Locating Spatial Changes of Seismic Scattering Property by Sparse Modeling of Seismic Ambient Noise Cross-Correlation Functions: Application to the 2008 Iwate-Miyagi Nairiku (Mw 6.9), Japan, Earthquake

Takashi Hirose, Hisashi Nakahara, Takeshi Nishimura, Michel Campillo

To cite this version:

Takashi Hirose, Hisashi Nakahara, Takeshi Nishimura, Michel Campillo. Locating Spatial Changes of Seismic Scattering Property by Sparse Modeling of Seismic Ambient Noise Cross-Correlation Functions: Application to the 2008 Iwate-Miyagi Nairiku (Mw 6.9), Japan, Earthquake. Journal of Geophysical Research : Solid Earth, 2020, 125 (6), pp.e2019JB019307. 10.1029/2019JB019307 . hal-02928306

HAL Id: hal-02928306
https://hal.univ-grenoble-alpes.fr/hal-02928306
Submitted on 2 Sep 2020

JGR Solid Earth

RESEARCH ARTICLE
10.1029/2019JB019307

Key Points:
- We detected seismic scattering property changes associated with the 2008 Iwate-Miyagi Nairiku, Japan, earthquake
- The largest change was located to the south of the epicenter using 17 seismic stations
- This change region was also retrieved from only five seismic stations by using sparse modeling

Supporting Information:
- Supporting Information S1
- Table S1

Correspondence to:
T. Hirose,
takashi.hirose@bosai.go.jp

Citation:
Hirose, T., Nakahara, H., Nishimura, T., & Campillo, M. (2020). Locating spatial changes of seismic scattering property by sparse modeling of seismic ambient noise cross-correlation functions: Application to the 2008 Iwate-Miyagi Nairiku (*Mw* 6.9), Japan, earthquake. *Journal of Geophysical Research: Solid Earth*, *125*, e2019JB019307. https://doi.org/10.1029/2019JB019307

Received 26 DEC 2019
Accepted 12 MAY 2020
Accepted article online 21 MAY 2020

Locating Spatial Changes of Seismic Scattering Property by Sparse Modeling of Seismic Ambient Noise Cross-Correlation Functions: Application to the 2008 Iwate-Miyagi Nairiku (*Mw* 6.9), Japan, Earthquake

Takashi Hirose[1], Hisashi Nakahara[2], Takeshi Nishimura[2], and Michel Campillo[3]

[1]National Research Institute for Earth Science and Disaster Resilience, Tsukuba, Japan, [2]Department of Geophysics, Graduate School of Science, Tohoku University, Sendai, Japan, [3]University Grenoble Alpes, CNRS, ISTerre, Grenoble, France

Abstract Locating change regions of seismic velocities and seismic scattering properties associated with volcanic activities and earthquakes is important for structural monitoring. To increase such applications, we propose to use sparse modeling to estimate spatial distributions of seismic scattering property changes. The sparse modeling is an inversion technique that enables us to estimate model parameters from a small data set with sparsity condition such as ℓ^1 norm regularization. We apply this technique to seismic ambient noise cross-correlation functions from 17 Hi-net stations around the epicenter of the 2008 Iwate-Miyagi Nairiku, Japan, earthquake ($Mw = 6.9$). We compute waveform decoherences at the 0.5–1 Hz band and invert the waveform decoherences for the spatial distributions of seismic scattering property changes. Just after the main shock, the largest change occurred at the south of the epicenter, and the maximum change of the scattering coefficient in this region is estimated to be 0.032 km^{-1}. The result from an ordinary linear least squares inversion with the ℓ^2 norm regularization is almost consistent with that from the sparse modeling. Moreover, we confirm the superiority of sparse modeling in imaging with smaller data sets. Only five seismic stations that are deployed near the epicenter so as to surround the change regions are necessary to retrieve the result from 17 stations. On the other hand, in the case of the ℓ^2 norm regularization, we need at least 15 stations. The sparse modeling will be helpful to estimate the spatial distribution of seismic scattering property changes from a small data set.

1. Introduction

Seismic interferometry is a technique to estimate changes of seismic velocity and/or seismic scattering property using cross-correlation functions (CCFs) of seismic ambient noise (e.g., Curtis et al., 2006; Shapiro, 2004). In recent years, seismic velocity changes associated with volcanic activities and earthquakes have been estimated by many previous studies, and a dense seismic network enables us to locate change regions of seismic velocities (e.g., Brenguier et al., 2008, 2014; Nimiya et al., 2017; Wang et al., 2019). As an example at a volcano, Brenguier et al. (2008) succeeded in detecting seismic velocity decreases of about 0.1% preceding eruptions to the east of the crater at Piton de la Fournaise, Reunion Island, and attributed these changes to dilatation of the edifice due to pressure increases in the magma chamber. As a recent example at an earthquake source region, Nimiya et al. (2017) estimated spatiotemporal seismic velocity changes during the 2016 *Mw* 7.0 Kumamoto earthquake, Japan, and the regions of large seismic velocity decreases were located near the rupture faults and Aso volcano. Their interpretation is as follows: The velocity decreases near the faults (0.3–0.4 %) might have been due to damages in regions shallower than 10 km, a change in the stress state, and/or an increase in pore pressure. The seismic velocity reduction observed at Aso volcano (0.7–0.8%) was likely caused by pressurized volcanic fluids. Seismic scattering property changes cause waveform changes of seismic ambient noise CCFs. There are still only a few previous studies based on this principle. Spatial distributions of seismic scattering property changes associated with volcanic activities and earthquakes have been estimated by using CCFs and linear least squares inversions using sensitivity kernels (e.g., Hillers et al., 2015; Machacca-Puma et al., 2019; Obermann et al., 2014, 2013). Obermann et al. (2013) succeeded in detecting changes in seismic scattering property around the crater of the October 2010 eruption at Piton de la Fournaise, Reunion Island. Obermann et al. (2014) found changes of seismic scattering property associated with the 2008 Wenchuan, China, earthquake. They succeeded in locating decoherence, a measure of dissimilarity between two waveforms and is defined as 1-(coherence), around the earthquake epicenter using seismic ambient noise CCFs in the 1–3 s period band. They interpreted that the coseismic changes can be directly related to severe damages at shallow depths caused by the main shock and many aftershocks. Obermann et al. (2014) is the only previous study that estimated spatial distributions of seismic scattering property changes associated with a large earthquake by applying seismic interferometry analysis. To understand seismic scattering property changes, it is necessary to increase the number of applications.

Although locating change regions of seismic velocities and seismic scattering properties is important for structural monitoring, one limiting factor is the number of seismic stations. Machacca-Puma et al. (2019) estimated spatial distributions of seismic velocity and seismic scattering property changes at Ubinas volcano, Peru, only from six seismic stations. However, a large number of seismic stations are thought to be necessary to conduct imagings in general. For example, Obermann et al. (2013) used 19 broadband seismic stations that are deployed in the area with about 15 km of the east-west and 10 km in the north-south. Obermann et al. (2014) used 114 seismic stations in the target area (29°N to 32°N and 100°E to 105°E). Conducting such estimations is not easy in regions where only a small number of stations are set up.

The ℓ^1 norm regularization approaches have been applied in reflection seismology problems for denoising in the last four decades (e.g., Claerbout & Muir, 1973; Santosa & Symes, 1986). Sparse modeling is a technique to estimate model parameters from a small data set with sparsity constraints. ℓ^1 norm regularization is one of the approaches in the sparse modeling, and Tibshirani (1996) proposed a “least absolute shrinkage and selection operator (LASSO)” that is a representative method to solve a linear regression problem by using ℓ^1 norm regularization. LASSO has been widely used in sparse modeling. In recent years, this technique has been used for earthquake source inversions or geodetic inversions (e.g., Evans & Meade, 2012; Nakata et al., 2016; Yao et al., 2011). Yao et al. (2011) estimated the distribution of source powers during the rupture of the 2011 Tohoku-Oki, Japan, earthquake from teleseismic *P* waves recorded by an array of stations in the United States. Their estimation results of the distribution of source powers at 0.5–1 and 0.2–0.5 Hz bands shared similar features as back projection results in other previous studies that revealed dominant seismic energy radiation in the downdip. They also confirmed that the distribution of source powers at 0.1–0.2 and 0.05–0.1 Hz bands was almost consistent with slip inversion results from seismic data and geodetic data that revealed large slip patches dominantly close to the trench. Evans and Meade (2012) estimated spatial distributions of coseismic slip and afterslip associated with the 2011 Tohoku-Oki earthquake using Global Navigation Satellite System (GNSS) data and the sparse modeling. They compared results from the sparse modeling with those from the damped least squares inversion and showed that the sparse modeling recovered more compact and sharply varying slip distributions than those from the damped least squares inversion. Although sparse modeling has been used in various kinds of analyses, there is no previous study that applies this technique to estimate spatial distributions of seismic scattering property changes.

We apply seismic interferometry to seismic ambient noise data before and after the 2008 *Mw* 6.9 Iwate-Miyagi Nairiku, Japan, earthquake, and estimate the spatial distribution of seismic scattering property changes from decoherences of seismic ambient noise CCFs by using the sparse modeling. For comparison, we also apply an ordinary linear least squares inversion with ℓ^2 norm regularization to locate change regions of seismic scattering properties. To interpret possible causes of seismic scattering property changes associated with the Iwate-Miyagi Nairiku earthquake, we compare the results with other observations. Finally, we discuss the minimum number of seismic stations and their appropriate configurations to retrieve a reliable result. We conduct imaging by using both of sparse modeling and ℓ^2 norm regularization by changing the numbers and distributions of seismic stations and clarify the superiority of the sparse modeling in the imaging with a small data set.

