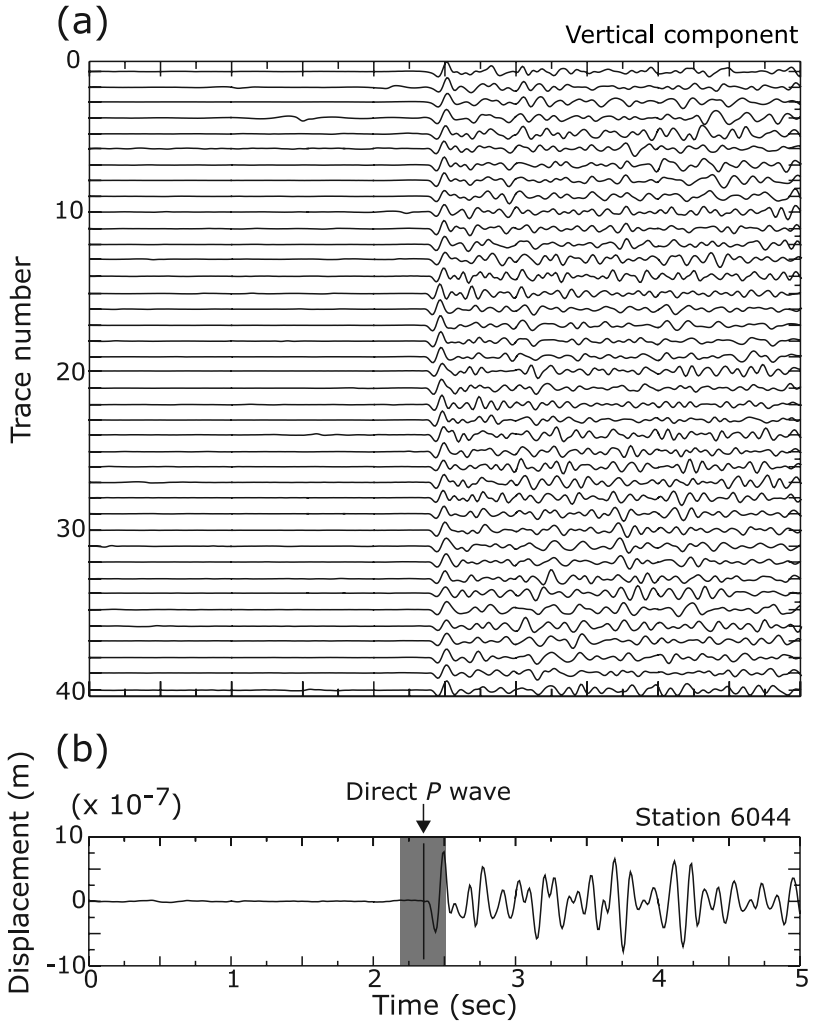

Figure 6. Bandpass filtered (2–16 Hz) seismograms in the vertical component for the actual test data (see text) recorded at (a) the dense seismic array (see Figure 7) and (b) one station of the seismic array. Shaded area represents the time window for calculating the 2-D slowness vector of the direct *P*-wave. Note that we exclude the seismograms with poor signal-to-noise ratios.

1, 2, $\cdots$, N, with a constant interval Δt in a specific time window T,

$$\mathbf{u}(t_n) = \left(u_x(t_n), u_y(t_n), u_z(t_n)\right)^{\mathrm{T}}, \tag{19}$$

where N represents the number of data and the superscript T represents the vector transpose.

[48] We can define the covariance matrix $\mathbf{G}(\omega, T)$ of $\mathbf{u}$ (t_n) in the time-frequency domain, for a time window T over the time series vector $\mathbf{u}(t_n)$ as follows:

$$\mathbf{G}(\omega, T) = \begin{pmatrix} S_{xx}(\omega, T) & S_{xy}(\omega, T) & S_{xz}(\omega, T) \\ S_{yx}(\omega, T) & S_{yy}(\omega, T) & S_{yz}(\omega, T) \\ S_{zx}(\omega, T) & S_{zy}(\omega, T) & S_{zz}(\omega, T) \end{pmatrix} \tag{20}$$

$$S_{ij}(\omega, T) = U_i(\omega, T)U_j(\omega, T)^* \quad (i,j = x,y,z), \tag{21}$$

where $U_i(\omega, T)$ indicates the Fourier transform of $u_i(t_n)$ within the time window T. ω is the angular frequency and the asterisk represents the complex conjugate vector. This Fourier transform approach, however, should not be effective in estimating the degree of elliptical polarization due to the limited accuracy if we like to use a short time window, as discussed in the previous section. We then introduce the following polarization analysis technique with a stationary MAR model, in order to estimate a covariance matrix of a given set of 3-C seismograms in the time-frequency domain of high resolution.

[49] First, we describe a stationary MAR model briefly. We follow a technique proposed by *Akaike and Kitagawa* [1994] to compute the covariance matrix, $\mathbf{G}$. The 3-C seismograms $\mathbf{u}(t_n; r)$ at the r-th station may be expressed by a stationary three-dimensional AR model (i.e., the

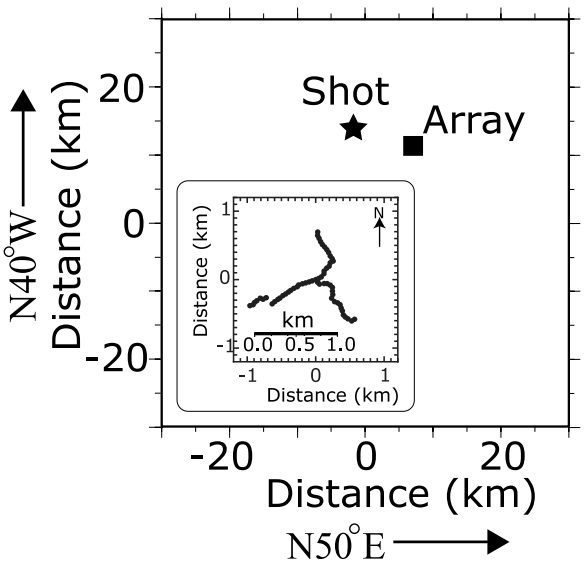

Figure 7. Geometry of the shot (solid star) and array (solid square) for the actual test data. The inserted figure shows the array configuration.

present value is expressed by a linear combination of past values) as follows:

$$\mathbf{u}(t_n; r) = \sum_{m=1}^{M} \mathbf{A}(m; r)\mathbf{u}(t_{n-m}; r) + \mathbf{w}(t_n; r), \tag{22}$$

where M denotes the order of a stationary three-dimensional AR model. $\mathbf{w}(t_n; r)$ is assumed to be the three-dimensional Gaussian white noise vector whose each element has the zero mean value,

$$E(\mathbf{w}(t_n; r)) = (0, \cdots, 0)^{\mathrm{T}}, \tag{23}$$

where E represents the average over the number of data N, in a short time window T.

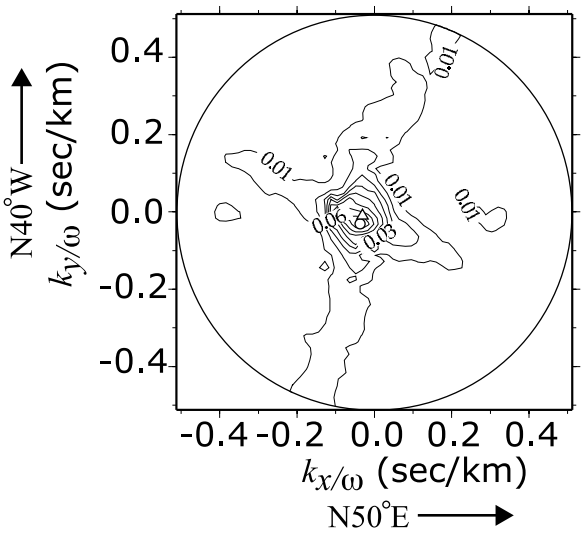

Figure 8. Estimated the *f-k* power spectrum for the direct *P*-wave in Figure 6. Open circle and triangle indicate the estimated and theoretical maximum points, respectively.

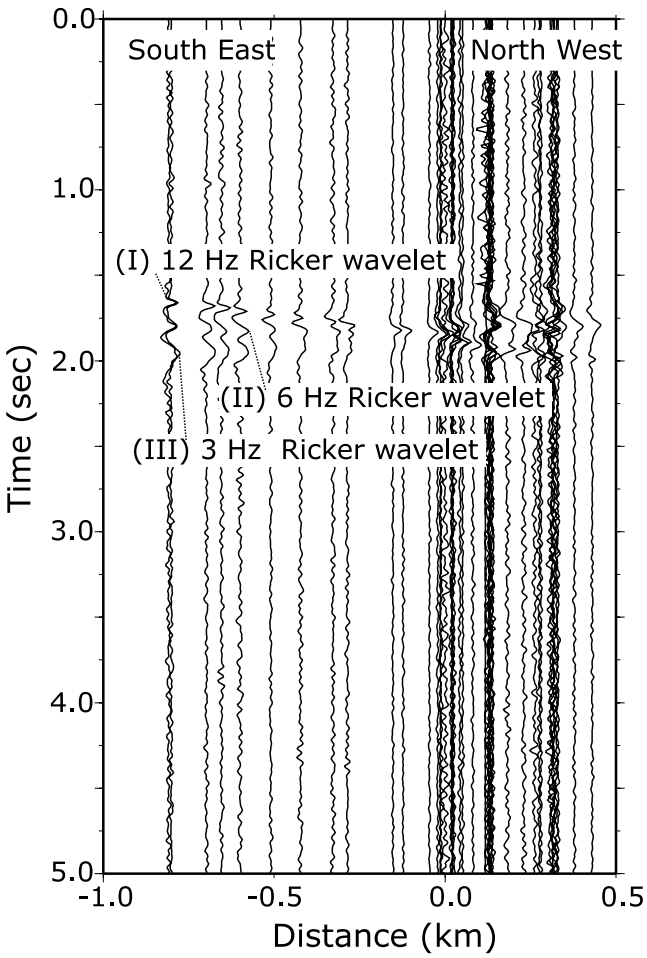

Figure 9. Synthetic seismograms for examining the resolution in the frequency-wave number domain for the array. The seismogram contains three Ricker wavelets with different frequencies: 3 Hz, 6 Hz, and 12 Hz. The travel times of these three wavelets are the same as each other at the central station at distance 0 km.

[50] A 3 × 3 AR coefficient matrix, $\mathbf{A}(m; r)$, defined by equation (22) is expressed by

$$\mathbf{A}(m; r) = \begin{bmatrix} a_{xx}(m; r) & a_{xy}(m; r) & a_{xz}(m; r) \\ a_{yx}(m; r) & a_{yy}(m; r) & a_{yz}(m; r) \\ a_{zx}(m; r) & a_{zy}(m; r) & a_{zz}(m; r) \end{bmatrix}, \tag{24}$$

where $a_{ij}(m; r)$ is called the ij-th AR coefficient. The coefficients are estimated by solving a normal equation for the least squares estimation of equation (22), minimizing the variance of $\mathbf{w}(t_n; r)$.

[51] The order of AR coefficients M can be determined by the AIC, in a similar manner to the *f-k* analysis with a stationary AR model explained in the previous section. In this case, the AIC value is defined by

$$\mathrm{AIC}(M; r) = N \log|\mathbf{C}(M; r)| + 2M \cdot 3^2 \tag{25}$$

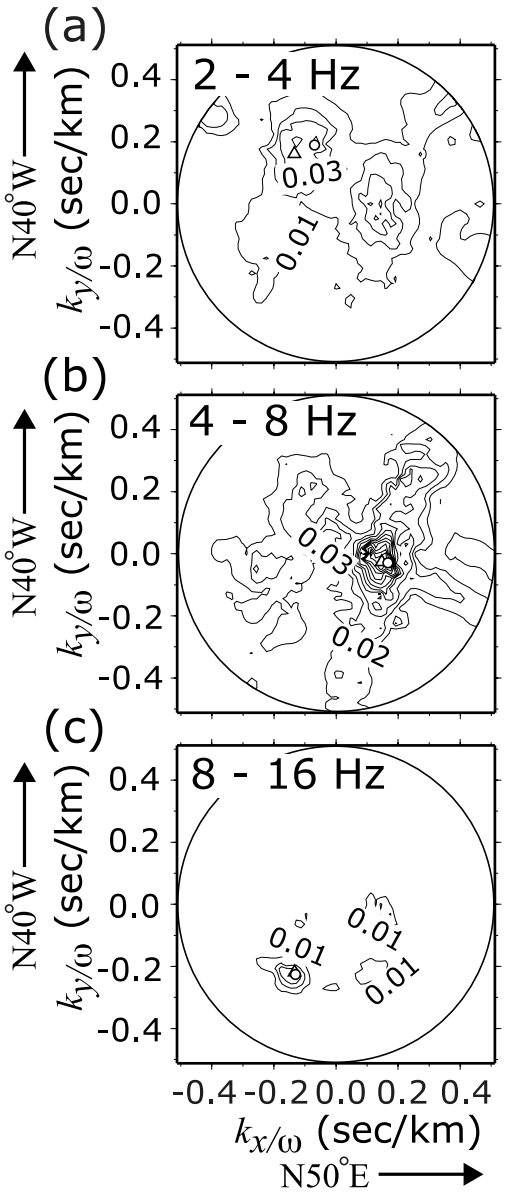

Figure 10. Estimated *f-k* power spectra in frequency ranges of (a) 2–4 Hz, (b) 4–8 Hz, and (c) 8–16 Hz for the numerical experiment with the seismogram in Figure 9. Open circles and triangles indicate the estimated and theoretical maximum points, respectively.

$$E\left(\mathbf{w}(t_n;r)\mathbf{w}(t_n;r)^{\mathrm{T}}\right) = \begin{bmatrix} \sigma_{xx}(r) & \sigma_{xy}(r) & \sigma_{xz}(r) \\ \sigma_{yx}(r) & \sigma_{yy}(r) & \sigma_{yz}(r) \\ \sigma_{zx}(r) & \sigma_{zy}(r) & \sigma_{zz}(r) \end{bmatrix} \equiv \mathbf{C}(M;r), \tag{26}$$

where $\mathbf{C}(M; r)$ represents the covariance matrix of $\mathbf{w}(t_n;r)$ with the covariance $\sigma_{ij}(r)$ of the i-th and j-th components as each element. The order M is selected so that the AIC value of equation (25) becomes minimum.

[52] The power spectral density in the form of a 3 × 3 matrix, $\mathbf{S}(\omega; M, r)$, of the 3-C seismogram $\mathbf{u}(t_n; r)$ is now expressed by

$$\mathbf{S}(\omega;M,r) = \mathbf{A}(\omega;M,r)^{-1}\mathbf{C}(M;r)\widetilde{\mathbf{A}}(\omega;M,r)^{-1}, \tag{27}$$

where $\mathbf{A}(\omega; M, r)$ is the discrete Fourier transform of the matrix $\mathbf{A}(m; r)$ with $m = M$ in equation (24). The tilde denotes the complex conjugate and transport operator. The power spectra $S_{ii}(\omega; M, r)$ (i.e., the i-th diagonal element of $\mathbf{S}(\omega; M, r)$) of the i-th component is given by

$$S_{ii}(\omega;M,r) = \sum_{j=1}^{N} \left|A_{ij}(\omega;M,r)^{-1}\right|^2 C_{jj}(M;r) \qquad (i = x,y,z), \tag{28}$$

where $C_{jj}(M; r)$ is the j-th diagonal element of the covariance matrix $\mathbf{C}(M; r)$ and $A_{ij}(\omega; M, r)^{-1}$ is the (i, j) element of the inverse 3 × 3 matrix of $\mathbf{A}(\omega; M, r)$ (see *Akaike and Kitagawa* [1994] for details).

[53] Using $S_{ij}(\omega, T; M, r)$ as the ij-th element of the power spectral matrix $\mathbf{S}(\omega; M, r)$ of equation (27) for a given time window T, we can rewrite the covariance matrix $\mathbf{G}(\omega, T; M, r)$ at the r-th station with a stationary three-dimensional AR model of degree M as follows:

$$\mathbf{G}(\omega,T;M,r) = \begin{pmatrix} S_{xx}(\omega,T;M,r) & S_{xy}(\omega,T;M,r) & S_{xz}(\omega,T;M,r) \\ S_{yx}(\omega,T;M,r) & S_{yy}(\omega,T;M,r) & S_{yz}(\omega,T;M,r) \\ S_{zx}(\omega,T;M,r) & S_{zy}(\omega,T;M,r) & S_{zz}(\omega,T;M,r) \end{pmatrix}. \tag{29}$$

[54] Note that we conduct the above procedure for a 3-C seismogram at each station of an array. We here consider an effect of a reverberation in a shallow subsurface layer on the estimation of polarization vectors. To suppress this effect, we use the average covariance matrix $\overline{\mathbf{G}}(\omega, T; M)$ among the stations for a specific scattered phase [e.g., *Jurkevics*, 1998],

$$\overline{\mathbf{G}}(\omega,T;M) = \frac{1}{R}\sum_{r=1}^{R}\mathbf{G}(\omega,T;M,r), \tag{30}$$

where R indicates the number of stations. The singular value decomposition of the averaged covariance matrix $\overline{\mathbf{G}}(\omega, T; M)$ is given by

$$\overline{\mathbf{G}} = \mathbf{U}\mathbf{\Lambda}\widetilde{\mathbf{V}}, \tag{31}$$

where $\mathbf{U}$ and $\mathbf{V}$ are 3 × 3 unitary matrices, and $\mathbf{\Lambda}$ is a 3 × 3 diagonal matrix with three singular values of $\overline{\mathbf{G}}$ [e.g., *Aki and Richards*, 1980, chapter 11]. Since $\overline{\mathbf{G}}$ is Hermite, $\mathbf{\Lambda}$ is composed of three real eigenvalues ($|\lambda_0| \geq |\lambda_1| \geq |\lambda_2|$).