2. Data and Methods

2.1. Data

The 2008 *Mw* 6.9 Iwate-Miyagi Nairiku earthquake occurred on 14 June 2008 (Japan Standard Time) in the central part of the Tohoku region in northeast Japan. The depth of the hypocenter was approximately 6 km, and the focal mechanism of the main shock was a reverse fault type. The hypocenter of this earthquake

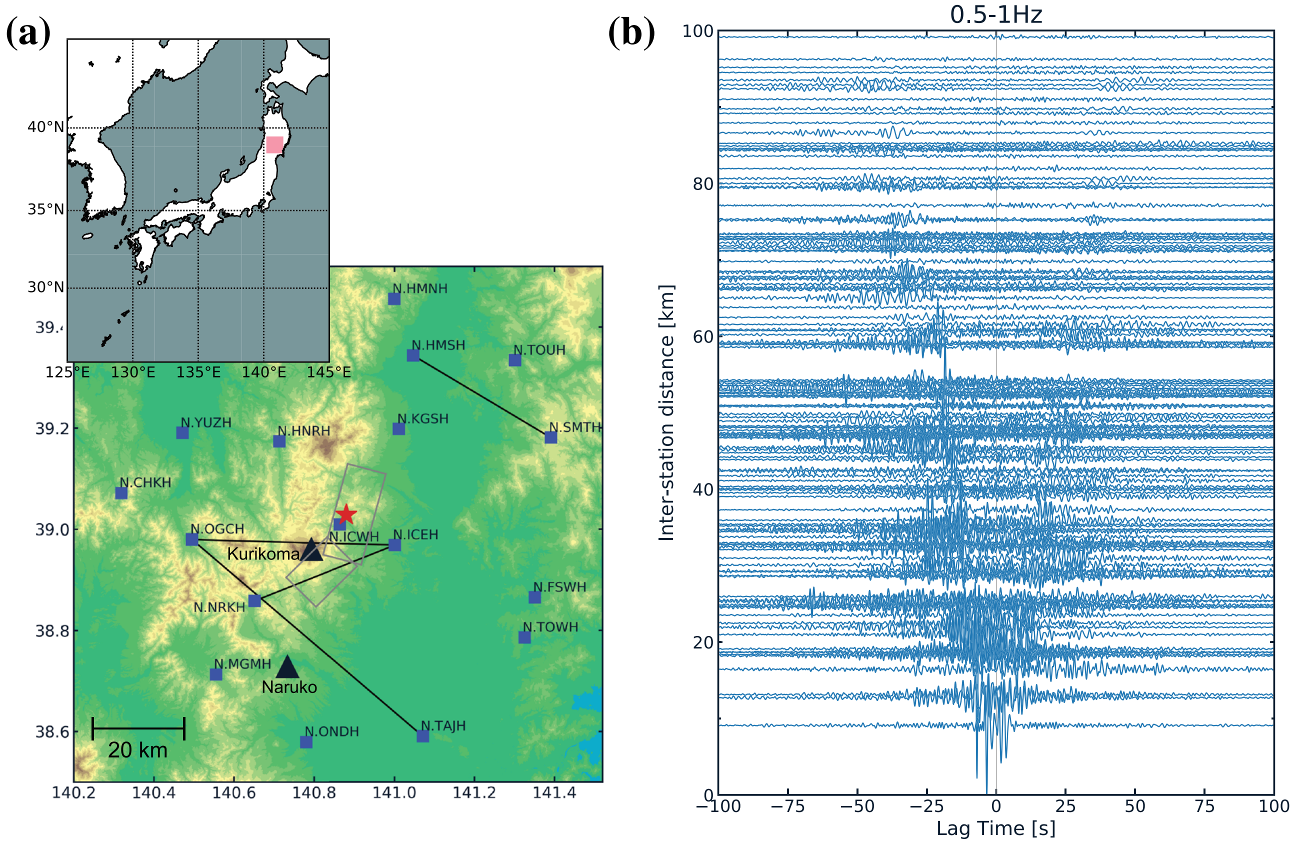

Figure 1. (a) Study area (pink in the inset), locations of 17 Hi-net stations (blue squares in the closeup) and the epicenter of the 2008 Iwate-Miyagi Nairiku, Northeast Japan, earthquake (red star in the close-up). The two subfaults estimated by Ohta et al. (2008) are shown by gray rectangles. The two black triangles represent active volcanoes, Kurikoma and Naruko. (b) Record section of seismic ambient noise CCFs in the 0.5–1 Hz band. These CCFs are computed by stacking daily CCFs over 1 year in 2007.

is located among the dense seismic network, Hi-net, which is operated by the National Research Institute for Earth Science and Disaster Resilience (NIED) (National Research Institute for Earth Science and Disaster Resilience, 2019). Strong ground motions were observed at accelerometers that were deployed near the epicenter; the peak ground acceleration (PGA) reached approximately 4,000 gal (Aoi et al., 2008). We use seismograms on the vertical component during the years of 2007 and 2008 from 17 Hi-net stations. The locations of seismic stations are shown by blue squares in Figure 1a. A three-component short-period seismometer with a natural frequency of 1 Hz is installed at each station. All the seismograms are recorded with a sampling frequency of 100 Hz.

2.2. Calculation Procedure of Seismic Ambient noise CCFs

First, we filter seismic ambient noise records at the 0.5–1 Hz band and select 10 min long data having amplitudes of smaller than 10 times the average root-mean-square (RMS) amplitude of seismic ambient noise recorded in 2007. Second, we apply temporal flattening technique (Weaver, 2011) and spectral whitening (e.g., Bensen et al., 2007; Shapiro et al., 2006) for suppressing the influence of earthquakes and persistent monochromatic noise sources on seismic ambient noise CCFs. Third, we compute CCFs every 10 min and calculate daily CCFs (hereafter called DCCFs) by stacking these 10-min CCFs every day for 136 pairs from 17 stations. The obtained CCFs are represented to be Green's functions recorded at one of the station pairs from a virtual source at the other station of the pair. Figure 1b is a record section of CCFs in the 0.5–1 Hz band. The Rayleigh waves that propagate with 3 km/s are seen. These CCFs are computed by stacking DCCFs over 1 year in 2007. In this study, we use these 1 year stacked CCFs as reference CCFs (hereafter called RCCFs). Figure 2 shows examples of temporal changes in daily coherence values over 2 years for four station pairs (N.ICEH-N.NRKH, N.OGCH-N.ICEH, N.HMSH-N.SMTH, and N.OGCH-N.TAJH). These coherence

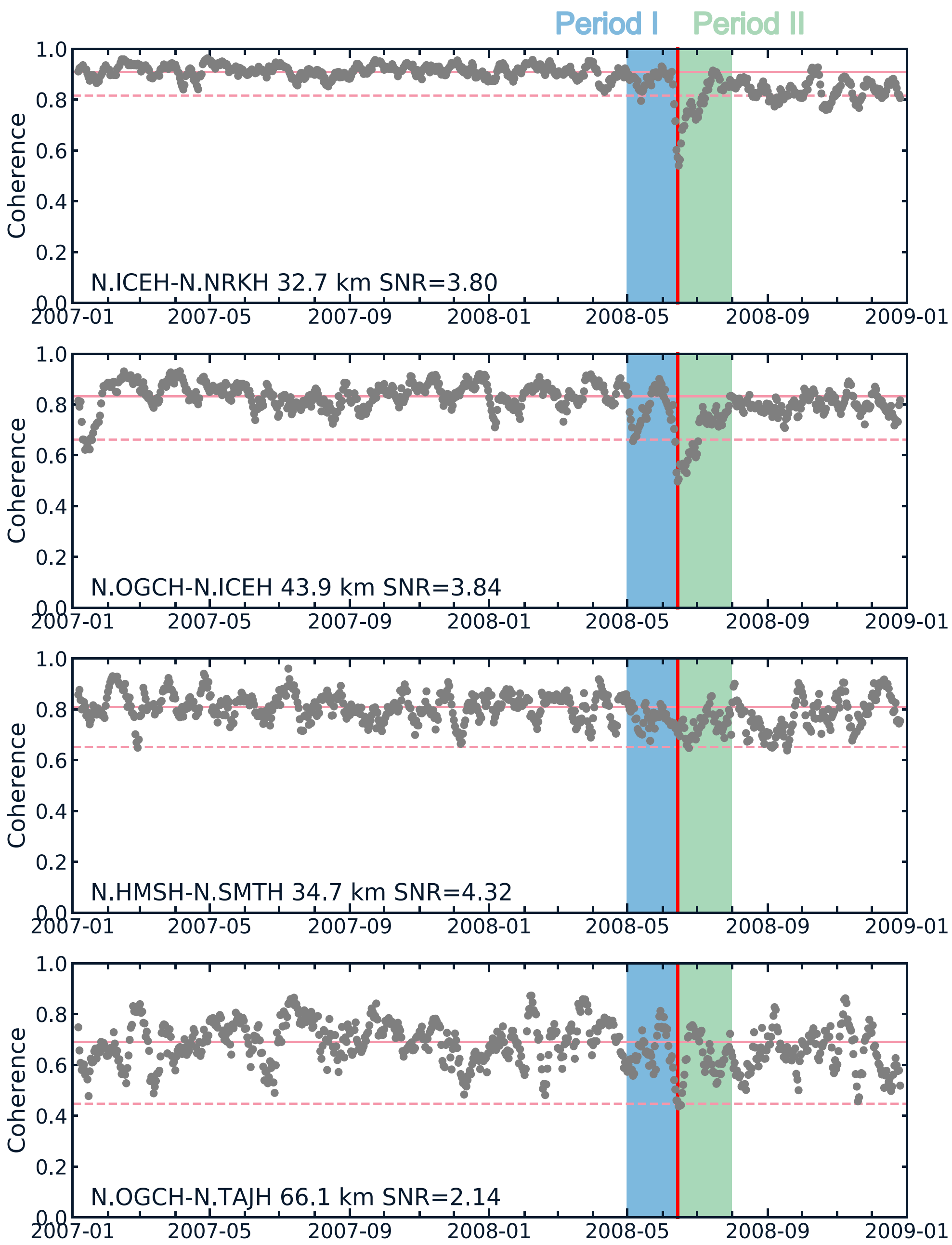

Figure 2. Temporal changes of coherences over 2 years for station pairs of N.ICEH-N.NRKH (first panel), N.OGCH-N.ICEH (second panel), N.HMSH-N.SMTH (third panel), and N.OGCH-N.TAJH (fourth panel). These station pairs are shown by black solid lines in Figure 1a. Coherence values are computed between the CCF stacked over 1 year in 2007 and 10 days stacked CCFs by sliding 10 days long time windows by 1 day. The red lines represent the day of the Iwate-Miyagi Nairiku earthquake. The pink solid lines represent mean coherence values before the earthquake (1 January 2007 to 13 June 2008), and dashed ones represent the mean coherence minus 3 times standard deviations (3σ) of coherence values. SNR represents the mean signal-to-noise ratio of the 10 days stacked CCFs.

values are computed between the RCCF and 10 days stacked CCFs with sliding 10 days long time windows by 1 day. In this study, the coherence value is computed in the frequency domain using the relation, $Coh_{xy}(f) = \left|P_{xy}(f)\right|^2 / P_{xx}(f)P_{yy}(f)$, where $Coh_{xy}(f)$ is the coherence value between signals x and y at a frequency f. $P_{xx}(f)$ and $P_{yy}(f)$ are power spectral densities of x and y, respectively, and $P_{xy}(f)$ is the cross-spectral density of x and y. In this study, we use the mean values of $Coh_{xy}(f)$ between the 0.5 and 1 Hz band. The significant decoherences were detected associated with the Iwate-Miyagi Nairiku earthquake (red lines). The pink solid lines represent mean coherence values before the earthquake (1 January 2007 to 13 June 2008), and pink dashed lines represent the mean coherence minus 3 times standard deviations (3σ). The station pairs of N.ICEH-N.NRKH and N.OGCH-N.ICEH show a sharp drop of coherences associated with the earthquake (see top two panels). The coherences gradually increased after the earthquake and recovered to around the mean coherences within 1 month after the earthquake. Though seasonal effects may exist, we cannot see significant decoherences, exceeding mean coherence minus 3σ, in the same season in 2007. Therefore, the significant decoherences in June 2008 should have been caused by the Iwate-Miyagi Nairiku earthquake. In contrast, we do not see such significant decoherences for the station pairs of N.HMSH-N.SMTH and N.OGCH-N.TAJH, which do not cross around the epicenter or both of the stations located far from the epicenter (see bottom two panels). Coherence for the station pair of N.OGCH-N.TAJH shows large temporal fluctuation. We also show mean signal-to-noise ratios (SNRs) of 10 days stacked CCFs for each station pair. A root-mean-square (RMS) amplitude of a CCF from +90 to +100 s and from -100 to -90 s in lag times is assumed as a noise part and that from 16 to 64 s after the direct Rayleigh wave arrival time is assumed as a signal part. Mean SNR of 10 days stacked CCFs for this station pair is a little bit lower than those for other station pairs. This can be because the interstation distance of this station pair is longer compared to the other station pairs. The large temporal fluctuation of coherences for this station pair can be caused by the low SNRs of CCFs.

2.3. Imaging Method of Seismic Scattering Property Changes

First, we stack DCCFs over two periods. The first period is before the Iwate-Miyagi Nairiku earthquake, 1 May 2008 to 13 June 2008 (hereafter called Period I), and the second one is after the earthquake, 15 June 2008 to 31 July 2008 (Period II). These two periods are shown in Figure 2 by blue and green shaded areas, respectively. After that, we compute decoherence values between RCCFs and stacked CCFs in the two periods, respectively. As already mentioned in section 1, the decoherence is defined as 1-(coherence). These decoherence values are measured by sliding a 16 s long short time windows from 16 to 64 s after the direct Rayleigh wave arrival time without overlapping. This is because Obermann et al. (2016) calculated the lapse time dependence of energy partition ratio between body waves and surface waves, and they pointed out that surface waves were predominant until 6 to 8 times the mean-free time in lapse time. In this study, we assume the spatially uniform Rayleigh wave velocity to be 3 km/s. This is justified because, according to Nishida et al. (2008), the spatial variation of S wave velocities at a depth of 2 km is a few percent in our study area. The scattering mean-free path in our study area is estimated to be 30 km (Appendix A). In this case, the mean-free time is 10 s. Therefore, we assume that surface waves are predominant for most of the short time windows. We select those decoherence values according to SNRs of stacked CCFs in Period II. We only use decoherence values from time windows whose SNRs are larger than 1.5.

After the data selection, we take differences of decoherence values in Period I ($dc_1^{\mathrm{obs}}(t)$) and Period II ($dc_2^{\mathrm{obs}}(t)$):

$$dc_{12}^{\mathrm{obs}}(t) = \left| dc_2^{\mathrm{obs}}(t) - dc_1^{\mathrm{obs}}(t) \right| . \tag{1}$$

Here, t is a lag time. As shown in Figure 2, the coherence levels before the earthquake are different for different station pairs. For example, the coherences are low even before the earthquake for the station pair of N.OGCH-N.TAJH whose interstation distance is as long as 66.1 km. Therefore, we use differences of decoherence values before and after the earthquake to focus on the coseismic changes of decoherence. A decoherence value represents changes in waveforms. We cannot know whether the decoherence is caused by an increase or decrease of the scattering coefficient. Hence, in this study, we use absolute values of differences in decoherence values.

Decoherence values due to a scattering property change at a point is expressed using the following relationship (e.g., Larose et al., 2015; Margerin et al., 2016):

$$dc_{12}^{\mathrm{syn}}(t) = \frac{c\Delta g_{12}(\mathbf{x})}{2} K_{\mathrm{dc}}(\mathbf{x}, t; \mathbf{x}_R, \mathbf{x}_S). \tag{2}$$

Here, $dc_{12}^{\mathrm{syn}}(t)$ is a synthesized decoherence value at the lag time t, c is the Rayleigh wave velocity (assumed to be 3 km/s in this study), and $\Delta g_{12}(\mathbf{x})$ is an amount of a change in scattering coefficient at a location $\mathbf{x}$ between Periods I and II. Scattering coefficient of Rayleigh wave means scattering power per unit area and is given by the product of the number density of scatterers and the differential scattering cross section. Small-scale

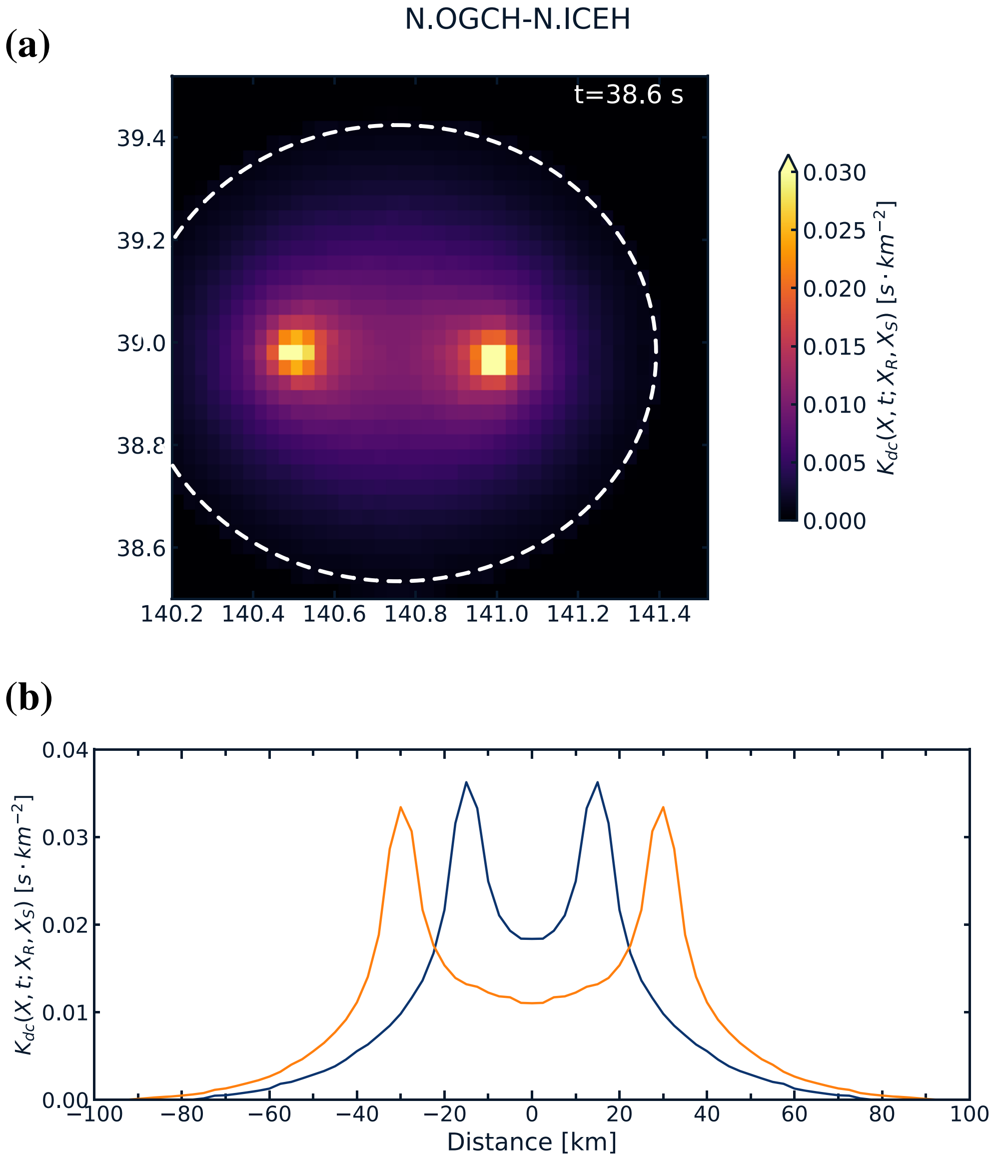

Figure 3. (a) An example of the decorrelation kernel for the station pair of N.OGCH-N.ICEH at 38.6 s in lag time. The white dashed ellipse represents a single scattering ellipse. The kernel is calculated assuming the scattering mean-free path of Rayleigh wave to be 30 km. This parameter is estimated from envelopes of seismic ambient noise CCFs (see Appendix A). (b) Examples of cross sections of sensitivity kernels for two different interstation distances. Blue and orange solid lines represent the kernels for the station pairs whose interstation distances are 30 and 60 km, respectively. They are calculated at a lapse time of 40 s after the direct Rayleigh wave arrivals.

fluctuations in mass density and seismic velocities work as scatterers (e.g., Sato et al., 2012). $K_{\mathrm{dc}}(\mathbf{x}, t; \mathbf{x}_R, \mathbf{x}_S)$ is a decorrelation kernel (Margerin et al., 2016). $\mathbf{x}$, $\mathbf{x}_S$, and $\mathbf{x}_R$ are the locations of a change region, source, and receiver, respectively. The decorrelation kernel is the spatial weighting function for locating seismic scattering property changes and is calculated by using the following equation:

$$K_{\mathrm{dc}}\left(\mathbf{x}, t; \mathbf{x}_R, \mathbf{x}_S\right) = \frac{\int_0^t I(\mathbf{x}_R, t-t'; \mathbf{x}) I(\mathbf{x}, t'; \mathbf{x}_S) \mathrm{d}t'}{I(\mathbf{x}_R, t; \mathbf{x}_S)}, \tag{3}$$

where $I(\mathbf{x}_R, t; \mathbf{x}_S)$ is the energy propagator calculated from the 2-D radiative transfer model which assumes isotropic scattering and source radiation of scalar waves in an infinite medium (e.g., Shang & Gao, 1988;

Sato, 1993). The energy propagator is described as follows:

$$I(r,t) = \frac{e^{-ctg_0}}{2\pi cr}\delta\left(t-\frac{r}{c}\right) + g_0\frac{e^{g_0\left(\sqrt{c^2t^2-r^2}-ct\right)}}{2\pi\sqrt{c^2t^2-r^2}}H\left(t-\frac{r}{c}\right), \tag{4}$$

where r is the distance between source and receiver, g_0 is the total scattering coefficient, and H is the Heaviside function.