[55] Next, we redefine Vidale's elliptical component of polarization P_E [*Vidale*, 1986], using the above the eigenvalues of the averaged covariance matrix $\overline{\mathbf{G}}(\omega, T; M)$ of equation (30) as follows:

$$P_E = \frac{\sqrt{1-L^2}}{L}, \tag{32}$$

with

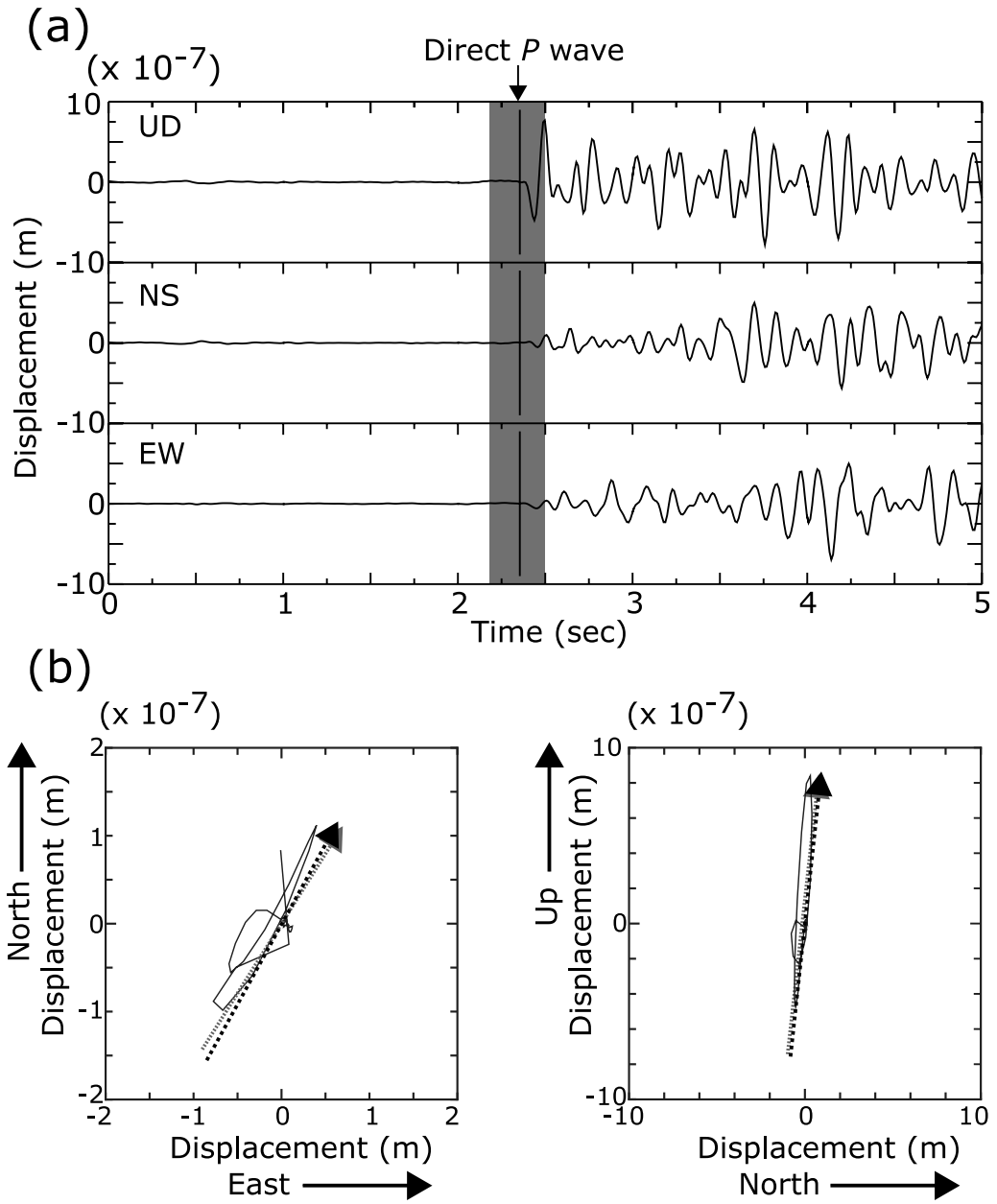

Figure 11. (a) Three-component seismogram at one station of the array. The shaded area represents the time window for measuring the polarization vector of the direct *P*-wave. (b) Result of the reliability test of the polarization analysis technique proposed in this paper. The estimated (black arrow) and theoretical (gray arrow) polarization vectors. Also shown in the particle motions of the direct *P*-wave in the analyzed time window shown in Figure 11a.

$$L = \sqrt{(\mathrm{Re}(x_0 \cdot \mathrm{cis}\alpha))^2 + (\mathrm{Re}(y_0 \cdot \mathrm{cis}\alpha))^2 + (\mathrm{Re}(z_0 \cdot \mathrm{cis}\alpha))^2}, \tag{33}$$

where the complex eigenvector $\mathbf{v}_0\ (\omega, T; M) = (x_0, y_0, z_0)^{\mathrm{T}}$, associated with the largest eigenvalue λ_0, pointing in the direction of the maximum polarization. $\mathrm{cis}\alpha = \cos\alpha + i\sin\alpha$, and Re() represents the real part. We search for the value of α to maximize the length of the real component of the eigenvector $\mathbf{v}_0$. P_E becomes 1 for a circular polarization, while 0 for a perfectly linear one (see *Vidale* [1986] for details). We use P_E of equation (32) to distinguish body-wave scattered phases from surface waves and noise.

[56] Using the above the eigenvector $\mathbf{v}_0$, we also define the strike, Θ_{H}, and dip, Θ_{V}, of the direction of the maximum polarization projected on the horizontal plane as follows:

$$\Theta_{\mathrm{H}} = \tan^{-1}\left(\frac{\mathrm{Re}(y_0)}{\mathrm{Re}(x_0)}\right) \tag{34}$$

$$\Theta_{\mathrm{V}} = \tan^{-1}\left(\frac{\mathrm{Re}(z_0)}{\sqrt{\mathrm{Re}(x_0)^2 + \mathrm{Re}(y_0)^2}}\right). \tag{35}$$

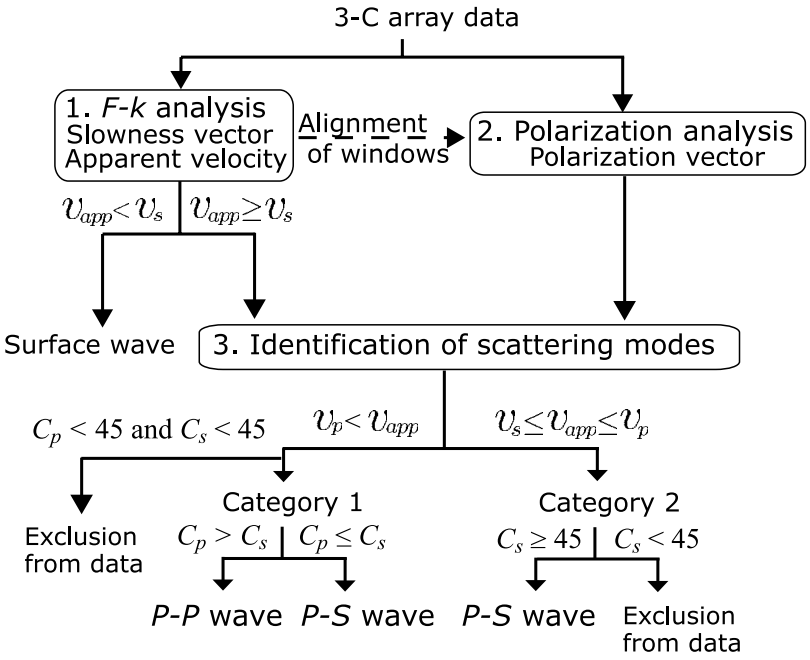

Figure 12. A flowchart for estimating the scattering mode of scattered phases for an explosion source.

[57] The polarization analysis technique with a stationary three-dimensional AR model explained in this section enables us to estimate a covariance matrix in the time-frequency domain of high resolution in the similar manner to the spectrogram shown in Figure 5. We are able to obtain the degree of elliptical polarization and direction, that is, the azimuth and dip of the maximum polarization more precisely and stably than ever.

[58] Using the actual test data described in section 4.1, we performed a reliability test of the polarization analysis technique with a stationary MAR model introduced in this paper. We estimated the polarization vector for the direct *P*-wave (Figure 11) whose the theoretical polarization vector can be calculated on the basis of the geometry of the shot and array as well as a 3-D *P*-wave velocity structure. The time window length was set to be 0.32 s, similar to our analysis in Paper II. Figure 11b shows the determined polarization vector to be nearly identical to the theoretical one.