To compute sensitivity kernels, the value of scattering coefficient in a study area is necessary. Hence, we estimate it using envelopes of seismic ambient noise CCFs based on the passive method developed by Hirose et al. (2019). As a result, the total scattering coefficient g_0 (or the scattering mean-free path ℓ) is estimated to be 0.033 km^{-1} (or 30 km) at the 0.5–1 Hz band (Appendix A). Figure 3a shows an example of a sensitivity kernel for the station pair N.OGCH-N.ICEH. The sensitivity kernel shows two peaks around the two seismic stations. This means that if a change region of seismic scattering properties is close to the station, large decoherences should occur. The widths of those two peaks are proportional to the mean-free path around the two stations (Figure 3b). The white dashed ellipse in Figure 3a represents a single scattering ellipse. The sensitivity kernel is computed using energy propagators which represent energy transport from a source to a receiver at a lapse time t. Seismic energies outside the single scattering shell do not contribute to energy transportations from the source to the receiver at a lapse time t. Therefore, values of the sensitivity kernel are zero outside the single scattering shell.

For spatially distributed changes in scattering coefficient, decoherence values are expressed as the contribution from all points of scattering property changes as follows:

$$dc_{12}^{\mathrm{syn}}(t) = \frac{c}{2}\int \Delta g_{12}(\mathbf{x})K_{\mathrm{dc}}(\mathbf{x},t;\mathbf{x}_R,\mathbf{x}_S)\mathrm{d}S(\mathbf{x}). \tag{5}$$

Here $\mathrm{d}S(\mathbf{x})$ is a small surface. Discretizing Equation 5 by dividing the study area into small square cells, we set up a linear inverse problem to estimate scattering property changes (Δg values) at all the square cells as shown in Equations 6 and 7. This can be considered as a tomographic inversion for the spatial distribution of Δg values.

$$\mathbf{d} = \mathbf{Gm}, \tag{6}$$

$$d_i = dc_i^{\mathrm{obs}}, G_{ij} = \frac{c\Delta s}{2}K_{ij}, m_j = \Delta g_j, \tag{7}$$

$\mathbf{d}$ is a data vector for which each component d_i ($i = 1, \ldots, n$) corresponds to observed decoherence values for a given station pair and lag time. n is the total number of measurements. $\mathbf{G}$ is a matrix for which each component G_{ij} corresponds to the sensitivity kernel K_{ij} for a given station pair and lag time weighted by the surface area of the cells Δs and Rayleigh wave velocity c. $\mathbf{m}$ is a model vector for which each component m_j contains the Δg for each small cell j. In this study, we estimate the model vector $\mathbf{m}$ which minimizes the following objective function:

$$F = \frac{1}{2}\|\mathbf{d}-\mathbf{Gm}\|_2^2 + \lambda\|\mathbf{m}\|_1. \tag{8}$$

The first term of the right-hand side of Equation 8 is the sum of squared residuals (SSRs) between observed and synthesized decoherence values. The second term of the right-hand side is the ℓ^1 norm of a model vector, and λ is a hyperparameter. The optimization using ℓ^1 norm regularization (e.g., Tibshirani, 1996) is a means of sparse modeling and is also called "LASSO." Figures 4a and 4b are schematic pictures for the sparse modeling and ℓ^2 norm regularization in a 2-D case, respectively. The blue ellipses represent the contours of SSR, and the red diamond and circle represent the constraint of the ℓ^1 and ℓ^2 norms, respectively. The minimum value of the objective function (green star) is the intersection of the ellipse and the diamond/circle. In the case of sparse modeling, the diamond region is more likely to produce an intersection that has one component of the solution that is zero. This is because a diamond is convex so it is easier to intersect with an ellipse at an apex. In sparse modeling, most of the model parameters are estimated to be zero. This means that we can reduce the number of model parameters, and hence, we can estimate a model vector $\mathbf{m}$ from a small data set.

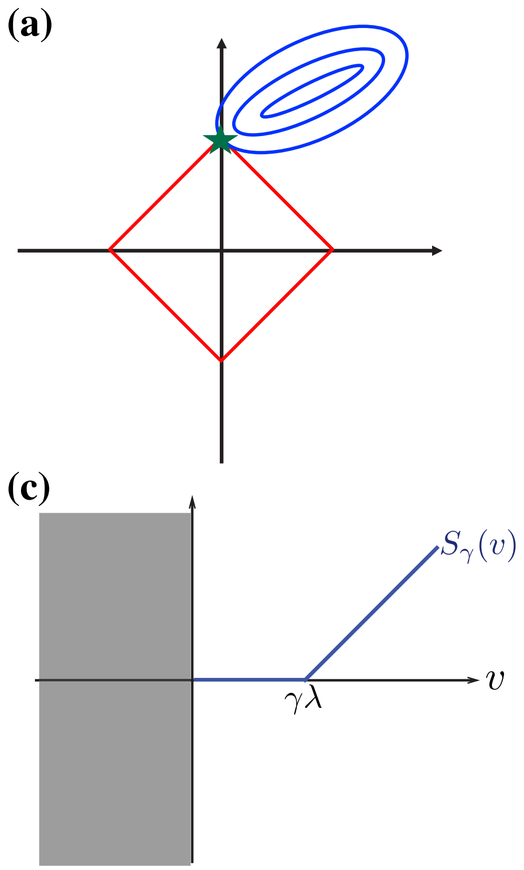

Figure 4. (a) Schematic picture for the sparse modeling. The red diamond represents ℓ^1 norm regularization term. The blue contours represent the sum of squared residuals. The green star is an optimal point. (b) That for the ℓ^2 norm regularization. The red circle represents ℓ^2 norm regularization term. (c) An example of the soft-thresholding function. $\gamma\lambda$ is a threshold value.

In the ℓ^1 norm regularization, the second term of the right-hand side of Equation 8 is nondifferentiable. Therefore, we use the Iterative Shrinkage Thresholding Algorithm (ISTA) (e.g., Beck & Teboulle, 2009) and minimize the objective function by repeating the following procedure:

$$\mathbf{m}_{k+1} = S_\gamma\left(\mathbf{m}_k - \gamma\mathbf{G}^\top(\mathbf{G}\mathbf{m}_k - \mathbf{d})\right). \tag{9}$$

$\mathbf{m}_k$ is the estimated model vector on the kth iteration. $\mathbf{G}^\top(\mathbf{G}\mathbf{m}_k - \mathbf{d})$ is a gradient of $\frac{1}{2}\|\mathbf{d} - \mathbf{G}\mathbf{m}\|_2^2$ (the first term of the right-hand side of Equation 8). γ (>0) is a step size. The ISTA will converge when the step size $\gamma \in (0, 1/L]$ is used (e.g., Parikh & Boyd, 2013), where L is a Lipschitz constant. We use the maximum eigenvalue of the matrix $\mathbf{G}^\top\mathbf{G}$ as the Lipschitz constant L (e.g., Beck & Teboulle, 2010). On the ISTA, most of the model parameters are kept to be 0 due to the following soft-thresholding function (Figure 4c):

$$S_\gamma(v) = \begin{cases} v - \gamma\lambda & , v \geq \gamma\lambda \\ 0 & , 0 < v < \gamma\lambda \end{cases}, \tag{10}$$

where v is a real number. The model parameters are kept to be positive during iterations. The soft-thresholding function returns 0 in the case of $0 < v < \gamma\lambda$. Accordingly, we can estimate a sparse model vector that minimizes Equation 8.

3. Results

3.1. Spatial Distribution of Seismic Scattering Properties

In this part, we use all 17 Hi-net stations to estimate the spatial distribution of seismic scattering property changes. As mentioned in section 2, we select decoherence values according to the SNRs of stacked CCFs in

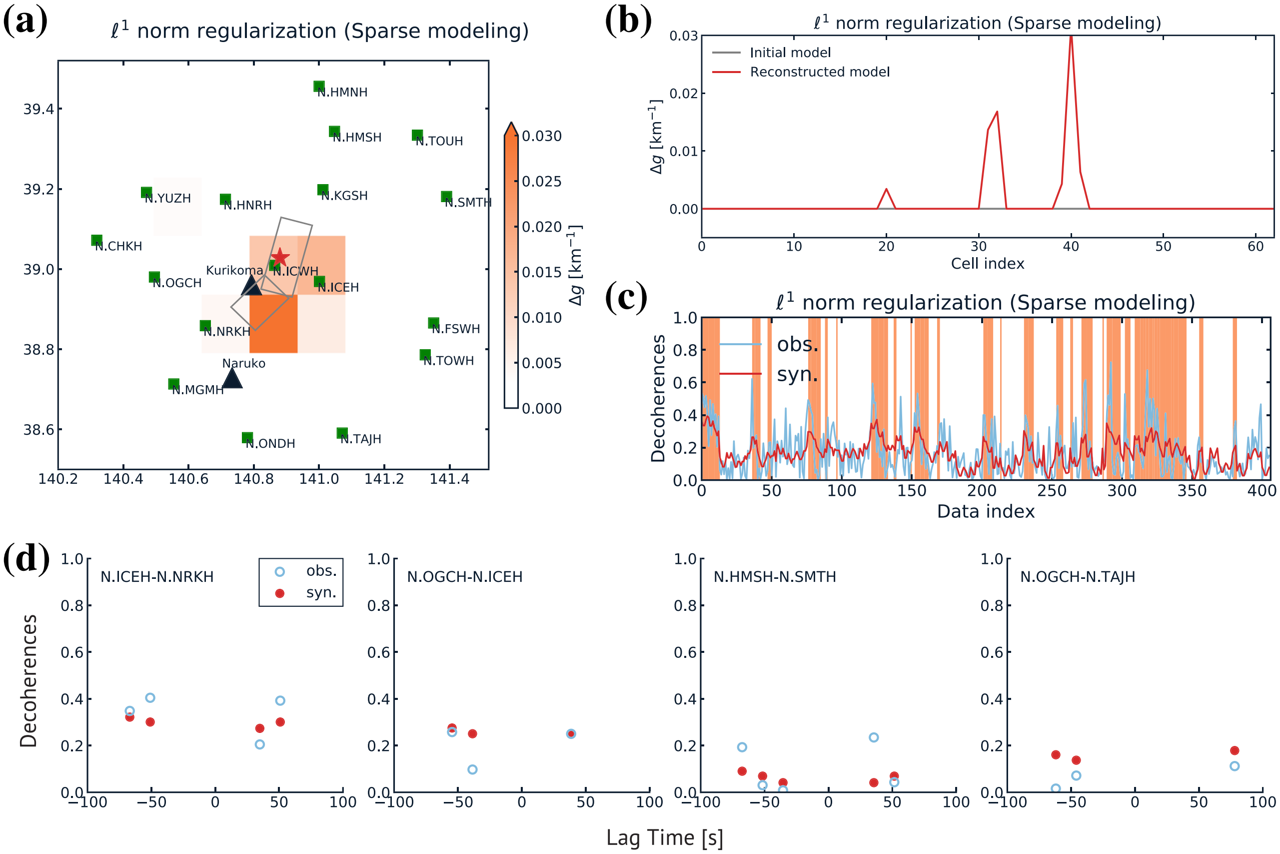

Figure 5. (a) Spatial distribution of Δg values inferred from the sparse modeling. The red star represents the epicenter. The two subfaults estimated by Ohta et al. (2008) are shown by gray rectangles. (b) The initial (gray) and estimated (red) models. All of the initial model parameters are fixed to zero (see Text S1 and Figure S1 in the supporting information). (c) Observed (blue) and synthesized (red) decoherence values for all station pairs and lag times (Table S1). Orange shaded parts correspond to decoherence values that satisfying the following: (1) The station pairs that cross the three small cells with large Δg values (31, 32, and 40 of cell index). (2) The station pairs include N.ICWH or N.ICEH that locate within the region of large change. (d) Observed (blue circle) and synthesized (red circle) decoherence values at each lag time for the four station pairs in Figure 2.

Period II (after the earthquake). In this study, we use 406 observed decoherence values. To use Equation 6, we need to divide the study area into small cells whose size is smaller than the scattering mean-free path in the study area to assume that scattering occurs only once within each cell. As shown in Appendix A, the scattering mean-free path of Rayleigh wave in our target region is estimated to be 30 km. Therefore, we divide the study area into 63 small cells with a size of 0.15° × 0.15° (approximately 13 km × 16 km) and estimate Δg values for all small cells. Figure 5a shows the estimated spatial distribution of Δg values using sparse modeling after 212 iterations. The region with the largest Δg value is estimated at the south of the epicenter of the main shock, and its value is 0.032 km^{-1} ($\Delta g/g_0 = 97\,\%$). Misfit (SSRs) between observed and synthesized decoherences is 5.38. The red line in Figure 5b represents the estimated Δg values for the small cells. The gray line represents the initial model. In this study, the initial model is a zero vector, as we do not possess any a priori information about the expected changes. In this case, only six small cells have nonzero values, and this indicates that the sparse modeling works well. Figure 5c shows observed and synthesized decoherence values for all station pairs and lag times. Orange shaded parts correspond to decoherence values satisfying the following conditions: (1) The station pairs that cross the three small cells with large Δg values (31, 32, and 40 of Cell index. See Text S1 and Figure S1 in the supporting information). (2) The station pairs include N.ICWH or N.ICEH that locate within the region of large change. Many spiky-shaped parts in observed decoherences correspond to the orange shaded parts. Synthesized decoherence values are calculated using the estimated model parameters and decorrelation kernels (Equation 5). Figure 5d shows observed and synthesized decoherence values at each lag time for the four station pairs in Figure 2. Synthesized decoherences almost explain observed ones. As shown in Equation 8, we consider both the SSRs (data fitness) and the ℓ^1 norm (sparsity of the model), and the ℓ^1 norm term is weighted by a hyperparameter λ. The value of λ affects estimated Δg values. The top panel of Figure 6 shows its effect: Pink lines on

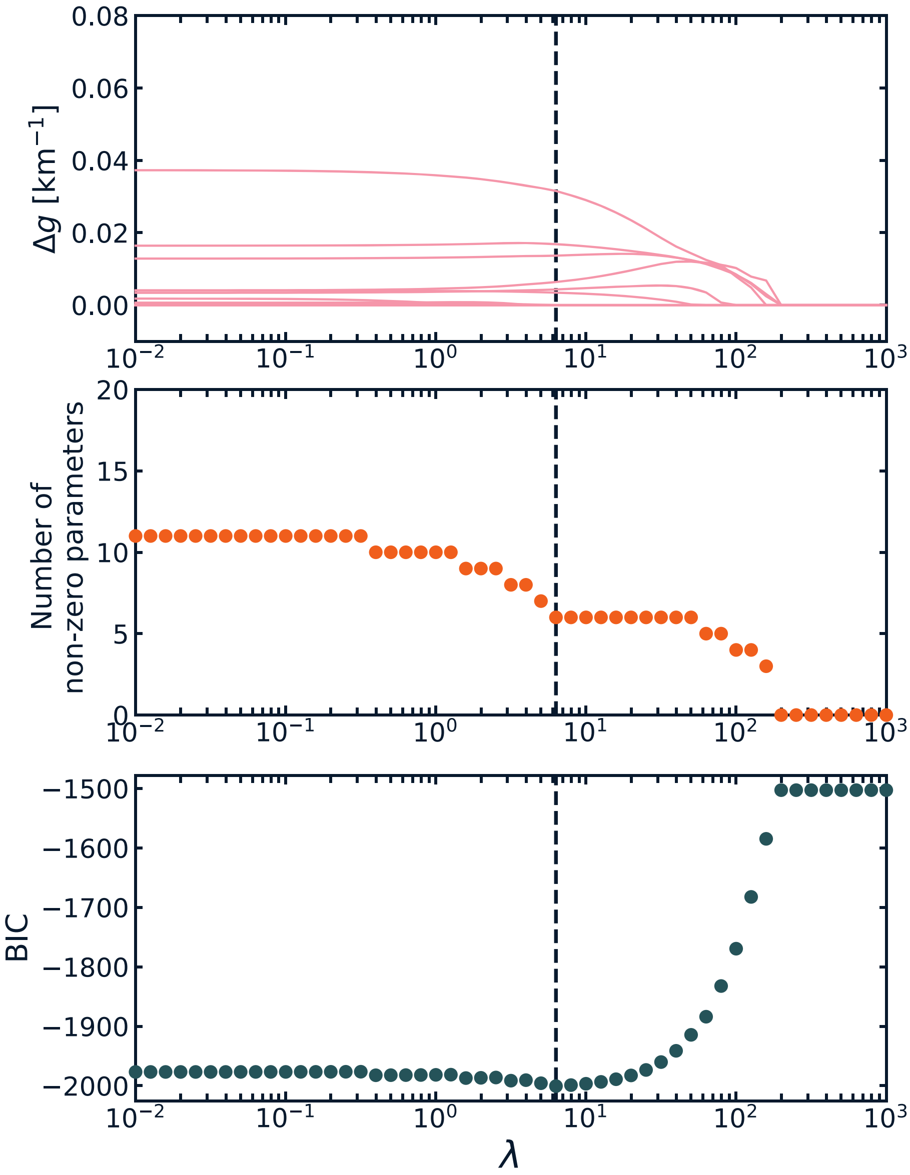

Figure 6. (top) Δg values for each small cell and λ. (middle) Numbers of nonzero estimated model parameters for each λ. (bottom) BIC values for each λ. The black dashed lines on each panel represent the value of λ with the minimum BIC.

the top panel represent estimated Δg values for the small cells. The middle panel shows the numbers of nonzero model parameters. For a large λ, only sparsity of a model is considered, and accordingly, all model parameters become 0 or nearly equal 0. We determine the optimal value of λ by calculating the Bayesian information criterion (BIC) (e.g., Kass & Raftery, 1995):

$$\mathrm{BIC} = n \ln\left(\frac{SSR}{n}\right) + \mathrm{d}f \times \ln(n). \tag{11}$$

Here, n is the number of the data and df is the degree of freedom of the estimated model. In the case of ℓ^1 norm regularization, df is the number of the nonzero model parameters (Wang et al., 2009). The bottom panel of Figure 6 shows BIC for all λ. In our case, the BIC becomes the minimum (BIC = -2000.8) for λ = 6.3, and we use this value.

3.2. Recovery Tests

To check the reliability of our inversion results, we conduct recovery tests. Figure 7a shows the input model of the recovery tests. We input Δg value of 0.033 km^{-1} ($\Delta g/g_0$ = 100%) to only one small cell at the south of