[59] We next examined the threshold of P_E for extracting body-wave scattered phases from observed seismograms, by using a simple numerical experiment. We generated waveforms only containing randomly amplitudes following a Gaussian distribution. To find the threshold of P_E statistically, we generated 30 different sets of traces. P_E of noise is always higher than 0.4, and then we selected $P_E = 0.4$ for our application in Paper II, confirming no significant noise observed seismograms based on the threshold of P_E.

[60] We additionally discuss the stability of a methodology with stationary AR or MAR models for analyzing seismic phases. As mentioned in section 1, our purpose is to detect scattered phases and to determine their scattering points/strengths as well as identify the scattering modes, by combining the 2-D slowness vectors (see section 4.1) and the polarization vectors (this section) of the detected scattered phases in addition to their travel times (see section 5).

[61] One important issue is the effects of scattered phase durations on the estimations of the slowness and polarization vectors. As mentioned above, we assumed that a seismic signal in a short time window (0.32 s for our application) can be treated as a locally stationary time series [*Takanami*, 1991]. If this assumption is not valid for a specific scattered phase, however, the estimations of the slowness and polarization vectors of this phase could be unstable. Our application (Paper II) used scattered phases identified by at least two shot-array pairs. Through this procedure, we expect the unstable estimations of slowness and polarization vectors should be suppressed.

4.3. Identification of Scattering Modes

[62] Seismograms recorded by a small-aperture 3-C seismic array have several kinds of advantage over those by a vertical-component seismic array or arrays of isolated 3-C stations [e.g., *Jepsen and Kennett*, 1990; *Wagner and Langston*, 1992; *Wagner and Owens*, 1996]. Using receiver function and shear wave splitting analyses of 3-C seismograms, some studies have revealed, for example, the depth of Moho discontinuity [e.g., *Abers*, 1998] and seismic anisotropy structure of the upper mantle [e.g., *Bear et al.*, 1999] quantitatively.

[63] As an example of studies on high-frequency late arrivals or coda waves, *Wagner* [1997] attempted to reveal a coda excitation process in *P*- and *S*-coda waves for local earthquakes near the San Jacinto fault zone in southern California, using their wave propagation directions and polarization characteristics. He distinguished between coherent and incoherent scattered waves in both *P*- and *S*-coda waves, suggesting that coherent *S*-coda waves are more sensitive in source characteristics, such as the depth, distance, and radiation pattern of source, than coherent *P*-coda waves. *Poppeliers and Pavlis* [2002] deployed a 3-C seismic array to evaluate topographic seismic site responses at Glendora Lake in southwestern Indiana, showing that the

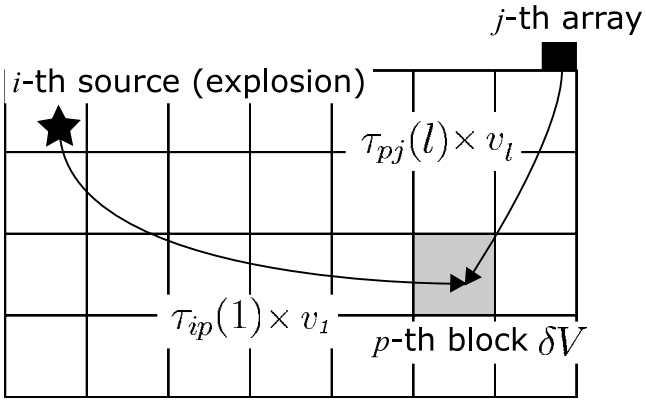

Figure 13. A model geometry and a scattering raypath of a scattered wave by the p-th block of volume δV for the i-th source at the j-th array. Note that we compute the take-off angle from the j-th array with the 3-D slowness vector by utilizing the 2-D slowness vectors determined by the *f-k* analysis and the identified wave-types of scattered waves.

interaction of both body and surface waves has significant effects on the site responses, using the geometrical variation of hodograms along the seismic array.

[64] The combination of slowness and polarization vectors with 3-C seismic array observations may give some useful information on scattered phases in coda waves. We attempt to identify scattering modes of scattered phases in coda waves, using the information of both slowness and polarization vectors, as estimated in the previous sections 4.1 and 4.2. Assuming that we use seismograms only for explosion sources (see Paper II), the scattered modes should be either *P-P* or *P-S* scattering. The identification of these two scattered models is possible by measuring the angle between slowness and polarization vectors. Since we obtain the 2-D wave number vector of each phase by the *f-k* analysis in section 4.1, we additionally need to assume the corresponding *P*- or *S*-wave velocity beneath the 3-C seismic array, in order to estimate its 3-D slowness vector. This velocity likely corresponds to the average velocity of either *P* or *S* wave from the surface down to the scale of the wavelength (i.e., frequency) to be analyzed.

[65] We introduce the scheme of the identification of the scattering mode of each scattered phase, following the criterion of *Nakamura* [1993]. First, we estimate its apparent velocity v_{app} by the *f-k* analysis in section 4.1, and identify each detected phase to be either body wave or surface wave by comparing with the *S*-wave velocity v_s of the surface layer as the representative seismic wave velocity of each phase. If the apparent velocity v_{app} is smaller than the velocity v_s, the phase is considered to be composed mainly of surface waves so that we discard it from our data set (Figure 12).

[66] Next, we classify the remaining scattered waves of the above first procedure further into two categories from the comparison with the reference *P*-wave velocity. If the apparent velocity v_{app} is larger than the *P*-wave velocity v_p of the first layer, the scattered wave cannot be identified only with the information of slowness (category 1). On the other hand, it cannot be *P* wave in the case of $v_s \le v_{app} \le v_p$, so that we conclude it to be *S* wave (category 2).

[67] To identify the scattering mode to be whether *P-P* or *P-S* in the case of category 1 as the final procedure, we incorporate the polarization vector obtained in section 4.2, using the same time window as in the *f-k* analysis. We define the angle, Ψ, between its slowness and polarization vectors by

$$\Psi_{p,s} = \cos^{-1}\frac{\mathbf{s}_{p,s}\cdot\mathbf{p}}{|\mathbf{s}_{p,s}||\mathbf{p}|} \quad : 0 \le \Psi_{p,s} \le 90, \tag{36}$$

where **p** indicates the polarization vector. **s** represents the 3-D slowness vector from the 2-D wave number vector **k** estimated by the *f-k* analysis, assuming the surface layer velocities for *P*- and *S*-waves. The dot in the numerator denotes the *dot product* of two vectors, and $|\ |$ represents the magnitude of a vector. The polarization of *P* wave must be nearly parallel to the propagation direction while the one of *S* wave should be perpendicular. We evaluate the credibility of wave type, C defined as $C_p = 90-\Psi_p$ in degree for *P*- and $C_s = \Psi_s$ for *S*-wave. The scattered waves with C less than 45 are removed from our data set. In the category 1, we identify the scattered wave as *P* wave if C_p is greater than C_s while *S* wave for $C_p \le C_s$. Even in the case of category 2, we cross-check the credibility of one wave type, C_s, to confirm that the polarization does agree with *S* wave.

[68] As mentioned in section 4.2, we can exclude noise from our data set through the threshold of P_E, in addition to exclude surface waves on the basis of the apparent velocity. We are, therefore, able to use body-wave scattered phases to estimate a distribution of scatterers.

[69] As a demonstration of above scheme for identification of scattering mode, we computed the credibility of wave type for the direct *P*-wave of the actual test data. The 2-D slowness and polarization vectors were estimated by the reliability tests of the *f-k* and polarization techniques described in sections 4.1 and 4.2. The apparent velocity for this *P*-wave is 20 km/s because of the almost vertical incident wave. The estimated azimuth (clockwise from north) and dip angle (from vertical up) for the 3-D slowness vector are 30° and 10°, assuming *P*- and *S*-waves, respectively, while the azimuth and dip angle for the polarization vector are 29° and 6°. The angle, Ψ, between the slowness and polarization vectors is 4°, and the credibilities of wave types for *P* and *S* waves are 86 and 2, respectively. The polarization vector of this direct *P*-wave was determined to be nearly parallel to its slowness vector, and then the wave-type was determined as *P*-wave by using the wave-type identification scheme as we expected.