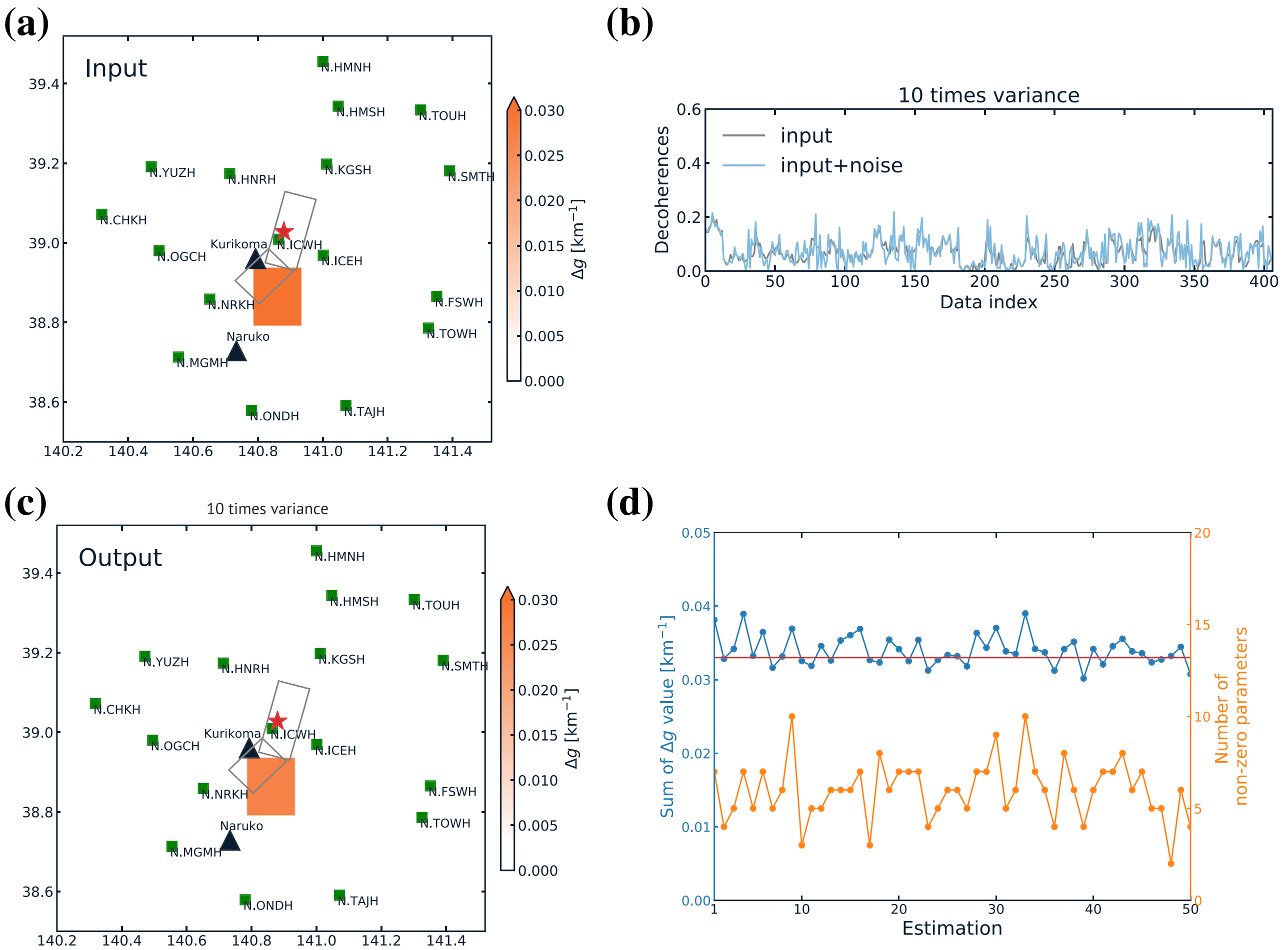

Figure 7. (a) Input model of the recovery tests. The Δg value of 0.033 km^{-1} ($\Delta g/g_0 = 100\%$) is assigned to only one small cell at the south of the epicenter. Input Δg values for the other small cells are zero. (b) An example of synthesized decoherences. The gray line represents the decoherences calculated from the input model shown in (a). The blue one represents decoherences after adding noises to the gray ones, and these decoherences are used in the recovery tests. (c) The result of the recovery tests using sparse modeling. The Δg values shown in the figure are averaged for 50 estimations. (d) Sum of estimated Δg values (blue) and the number of nonzero model parameters (orange) for all 50 estimations resulting from the sparse modeling. The red line represents input Δg value, 0.033 km^{-1}.

the epicenter. Input Δg values for the other small cells are set to be 0. First, we compute synthesized decoherence values for this input model using Equation 5. The gray line in Figure 7b represents the synthesize decoherence values from the input model. Second, we add noises to these synthesize decoherences:

$$dc_i^{syn\prime}(t) = dc_i^{syn}(t) + N(0, \sigma_i^2). \tag{12}$$

Here, $dc_i^{syn}(t)$ is the synthesized decoherence value from the input model for a station pair i. $N(0, \sigma_i^2)$ is the noise term and is sampled from a Gaussian distribution, which is defined for each station pair. We compute coherences between a RCCF and 10 days stacked CCFs from 1 January 2007 to 13 June 2008 and calculate a variance of these coherence values for each station pair. We use 10 times variance as σ_i^2 for each station pair. The blue line in Figure 7b is an example of the synthesized decoherences with noise. Third, we estimate a spatial distribution of Δg using $dc_i^{syn\prime}(t)$ as input data. We conduct the same estimation for 50 times with changing $dc_i^{syn\prime}(t)$.

Figure 7c shows the results of the inversion. The input change region at the south of the epicenter is well recovered. The Δg values shown in Figure 7c are averaged for 50 estimations. The Δg value at the south of the epicenter is estimated to be 0.027 ± 0.003 km^{-1}, and this value is 82% of the input value (0.033 km^{-1}). Figure 7d shows the sum of estimated Δg values for all small cells (blue line) and the number of nonzero model parameters (orange line) for all 50 estimations resulting from the sparse modeling. The sum of estimated Δg values for all small cells is 0.034 ± 0.002 km^{-1} on average, and this value is almost the same as the input value. Nonzero model parameters were estimated at six small cells on average; however, estimated Δg values at the other cells are significantly small compared to that at the south of the epicenter. We also conducted the same synthetic tests with higher noise levels. Although Δg values at the south of the epicenter are underestimated with increasing noise level, the result that maximum change is located at the south of the epicenter is retained even in the case of higher noise levels.

According to the results of the recovery tests, the region of large seismic scattering property changes that is estimated at the south of the epicenter should be reliable, though the value of Δg is slightly underestimated.

4. Discussion

4.1. Comparison With Other Observations

In some previous studies, seismic velocity decreases related to the 2008 Iwate-Miyagi Nairiku earthquake were detected by seismic interferometry analyses (e.g., Hobiger et al., 2012, 2014; Takagi et al., 2012; Wang et al., 2017). Takagi et al. (2012) reported that about 0.3–0.4% Rayleigh wave velocity decreases at the southern part of the fault zone in the 0.25–0.5 Hz band. Hobiger et al. (2012) detected about 0.6% velocity decrease at a maximum around the south of the epicenter in the 0.5–1 Hz band. These previous studies compared seismic velocity changes with PGA and concluded that such large seismic velocity decreases were caused by the strong ground motion. An accelerometer located near the epicenter (N.ICWH) recorded a strong motion of approximately 4,000 gal (Aoi et al., 2008). This region is located just at the north of the region of large seismic scattering property changes. Suzuki et al. (2010) estimated the fault slip distribution by waveform inversions using strong motion data recorded at K-NET and KiK-net stations. The maximum slip was estimated to be 6.2 m for the area approximately 10 km to the south of the hypocenter of the main shock, and this region is well consistent with the region of the largest Δg. The region of large slip extended from about 1 to 6 km at depth. Since the peak sensitivity of Rayleigh wave phase velocity to shear velocity is about one third of wavelength, Rayleigh waves at the 0.5–1 Hz band will be sensitive to structural changes to about 1–2 km. Therefore, a large structural change caused by the large fault slip might have introduced seismic scattering property changes. Such large coseismic slip at the south of the epicenter was also detected by GNSS observations (e.g., Iinuma et al., 2009; Ohta et al., 2008). The gray rectangles in Figures 5a and 9a represent the fault planes estimated by Ohta et al. (2008). The slip amounts were estimated to be 1.8 m for the northern fault plane and 3.5 m for the southern one, respectively. The region of the largest Δg includes the southern fault plane, supporting our interpretation that the large slip might have caused seismic scattering property changes.

Although the regions of large seismic scattering property changes are well consistent with the regions of large slip and strong ground motion, we cannot rule out other possible causes of large seismic scattering property changes. One candidate is fluid migration: Hillers et al. (2015) reported significant decoherences associated with geothermal reservoir stimulation in Switzerland. They interpreted that these decoherences were caused by fluid injection. In the case of the Iwate-Miyagi Nairiku earthquake, the Kurikoma volcano locates close to the epicenter of the main shock. The crustal fluid beneath the active volcano might have moved to a shallow region associated with the earthquake and that may have caused the seismic scattering property changes. The other possible cause of large seismic scattering property changes is landslides. They occurred especially around the Kurikoma volcano associated with the earthquake. The dimension of the largest one was estimated to be 900 × 1,300 m (Yonezawa et al., 2010). Such an area of material weakness may be related to the seismic scattering property changes.

These candidates must be related to the large seismic scattering property change that occurred at the south of the main shock; however, we are not able to evaluate the contributions each cause to the seismic scattering property changes. Depth sensitivity of Rayleigh waves may be used to clarify the main cause of seismic scattering property changes: Strong ground motion and landslide generally cause structural changes especially in a shallow region. If a seismic scattering property change is caused by these factors, estimated Δg values are expected to be larger at higher-frequency bands than those at lower-frequency bands when considering surface waves only. At frequency bands below 0.5 Hz, we could not constrain the values of scattering mean-free path by using envelopes of CCFs. However, it is interesting to see if we can find frequency dependence in the spatial distribution of Δg values. Therefore, we conducted the same inversions at 0.25–0.5 and 0.125–0.25 Hz bands by assuming the scattering mean-free path to be 100 km. The results are shown in the supporting information as a guidance (Text S2 and Figure S2). Both at two frequency bands, regions of large

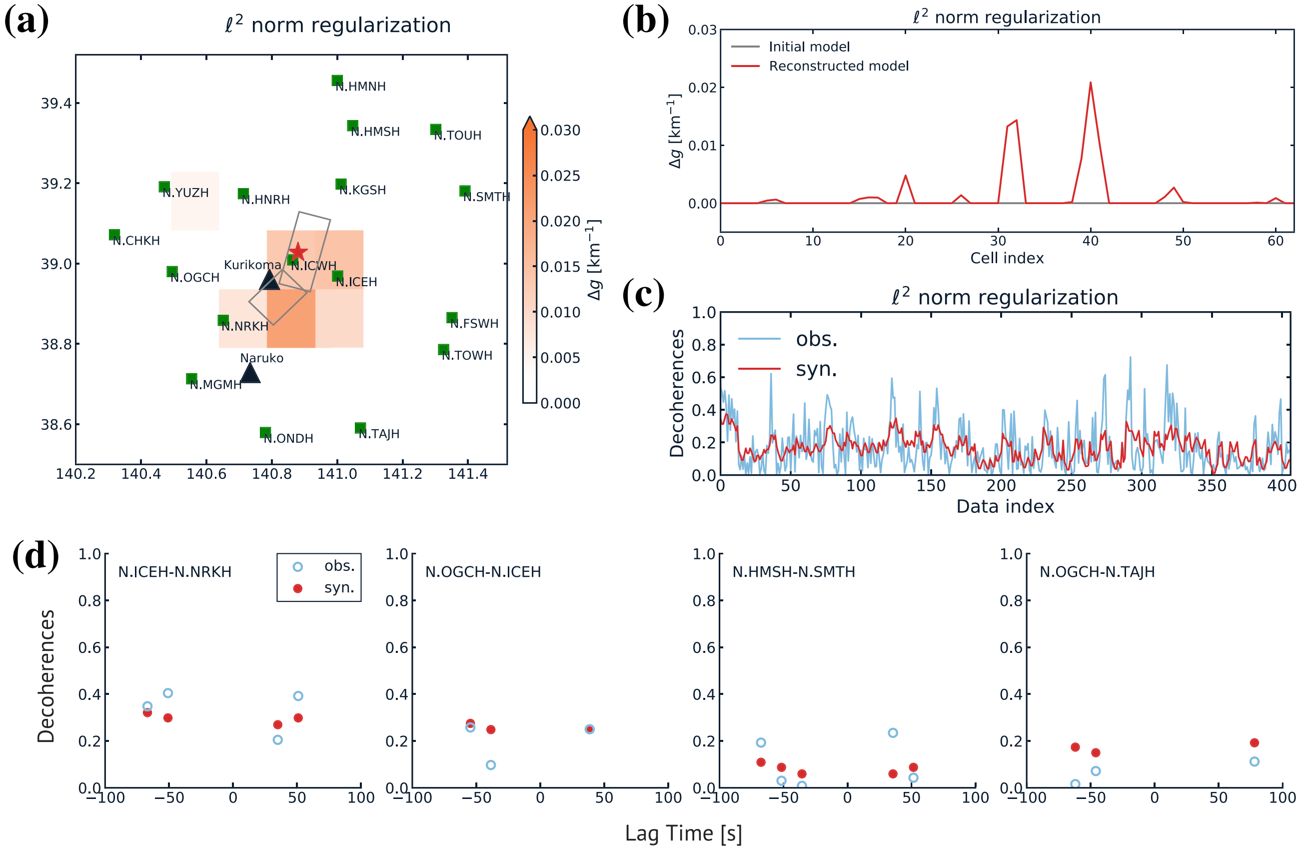

Figure 8. Similar to Figure 5 but obtained from the ℓ^2 norm regularization.

changes were located around the epicenter. The estimated maximum $\Delta g/g_0$ value at the 0.125–0.25 Hz band is a little bit smaller than those at the 0.25–0.5 and 0.5–1 Hz bands. This might reflect that seismic scattering property changes at higher-frequency bands are caused by some factors that can occur at shallow depths: large fault slip, strong ground motion, and landslide.

4.2. Comparison With ℓ^2 Norm Regularization

4.2.1. Result From the Linear Least Squares Inversion With ℓ^2 Norm Regularization

We estimate the spatial distribution of seismic scattering property changes from an ordinary linear least squares inversion with the ℓ^2 norm regularization which has been widely used to solve linearized inverse problems. In the case of the Iwate-Miyagi Nairiku earthquake, we have 17 seismic stations that are enough to conduct the ordinary linear least squares inversion. Therefore, we can compare the results from the sparse modeling with those from the ℓ^2 norm regularization (Figure 8).

Figure 8a shows the spatial distribution of Δg values. Here, we use the proximal gradient method (e.g., Beck & Teboulle, 2009; Nesterov, 2007) to estimate values of Δg. We selected the optimal value of λ which minimizes BIC. For the ℓ^2 norm regularization, df in Equation 11 is defined as the trace of the data resolution matrix, $\mathrm{d}f = \mathrm{tr}[\mathbf{G}(\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{G}^T]$ (Hastie et al., 2001). The BIC value of the estimated models from the ℓ^2 norm regularization is -1872.5. The regions of large Δg values are located around the epicenter of the main shock, and this result is well consistent with the result from the sparse modeling. As shown in Figure 8d, the synthesized decoherence values are relatively similar to observed ones as well as those from the sparse modeling. The maximum Δg value is about 0.021 km^{-1} ($\Delta g/g_0 = 64\%$) at the small cell which locates to the south of the epicenter. Misfit (SSRs) between observed and synthesized decoherences is 5.52. This is a little bit larger than that of the sparse modeling. Estimated Δg values from the ℓ^2 norm regularization are smaller than those from the sparse modeling. This is because the number of small cells with nonzero values is 18 in the case of the ℓ^2 norm regularization (Figure 8b). As shown in Figure 8c, the synthesized decoherence values roughly explain observed ones. However, the synthesized decoherences from the sparse modeling

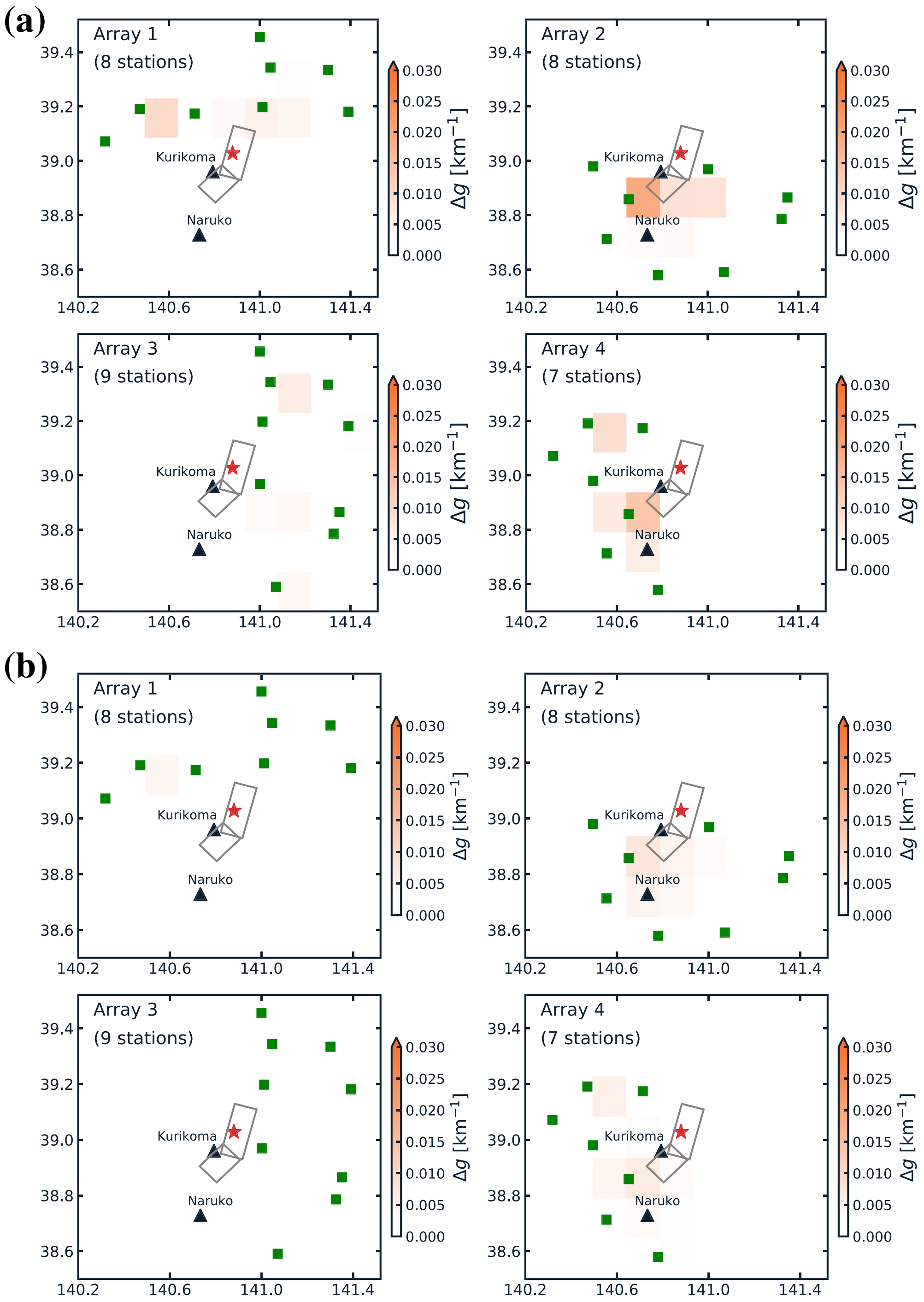

Figure 9. (a) Spatial distributions of Δg values obtained from the sparse modeling for four different distributions of seismic stations. The study area is divided into four areas as indicated by vertical and horizontal dashed lines in each panel. Those four seismic arrays are composed 3 (Array 1), southern part (Array 2), eastern part (Array 3), and western part (Array 4), respectively. (b) Similar to (a) but obtained from the ℓ^2 norm regularization.

can better explain the observed ones, especially spiky-shaped parts. This can be because the estimated Δg value from the sparse modeling is more sharply large at the south of the epicenter than that from the ℓ^2 norm regularization. The BIC value from the ℓ^1 norm regularization is 128.3 less than that from the ℓ^2 norm regularization. When the difference between BIC values from different two models is larger than 6, these two models are considered to be strongly different (Kass & Raftery, 1995). Based on the above discussions, we conclude that the estimated model from the ℓ^1 norm regularization is better than that from the ℓ^2 norm regularization.

4.2.2. Comparison Between Sparse Modeling and ℓ^2 Norm Regularization With Smaller Data Sets

To validate the applicability of sparse modeling with a smaller data set, we estimate spatial distributions of seismic scattering property changes by changing combinations of used stations.

First, we divide the study area into four regions around the epicenter and conduct estimations. Figure 9a shows the results for four different seismic arrays. These seismic arrays are composed of seismic stations in the northern part (Array 1), southern part (Array 2), eastern part (Array 3), and western part (Array 4), respectively. Although large seismic scattering property changes are estimated at the south of the epicenter for Arrays 2 and 4, such change regions are not estimated for Arrays 1 and 3. This is probably because Arrays 2 and 4 include station pairs which cross around the south of the epicenter while Arrays 1 and 3 do not. As shown in Figure 2, the station pairs of N.ICEH-N.NRKH (first panel) and N.OGCH-N.ICEH (second panel) which cross the south of the epicenter show significant decoherences related to the Iwate-Miyagi Nairiku earthquake. On the other hand, Arrays 1 and 3 are mainly composed of the station pairs which do not cross around the south of the epicenter. Most of these station pairs do not show significant decoherences related to the earthquake: For example, the station pairs of N.HMSH-N.SMTH (third panel in Figure 2) and N.OGCH-N.TAJH (bottom panel in Figure 2) do not show clear coherence drops. Those results indicate that we need station pairs which cross the south of the epicenter to detect the significant seismic scattering property change related to the Iwate-Miyagi Nairiku earthquake. We also conduct estimations using the same four seismic arrays and the ℓ^2 norm regularization (Figure 9b). In the case of the ℓ^2 norm regularization, the change regions at the south of the epicenter are not estimated significantly for all the four arrays. Although both of the sparse modeling and ℓ^2 norm regularization by using those four seismic arrays do not reproduce the results from all 17 stations, the results from the sparse modeling are better than those from the ℓ^2 norm regularization.

Second, we use other arrays and check how many seismic stations are needed to reproduce the result from all 17 stations: The top left panel of Figure 10 shows the result from three seismic stations closest to the epicenter (Array 5). The result from this seismic array still cannot retrieve the result from all 17 stations. The top right panel of Figure 10 shows the result from five seismic stations that are deployed to surround the epicenter at close range (Array 6). The large seismic scattering property change at the south of the epicenter is sharply retrieved for Array 6, and the maximum value of Δg is the same as that from all 17 stations (0.032 km^{-1}). This suggests that if seismic stations are deployed to surround the change region and near the epicenter, we can retrieve the change region from only five stations in the case of the Iwate-Miyagi Nairiku earthquake. For comparison, we conduct the same estimations using the ℓ^2 norm regularization. Figure 11 shows the estimation results. In the result from Array 6, the maximum value of Δg is estimated to be 0.0083 km^{-1}. The result from all 17 stations is 0.021 km^{-1}, and hence, five seismic stations are not enough to retrieve the result from all 17 stations in the case of the ℓ^2 norm regularization. When we use the ℓ^2 norm regularization, we need 15 stations to retrieve the result from all 17 stations (Array 9). Arrays 7 and 8 are composed of 5 and 12 seismic stations that are located far from the epicenter, respectively. Although there are some station pairs which cross around the change region at the south of the epicenter, no significant change region of seismic scattering property is estimated both in the cases of the sparse modeling and ℓ^2 norm regularization. Those results indicate that seismic stations near the change region are important to retrieve it, and seismic stations far from the change region cannot retrieve it well even if they cross around this region.