5. Mapping of Small-Scale Heterogeneities

[70] Seismic migration/back-projection is an effective tool for mapping the Earth's subsurface structure. Over last decades, several migration techniques have been proposed such as finite difference [*Claerbout and Doherty*, 1972], Kirchhoff [*Schneider*, 1978], and *f-k* [*Gazdag*, 1978; *Stolt*, 1978] migrations mainly in exploration seismology, and widely applied for imaging natural reservoirs (e.g., oil and gas hydrate) in the area of energy resource development. As mentioned in section 1, *Huang et al.* [1999a, 1999b] recently introduced new migration techniques to image

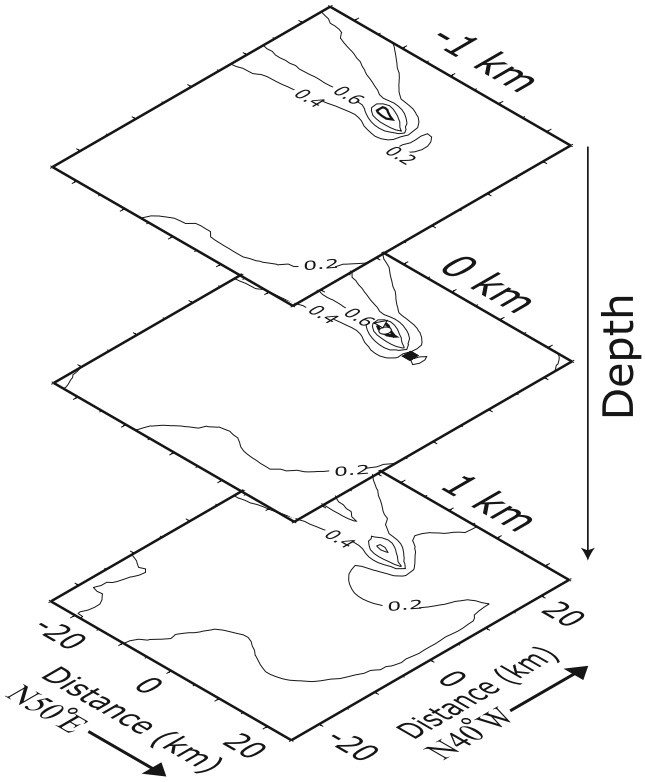

Figure 14. Depth slices of the distribution of the source location probabilities (see text) for the direct *P*-wave of the actual test data. Open circle and triangle represent the estimated and theoretical (i.e., the actual shot location) maximum points of the source location probabilities, respectively. Solid square shows the array location. The bold contour corresponds to the fit function values with greater than 0.95, which we calculate the error in the source location.

complex subsurface structures both stably and accurately, based on the Rytov and Born approximations, respectively. In earthquake seismology, *Ishii et al.* [2005] examined the rupture process of the 26 December 2004 Sumatra-Andaman earthquake, by using a back-projection method of teleseismic *P*-wave amplitude recorded at a seismic network in Japan.

[71] To enhance the resolving power of the distribution of small-scale heterogeneities by taking the advantage of 3-C seismic array data sets, we apply a slowness-weighted back-projection technique, which uses the observed slowness vector and travel time of each detected coherent phase with a three-dimensional seismic velocity structure. Under the assumption of a single scattering model, the travel time $\tau_{ijk}^{cal}(p, l)$ of each scattered phase is calculated by

$$\tau_{ijk}^{cal}(p,l) = \tau_{ip}(1) + \tau_{pj}(l), \tag{37}$$

as the sum of that from the *i*-th source to the *p*-th block $\tau_{ip}(1)$ plus that from the *p*-th block to the *j*-th array $\tau_{pj}(l)$, where *l* indicates the wave type (i.e., $l = 1$ and 2 for *P* and *S* waves).

[72] Assuming an explosion source (see Paper II), $l = 1$ is assigned for the travel time from the source, $\tau_{ip}(1)$, while $\tau_{pj}(l)$ may be either *P* ($l = 1$) or *S* ($l = 2$) wave, as identified by the scheme explained in section 4.3. Figure 13 shows a model geometry and a raypath of a given scattered wave. Small-scale heterogeneities are mapped as scattering coefficients located at each block of the volume δV (Figure 13) in our model space.

[73] In the present back-projection procedure, we first shoot a ray from the *j*-th array with the 3-D slowness vector **s** (i.e., take-off angle) of the *k*-th component, depending on the identified scattering mode (i.e., *P*- or *S*-wave velocity structure). In other words, we can put strong constraint on the raypath of each phase from a given array based on the observed slowness vector. Such well-constrained raypaths leave the very limited number of possible blocks in the three-dimensional model space responsible for each phase identified as a coherent scattered wave.

[74] We next shoot rays from the *i*-th source to the centers of all the possible blocks and compute $\tau_{ip}(1)$, on the basis of the *P*-wave velocity structure as well as calculate the travel times $\tau_{pj}(l)$ for blocks. We here consider an error in the determination of slowness vectors due to the effects of segmented blocks in a model space and finite time window analysis for *f*-*k* and polarization analyses. For this purpose, we introduce the following fit function $f(p, l)$ for travel time and slowness vector between the observed and calculated data, based on a 2-D Gaussian function,

$$f(p,l) = \exp\left(-\frac{\left(\tau_{ijk}^{obs}(l) - \tau_{ijk}^{cal}(p,l)\right)^2}{2\sigma_\tau}\right) \cdot \exp\left(-\frac{\left(\mathbf{s}_{ijk}^{obs}(l) - \mathbf{s}_{ijk}^{cal}(p,l)\right)^2}{2\sigma_{\mathbf{s}}}\right), \tag{38}$$

where σ_τ and $\sigma_{\mathbf{s}}$ are the standard deviations in the estimations of travel time and slowness vector, respectively. In addition, we define possible blocks, p^{sca}, responsible for scattering expressed by

$$p^{sca} = \{p | f(p,l) > \alpha\}, \tag{39}$$

where α denotes the lower limit of the fit value. The uncertainty in locations of scatterers can be evaluated by the value of α. We select $\alpha = 0.95$ for both the shot location with the direct *P*-wave of the actual test data and locations of scatterers with synthetic data.

[75] Since the ray from the station is well constrained by its slowness vector, we can pinpoint the location of the corresponding scatterer or model block very precisely, including its depth. Most of the previous studies on the imaging of heterogeneous structure used simple half-space velocity models [e.g., *Nishigami*, 1991; *Taira and Yomogida*, 2004]. In contrast, we adopt a three-dimensional seismic velocity structure of a target area inverted by seismic tomography and used the pseudo-bending method [*Um and Thurber*, 1987], in order to determine locations of seismic scatterers with much more precise raypaths and travel times.

[76] The scattering coefficient $g_{ij}(p, \omega, k, l)$ of the *p*-th block at the angular frequency ω in the *k*-th component is finally estimated from the observed *f*-*k* power spectra

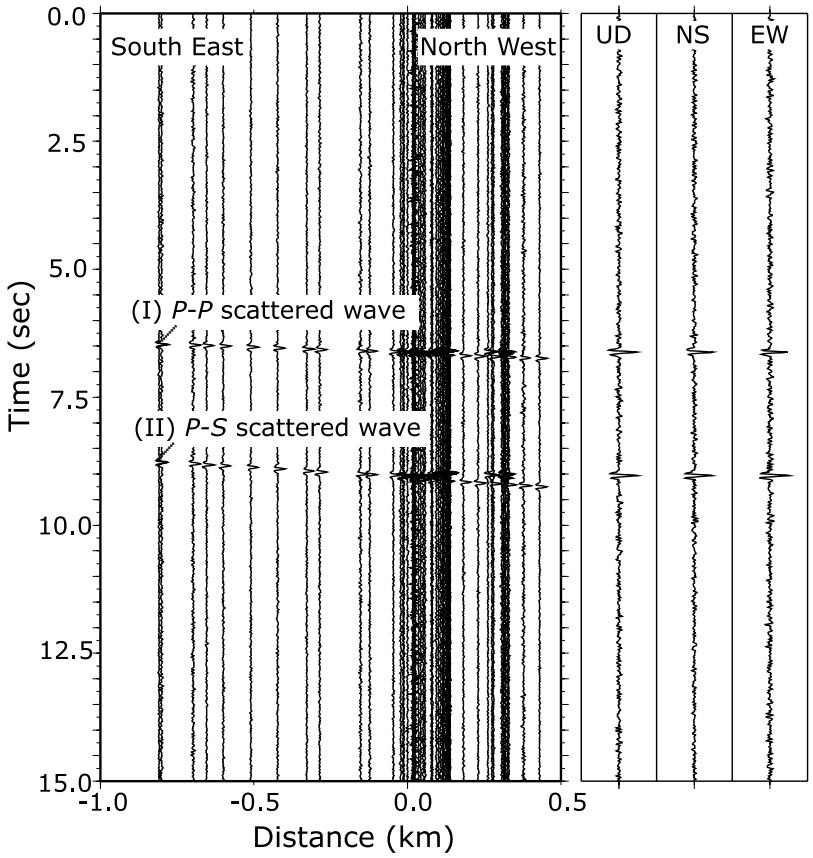

Figure 15. (left) Synthetic seismograms in the vertical component for examining the capability of locating scatterers and (right) the enlarged view of the three-component seismogram at the central station of the array.

$\widehat{P}^{l}_{ijk}(k_x, k_y, \omega)$ of the normalized seismograms $\widehat{u}^{l}_{ijk}(t)$ of equation (11), for a given source-array pair (i.e., the i-th source and the j-th array). To estimate the values of a reliable scattering coefficient, the f-k power spectrum $\widehat{P}^{l}_{ijk}(k_x, k_y, \omega)$ of equation (18) as a function of a lapse time, $\tau^{obs}_{ijk}(l)$, is mapped into the spatial variation for each identified scattering mode l at each ij-th f-k spectrogram,

$$g_{ij}(p, \omega, k, l) = \widehat{P}^{l}_{ijk}\left(k_x, k_y, \omega\right) \qquad \text{at } \tau^{obs}_{ijk}(l). \tag{40}$$

[77] For a given scatterer located at the p-th block, we calculate $g_{ij}(p, \omega, k, l)$ for all the source-array pairs (i.e., all the observed seismograms with the combination of i and j) and average them,

$$\widetilde{g}(p, \omega, k, l) = \frac{1}{N \times M} \sum_{i=1}^{N} \sum_{j=1}^{M} g_{ij}(p, \omega, k, l), \tag{41}$$

where N and M are the source and array numbers, respectively. $\widetilde{g}(p, \omega, k, l)$ means the scattering coefficient of the p-th block at the k-th component in the angular frequency ω for either P-P ($l = 1$) or P-S ($l = 2$) scattering. Note that the above formulations have been processed at each component, in order to discuss our imaging approach in general. In the end, we calculate the sum of values of $\widetilde{g}(p, \omega, k, l)$ for all the three components,

$$\widetilde{g}(p, \omega, l) \equiv \sum_{k=1}^{3} \widetilde{g}(p, \omega, k, l), \tag{42}$$

where $\widetilde{g}(p, \omega, l)$ is the final value to be mapped as scattering coefficient of the p-th block.

[78] As explained in section 1, the spatial distribution of small-scale heterogeneities has been revealed only after the recent availability of dense seismic networks. One important limitation of theses previous studies was the normalization of each seismogram in the first stage of their data analyses in a rather empirical manner. In contrast, our present approach maps the absolute power spectrum of each scattered phase into a scattering coefficient, after the corrections of source, station, and overall propagation effects discussed in sections 2 and 3.

[79] To investigate the location capability of the slowness-weighted back-projection technique, we located the explosion source, by using the direct P-wave of the actual test data. As demonstrated in previous sections, we estimated the 2-D slowness and the polarization vectors for this P-wave, and confirmed that the wave-type is to be P-wave. We set $\tau_{ip}(1) = 0$ for the direct P-wave. We normalized the scattering coefficients by its maximum value and called them source location probabilities.

[80] Figure 14 shows that depth slices of the distribution of source location probabilities for the direct P-wave.

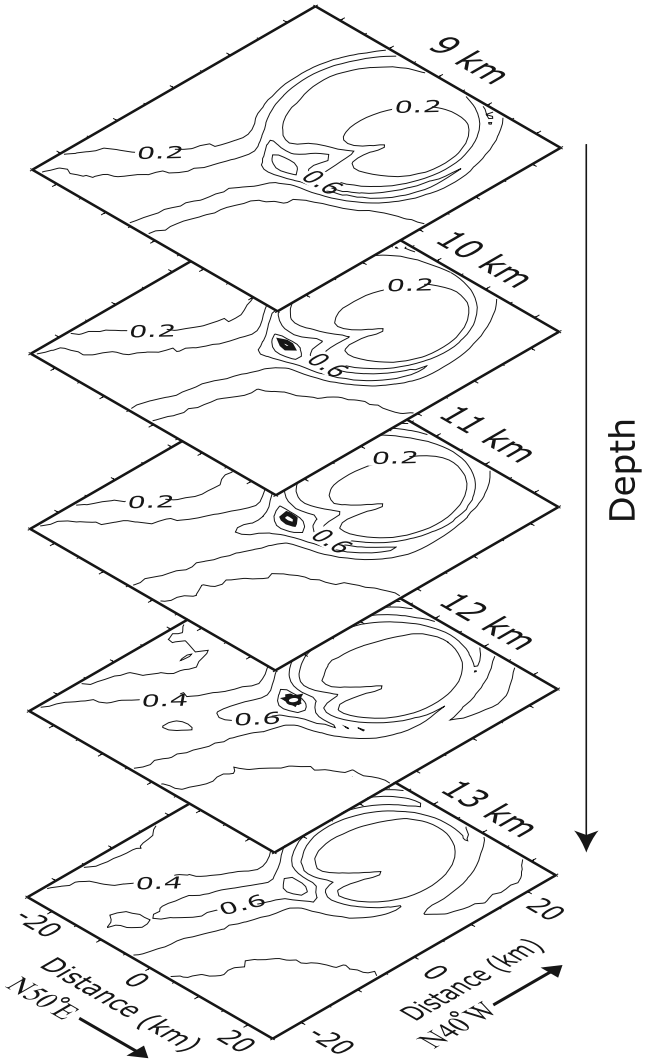

Figure 16. Depth slices of the distribution of the *P-P* scattering coefficients for the synthetic seismograms in Figure 15. Open circle and triangle represent the estimated and theoretical (i.e., the assigned *P-P* scatterer location) maximum points of the scattering coefficients, respectively. The bold contour corresponds to the fit function values with greater than 0.95. The *P-P* scattering coefficients are normalized by its maximum value.

The uncertainty in the source location is evaluated by equation (38). We consider the spreading of uncertainty in locations by the fit function greater than $\alpha = 0.95$. The maximum source location probability is obtained approximately a half kilometer east of its true shot location. The size of the uncertainty is about 3 km in horizontal and 1 km in vertical, so we can conclude that our approach succeeded in locating the shotpoint. Note that we only used the seismograms in the vertical component because our imaging approach can estimate, in principle, the source location by using one component seismograms recorded for a source-array pair.

[81] We also performed an numerical experiment with synthetic seismograms, in order to exam the capability of locating scatterers. Since we are interesting in heterogeneous structures around earthquake rupture zones in our application (Paper II), we placed one *P-P* and *P-S* scatterers at the coordinate center and 12 km depth where the $M = 5.2$ earthquake occurred (see Paper II). Each scatterer generates 10 Hz Ricker wavelet with the same amplitude. We synthesized three-component seismograms for the same shot and array as the actual test data (Figure 15), assuming the single isotropic scattering model [e.g., *Sato*, 1977]. Gaussian white noise 20% of the average signal level were added to the synthetic seismograms. Note that we assume that displacement of *SV* wave is the same as one of *SH* wave. The parameters for the *f-k* and polarization analyses were the same as those in the test for the direct *P*-wave mentioned above.

[82] Figures 16 and 17 show depth slices of the distributions of *P-P* and *P-S* scattering coefficients for this numerical experiment. We found the *P-P* scatterer is located at the assigned point, and the size of the uncertainty to be approximately 2 km in horizontal and 2 km in vertical. On the other hand, the *P-S* scatterer is located 1 km above the assigned point, and the size of uncertainty is the same as one in the *P-P* scatterer. These results show that our imaging approach is robust enough to locate scatterers.

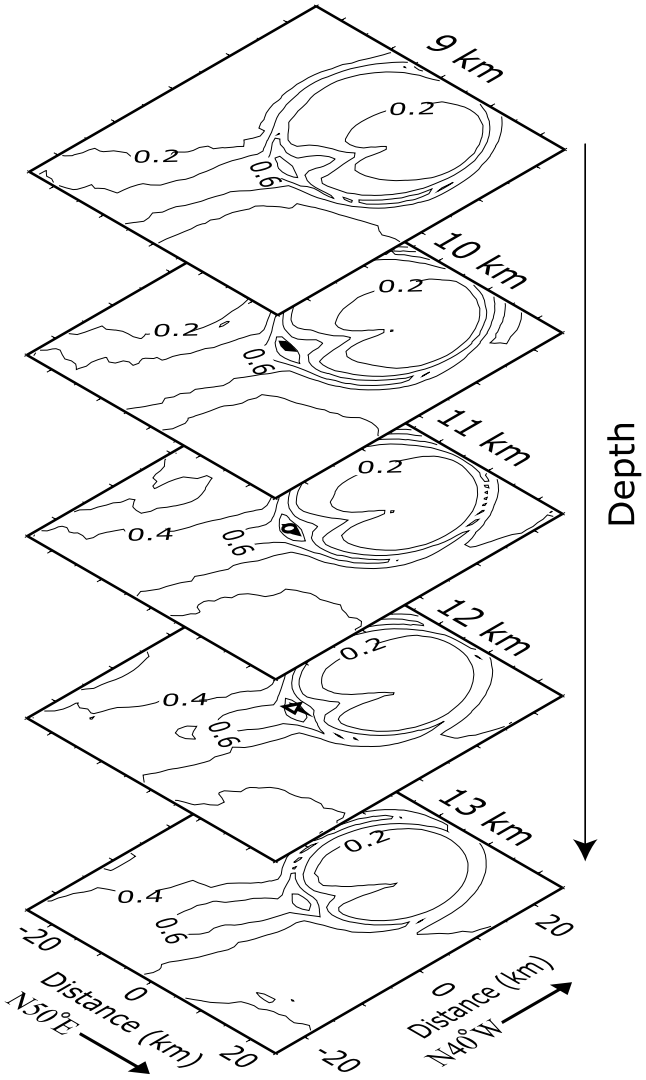

Figure 17. Same as Figure 16 except for *P-S* scattering coefficients.

6. Conclusions

[83] We have established a new imaging framework for small-scale heterogeneities in the crust, from the points of view of seismic scattering with the use of 3-C seismic array data. We have additionally extended the conventional analyses of coda waves. The key elements in our present development are as follows: (1) propagation effects of a complex background medium (i.e., anelastic attenuation) are corrected by a statistical amplitude recovery based on the AIC, (2) scattered phases are determined with high coherency and high linearity, by performing the *f-k* and polarization analyses with stationary AR and MAR models, respectively, (3) scattering modes of the detected coherent phases are estimated by combining slowness and polarization vectors, (4) source and station effects are corrected by a coda-normalization approach, and (5) *f-k* power spectra are mapped as scattering coefficients based on the travel times and slowness vectors of coherent phases with a three-dimensional velocity structure.

[84] One of our ultimate goals for studies on small-scale heterogeneities in the crust is to reveal physical properties of seismic scatterers in a quantitative manner, that is, to identify the materials responsible for observed scattered waves. The present approach is definitely compensates several kinds of weakness in conventional approaches of imaging small-scale heterogeneities. Let us summarize the advantages of the present imaging approach in terms of the above goal.

[85] 1. Previous studies on imaging of small-scale heterogeneities did not deal with complex propagation effects of high-frequency seismic waves such as intrinsic attenuation carefully. For example, these studies adopted a given decay curve of seismograms. On the contrary, our approach attempts to estimate and remove the complex propagation effect of each seismogram both statistically and quantitatively, on the basis of the AIC.

[86] 2. To improve resolution in *f-k* power spectra and covariance matrices in the time-frequency domain, we apply stationary AR and MAR models to *f-k* and polarization analyses. The improved resolution enables us to separate weak coherent signals of scattering (i.e., scattered phases) from superimposed noise.

[87] 3. By taking the advantage of a 3-C seismic array, we evaluate the scattering mode of each scattered phase. The information of the scattering mode will enable us to make more profound discussions on the nature of heterogeneities than ever.

[88] 4. We correct frequency-dependent source and station effects based on a coda-normalization approach with coda waves of large lapse time.

[89] 5. *F-k* power spectra of scattering phases are mapped as scattering coefficients in the model space accurately and stably by using their well-constrained slowness vectors with a three-dimensional velocity structure.

[90] In the companion paper, Paper II, we shall apply the present approach to the 3-C seismic array data sets recorded around the Nagamachi-Rifu fault, northeastern Japan, area, in order to observe how it works effectively, together with how powerful the present development is in obtaining information on seismic activities and tectonics in this region, as an concrete example.

[91] **Acknowledgments.** We wish to express our gratitude to Yasuto Kuwahara, Hisao Ito, and Kazutoshi Imanishi at Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology, for reading the draft and making a number of helpful suggestions. We thank Junji Koyama, Takeo Moriya, Tsutomu Sasatani, and Tetsuo Takanami at Hokkaido University for their valuable and helpful advice and suggestions. We would also appreciate many suggestions on *f-k* and polarization analyses from Kazunori Yoshizawa at Hokkaido University. Valuable comments by Justin Revenaugh at University of Minnesota, Charles K. Wilson at Lamont-Doherty Earth Observatory of Columbia University, and anonymous reviewers are greatly appreciated. A software package, Generic Mapping Tools (GMT), [*Wessel and Smith*, 1995] was used to plot some figures. This work was partially supported by the Earthquake Research Institute Cooperative Research Programs (2000-B-07, 2001-B-02, and 2003-B-04). The second author (K. Y.) was also supported in part by the bilateral treaty between the Royal Society and Japan Society of Promotion of Science.

References

Abers, G. A. (1998), Array measurements of phases used in receiver-function calculations: Importance of scattering, *Bull. Seismol. Soc. Am.*, *88*, 313–318.

Akaike, H. (1969), Fitting autoregressive models for prediction, *Ann. Inst. Stat. Math.*, *21*, 243–247.

Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, in *2nd International Symposium on Information Theory*, edited by B. N. Petrov and F. Csaki, pp. 267–281, Akad. Kiado, Budapest.

Akaike, H., and G. Kitagawa (Eds.) (1994), *The Practice of Time Series Analysis*, Springer, New York.

Aki, K. (1969), Analysis of seismic coda of local earthquakes as scattered waves, *J. Geophys. Res.*, *74*, 615–631.

Aki, K., and B. Chouet (1975), Origin of coda waves: Source, attenuation and scattering effects, *J. Geophys. Res.*, *80*, 3322–3342.

Aki, K., and P. G. Richards (1980), *Quantitative Seismology, Theory and Methods*, W. H. Freeman, San Francisco, Calif.

Aki, K., A. Christoffersson, and E. S. Husebye (1976), Three dimensional seismic structure of the lithosphere under Montana LASA, *Bull. Seismol. Soc. Am.*, *66*, 501–524.

Almendros, J., B. Chouet, and P. Dawson (2001), Spatial extent of a hydrothermal system at Kilauea Volcano, Hawaii, determined from array analyses of shallow long-period seismicity: 1. Method, *J. Geophys. Res.*, *106*, 13,565–13,580.

Asano, Y., and A. Hasegawa (2004), Imaging the fault zones of the 2000 western Tottori earthquake by a new inversion method to estimate three-dimensional distribution of the scattering coefficient, *J. Geophys. Res.*, *109*, B06306, doi:10.1029/2003JB002761.

Bear, L. K., G. L. Pavlis, and G. H. R. Bokelmann (1999), Multi-wavelet analysis of three-component seismic arrays: Application to measure effective anisotropy at Piñion Flats, California, *Bull. Seismol. Soc. Am.*, *89*, 693–705.

Bostock, M. G., S. Rondenay, and J. Shragge (2001), Multiparameter two-dimensional inversion of scattered teleseismic body waves: 1. Theory for oblique incidence, *J. Geophys. Res.*, *106*, 30,771–30,782.

Burg, J. P. (1967), Maximum entropy spectral analysis, paper presented at the 37th Annual International SEG Meeting, Soc. of Explor. Geophys., Tulsa, Okla.

Capon, J. (1969), High-resolution frequency-wavenumber spectrum analysis, *Proc. IEEE*, *57*, 1408–1418.

Chávez-Pérez, S., and J. N. Louie (1998), Crustal imaging in southern California using earthquake sequences, *Tectonophysics*, *286*, 223–236.

Claerbout, J. F., and S. M. Doherty (1972), Downward continuation of moveout-corrected seismograms, *Geophysics*, *37*, 741–768.

Del Pezzo, E., M. La Rocca, and J. Ibanez (1997), Observation of high-frequency scattered waves using dense arrays at Teide Volcano, *Bull. Seismol. Soc. Am.*, *87*, 1637–1647.

Dodge, D. A., and G. C. Beroza (1997), Source array analysis of coda waves near the 1989 Loma Prieta, California, mainshock: Implications for the mechanism of coseismic velocity changes, *J. Geophys. Res.*, *102*, 24,437–24,458.

Fehler, M. C., and L. J. Huang (2002), Modern imaging using seismic reflection data, *Annu. Rev. Earth Planet Sci.*, *30*, 259–284, doi:101.1146/annurev.earth.30.091201.140909.

Frankel, A., S. Hough, P. Friberg, and R. Busby (1991), Observations of Loma Prieta aftershocks from a dense array in Sunnyvale, California, *Bull. Seismol. Soc. Am.*, *80*, 1900–1922.

Gazdag, J. (1978), Wave equation migration with the phase-shift method, *Geophysics*, *43*, 1342–1351.

Goldstein, P., and R. J. Archuleta (1991), Deterministic frequency-wavenumber method and direct measurements of rupture propagation during earth-

quakes using a dense array: Theory and methods, *Bull. Seismol. Soc. Am.*, *96*, 6173–6185.
Gupta, I. N., C. S. Lynnes, and R. A. Wagner (1990), Broadband *F-K* analysis of array data to identify sources of local scattering, *Geophys. Res. Lett.*, *17*, 183–186.
Hasegawa, A., H. Ito, T. Iwasaki, and T. Ikawa (2001), Deep structure of Nagamachi-Rifu fault as inferred from seismic expeditions, paper presented at International Symposium on Slip and Flow Processes In and Below the Seismogenic Region, Min. of Educ., Culture, Sports, Sci. and Technol., Sendai, Japan.
Huang, L. J., M. C. Fehler, P. M. Roberts, and C. C. Burch (1999a), Extended local Rytov migration method, *Geophysics*, *64*, 1535–1545.
Huang, L. J., M. C. Fehler, and R. S. Wu (1999b), Extended local Born migration method, *Geophysics*, *64*, 1524–1534.
Ishii, M., P. M. Shearer, H. Houston, and J. E. Vidale (2005), Rupture extent, duration, and speed of the 2004 Sumatra-Andaman earthquake imaged by the Hi-New array, *Nature*, *435*, 933–936, doi:10.1038/nature03675.
Iwata, T., and K. Irikura (1988), Source parameters of the 1983 Japan-Sea earthquake sequence, *J. Phys. Earth*, *36*, 155–184.
Jepsen, D. C., and B. L. N. Kennett (1990), Three component array analysis, *Bull. Seismol. Soc. Am.*, *80*, 2032–2052.
Jurkevics, A. (1998), Polarization analysis of three-component array data, *Bull. Seismol. Soc. Am.*, *78*, 1725–1743.
Kanasewich, E. R. (1973), *Time Sequence Analysis in Geophysics*, Univ. of Alberta Press, Edmonton, Alberta, Canada.
Kitagawa, G., and H. Akaike (1978), Procedure for the modeling of nonstationary time series, *Ann. Inst. Stat. Math.*, *30*, 351–363.
Korneev, V. A., R. M. Nadeau, and T. V. McEvilly (2003), Seismological studies at Parkfield IX: Fault-zone imaging using guided wave attenuation, *Bull. Seismol. Soc. Am.*, *93*, 1415–1426.
LaCoss, R. T., E. J. Kelly, and M. N. Toksöz (1969), Estimation of seismic noise structure using arrays, *Geophysics*, *34*, 21–38.
Lilly, J. M., and J. Park (1995), Multiwavelet spectral and polarization analyses of seismic records, *Geophys. J. Int.*, *122*, 1001–1021.
Matsumoto, S., K. Obara, and A. Hasegawa (1998), Imaging *P*-wave scatterer distribution in the focal area of the 1995 M7.2 Hyogo-ken Nanbu (Kobe) Earthquake, *Geophys. Res. Lett.*, *25*, 1439–1442.
Matsuoka, T., M. Hato, T. J. Ulrych, and T. Kawamura (1986), Optimal gain curve for true amplitude recovery processing (in Japanese with English abstract), *Butsuri-Tansa*, *39*, 26–36.
Montalbetti, J. F., and E. R. Kanasewich (1970), Enhancement of teleseismic body phases with a polarization filter, *Geophys. J. R. Astron. Soc.*, *21*, 119–129.
Nagano, K., H. Niitsuma, and N. Chubachi (1986), A new automatic source location algorithm for downhole tri-axial AE measurement, paper presented at Progress in Acoustic Emission III, Jpn. Soc. for Non-Destruct. Inspect., Tokyo.
Nakamura, H. (1993), Estimation of scatterer distribution in the earth's crust using three-component small-aperture array observation, M.S. thesis, Kyoto Univ., Kyoto, Japan.
Neidell, N. S., and M. T. Taner (1971), Semblance and other coherency measures for multichannel data, *Geophysics*, *36*, 482–497.
Nishigami, K. (1991), A new inversion method of coda waveforms to determine spatial distribution of coda scatterers in the crust and uppermost mantle, *Geophys. Res. Lett.*, *18*, 2225–2228.
Phillips, W. S., and K. Aki (1986), Site amplification of coda waves from local earthquakes in central California, *Bull. Seismol. Soc. Am.*, *76*, 627–648.
Poppeliers, C., and G. L. Pavlis (2002), The seismic response of a steep slope: High-resolution observations with a dense, three-component seismic array, *Bull. Seismol. Soc. Am.*, *92*, 3102–3115.
Rautian, T. G., and V. I. Khalturin (1978), The use of coda for determination of the earthquake source spectrum, *Bull. Seismol. Soc. Am.*, *68*, 923–948.
Revenaugh, J. (1995), A scattered-wave image of subduction beneath the Transverse Range, *Science*, *268*, 1888–1892.
Rost, S., and C. Thomas (2002), Array seismology: Methods and applications, *Rev. Geophys.*, *40*(3), 1008, doi:10.1029/2000RG000100.
Samson, J. C. (1977), Matrix and stokes velocity representations of detectors for polarized waveforms: theory, with some applications to teleseismic waves, *Geophys. J. R. Astron. Soc.*, *51*, 583–603.
Sato, H. (1977), Energy propagation including scattering effects: Single isotropic scattering approximation, *J. Phys. Earth*, *25*, 27–41.
Sato, H. (1984), Attenuation of envelope formation of three-component seismograms of small local earthquakes in randomly inhomogeneous lithosphere, *J. Geophys. Res.*, *89*, 1221–1241.
Sato, H., and M. C. Fehler (1998), *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*, Springer, New York.
Scherbaum, F., F. Krüger, and M. Weber (1997), Double beam imaging: Mapping lower mantle heterogeneities using combinations of source and receiver arrays, *J. Geophys. Res.*, *102*, 507–522.
Schmidt, R. O. (1986), Multiple emitter location and signal parameter estimation, *IEEE Trans. Antennas Propag.*, *34*, 276–280.
Schneider, W. A. (1978), Integral formulation for migration in two and three dimensions, *Geophysics*, *43*, 49–76.
Soma, N., H. Niitsuma, and R. Baria (2002), Reflection technique in time-frequency domain using multicomponent acoustic emission signals and application to geothermal reservoirs, *Geophysics*, *67*, 928–938.
Spudich, P., and T. Bostwick (1987), Studies of the seismic coda using an earthquake cluster as a deep buried seismograph array, *J. Geophys. Res.*, *92*, 10,526–10,546.
Spudich, P., and E. Cranswick (1984), Direct observation of rupture propagation during the 1979 Imperial Valley, California, earthquake using a short-baseline accelerometer array, *Bull. Seismol. Soc. Am.*, *74*, 2083–2114.
Stolt, R. H. (1978), Migration by Fourier transform, *Geophysics*, *43*, 23–48.
Taira, T., and K. Yomogida (2004), Imaging of three-dimensional small-scale heterogeneities in the Hidaka, Japan, region: Coda spectral analysis, *Geophys. J. Int.*, *158*(3), 998–1008, doi:10.1111/j.1356-246X.2004.02333.x
Taira, T., K. Yomogida, Y. Kuwahara, K. Imanishi, and H. Ito (2007), Imaging of crustal heterogeneous structures using a slowness-weighted back-projection with effects of scattering modes: 2. Application of the Nagamachi-Rifu fault, Japan, area, *J. Geophys. Res.*, doi:10.1029/2006JB004382, in press.
Takanami, T. (1991), A study of detection and extraction methods for micro-earthquake waves by autoregressive models, *J. Fac. Sci. Hokkaido Univ. Ser. 7*, *9*, 67–196.
Tsujiura, M. (1978), Spectral analysis of the coda waves from local earthquakes, *Bull. Earthquake Res. Inst. Univ. Tokyo*, *53*, 1–48.
Um, J., and C. Thurber (1987), A fast algorithm for 2-pint seismic ray tracing, *Bull. Seismol. Soc. Am.*, *77*, 972–986.
Vidale, J. E. (1986), Complex polarization analysis of particle motion, *Bull. Seismol. Soc. Am.*, *76*, 1393–1405.
Wagner, G. S. (1997), Regional wave propagation in southern California and Nevada: Observations from a three-component seismic array, *J. Geophys. Res.*, *102*, 8285–8311.
Wagner, G. S. (1998), Local wave propagation near the San Jacinto fault zone, southern California: Observations from a three-component seismic array, *Bull. Seismol. Soc. Am.*, *103*, 7231–7246.
Wagner, G. S., and C. A. Langston (1992), Body-to-surface-wave scattered energy in teleseismic coda observed at the NORESS seismic array, *Bull. Seismol. Soc. Am.*, *82*, 2126–2138.
Wagner, G. S., and T. J. Owens (1996), Signal detection using multi-channel seismic data, *Bull. Seismol. Soc. Am.*, *86*, 221–231.
Wessel, P., and W. H. F. Smith (1995), New version of the Generic Mapping Tools released, *Eos Trans. AGU*, *76*, 329.
Wu, R. S. (1982), Attenuation of short period seismic waves due to scattering, *Geophys. Res. Lett.*, *9*, 9–12.
Wu, R. S. (2003), Wave propagation, scattering and imaging using dual-domain one-way and one-return propagators, *Pure Appl. Geophys.*, *160*, 509–539.
Wu, R. S., and K. Aki (1988), Introduction: Seismic wave scattering in three-dimensionally heterogeneous Earth, *Pure Appl. Geophys.*, *128*, 1–6.
Yilmaz, Ö. (1987), *Seismic Data Processing*, Soc. of Explor. Geophys., Tulsa, Okla.
Yomogida, K., R. Benites, P. M. Robert, and M. Fehler (1997), Scattering of elastic waves in 2-D composite media II. Waveforms and spectra, *Phys. Earth Planet. Inter.*, *104*, 175–192.
Yoshizawa, K., K. Yomogida, and S. Tsuboi (1999), Resolving power of surface wave polarization data for high-order heterogeneities, *Geophys. J. Int.*, *138*, 205–220.
Yule, G. U. (1972), On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers, *Philos. Trans. R. Soc., Ser. A*, *226*, 267–298.

T. Taira, Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, DC 20015, USA. (taira@dtm.ciw.edu)
K. Yomogida, Division of Earth and Planetary Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan. (yomo@mail.sci.hokudai.ac.jp)