To confirm that interpretation, we compare sensitivity kernels for different interstation distances. Figure 3b shows cross sections of sensitivity kernels for interstation distances of 30 km (mean-free path) and 60 km (twice the mean-free path). In this study, we compute decoherence values by sliding a 16 s long time window from 16 to 64 s after the direct Rayleigh wave arrival time without overlapping. Decoherence values are calculated at later lapse times for station pairs with long interstation distances. The two sensitivity kernels that are shown in Figure 3b are computed at a lapse time of 40 s after direct Rayleigh wave arrivals. As mentioned in section 2, there are two peaks whose widths are proportional to the mean-free path around

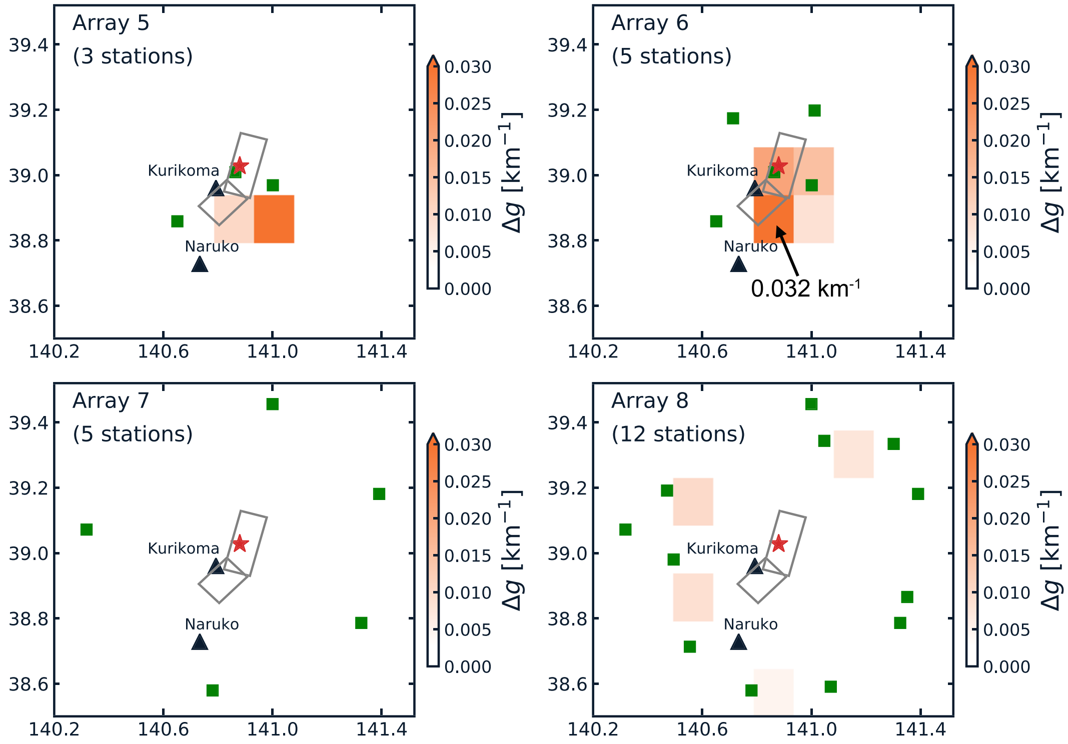

Figure 10. Spatial distributions of Δg values obtained from the sparse modeling for four different distributions of seismic stations (Arrays 5–8). Arrays 5 and 6 are composed of the three and five seismic stations closest to the epicenter, respectively. Arrays 7 and 8 are composed of 5 and 12 seismic stations that are located far from the epicenter, respectively.

the two stations, and this is the reason why seismic stations near the change region are important to retrieve the change region. As shown in Figure 3b, the peak values of the two sensitivity kernels are not significantly different from each other. On the other hand, the values of the sensitivity kernels in the region between two stations are found to be significantly different: That for the station pair with an interstation distance of 30 km is about 1.7 times larger than that for the station pair with an interstation distance of 60 km. This means that the station pair with a long interstation distance is less sensitive to a seismic scattering property change in a region between two stations. Accordingly, for the estimation of the spatial distribution of Δg values using sensitivity kernels, it should be better to use seismic stations near the change region and station pairs with short interstation distances. In other words, important data on the estimation are limited and the other data are redundant.

5. Conclusions

We applied seismic interferometry to the seismic ambient noise data before and after the 2008 Iwate-Miyagi Nairiku, Japan, earthquake and successfully detected significant decoherences associated with the earthquake. We estimated the spatial distribution of seismic scattering property changes using sparse modeling which enables us to locate change regions from a small data set. By using the sparse modeling, the region of the largest seismic scattering property changes was located at the south of the epicenter. The maximum value of Δg, variation of scattering coefficient, at this region was estimated to be 0.032 km^{-1} ($\Delta g/g_0 = 97\,\%$). We also conducted the ordinary linear least squares inversion with the ℓ^2 norm regularization. The results from the sparse modeling and that from the ℓ^2 norm regularization were consistent. The regions of large seismic scattering property changes are well consistent with the regions of large slip and strong ground motion, though fluid migration and landslides are not ruled out. We could not identify a cause of the seismic scattering property changes, results of previous studies and our study indicate that large structural changes occurred associated with the Iwate-Miyagi Nairiku earthquake. Finally, we explored the applicability of sparse modeling by reducing the number of seismic stations. We confirmed that we can retrieve the change regions from only five stations that are deployed to surround the change regions for this case. On the other

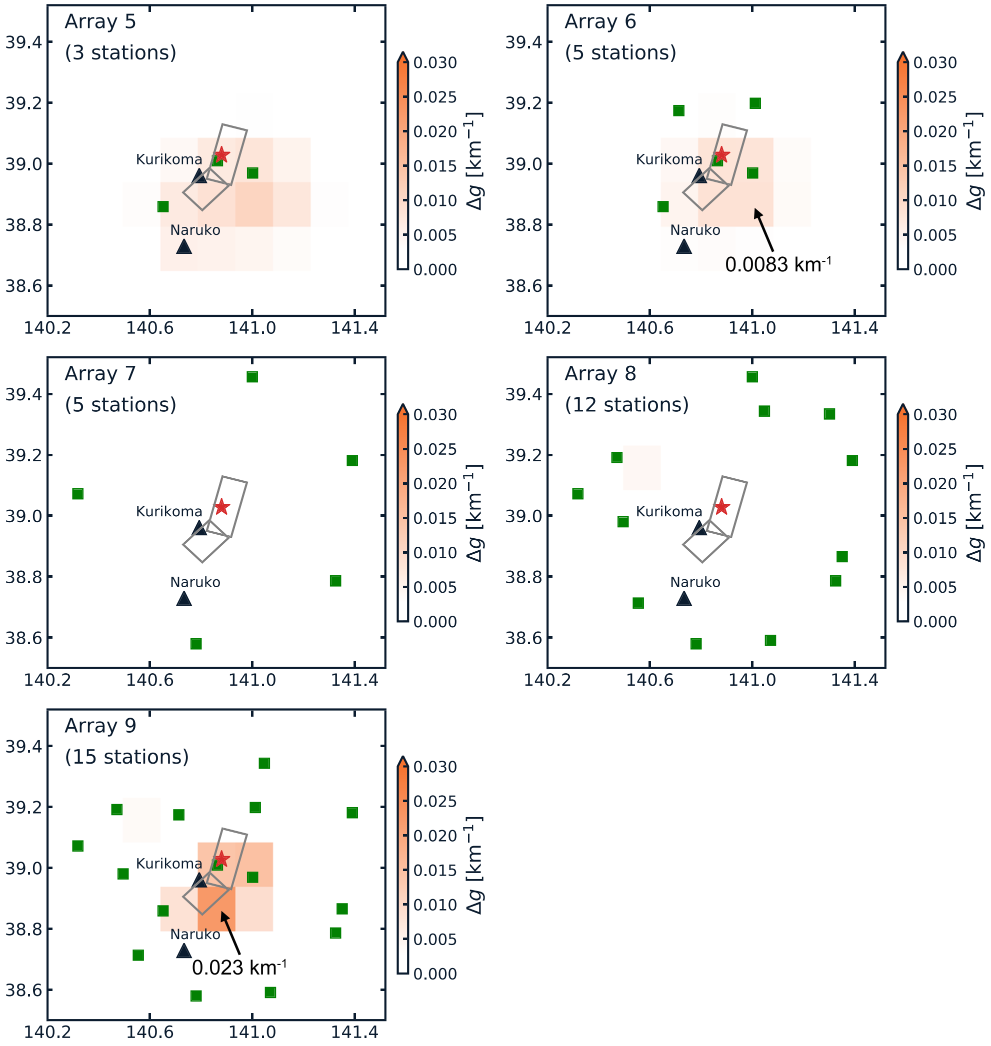

Figure 11. Similar to Figure 10 but obtained from the ℓ^2 norm regularization. The Array 9 is composed of 15 seismic stations closest to the epicenter.

hand, in the case of the ℓ^2 norm regularization, 15 seismic stations are necessary to reproduce the result from all 17 stations. The sparse modeling needs less data than the ℓ^2 norm regularization. We found that the station pairs with long interstation distances do not strongly affect the estimation result even if these pairs cross around the change regions. This is because, for a station pair with a long interstation distance, the value of its sensitivity kernel between two stations is small. This means that the important station pairs for the estimation are limited and the other pairs are redundant. The sparse modeling will be useful to estimate the spatial distribution of seismic scattering property changes from the small data set.

Appendix A: Scattering Mean-Free Path and Intrinsic Absorption Parameter in the Study Area

We estimated scattering mean-free path and intrinsic absorption parameter of Rayleigh wave at the 0.5–1 Hz band in the study area based on a passive method developed by Hirose et al. (2019). We computed envelopes of CCFs of seismic ambient noise on the vertical-vertical (ZZ), vertical-radial (ZR), vertical-transverse (ZT), radial-vertical (RZ), and transverse-vertical (TZ) components by stacking DCCFs over 1 year in 2007. We computed the mean squared (MS) envelopes by smoothing their squared amplitudes in 4 s. They were summed with respect to the three components and normalized by the envelopes at 55 s in lapse time for coda normalization. The best fit parameters were estimated by modeling the space-time distributions of energy densities using 2-D radiative transfer theory (e.g., Sato, 1993; Shang & Gao, 1988) (see also Equation 4). We

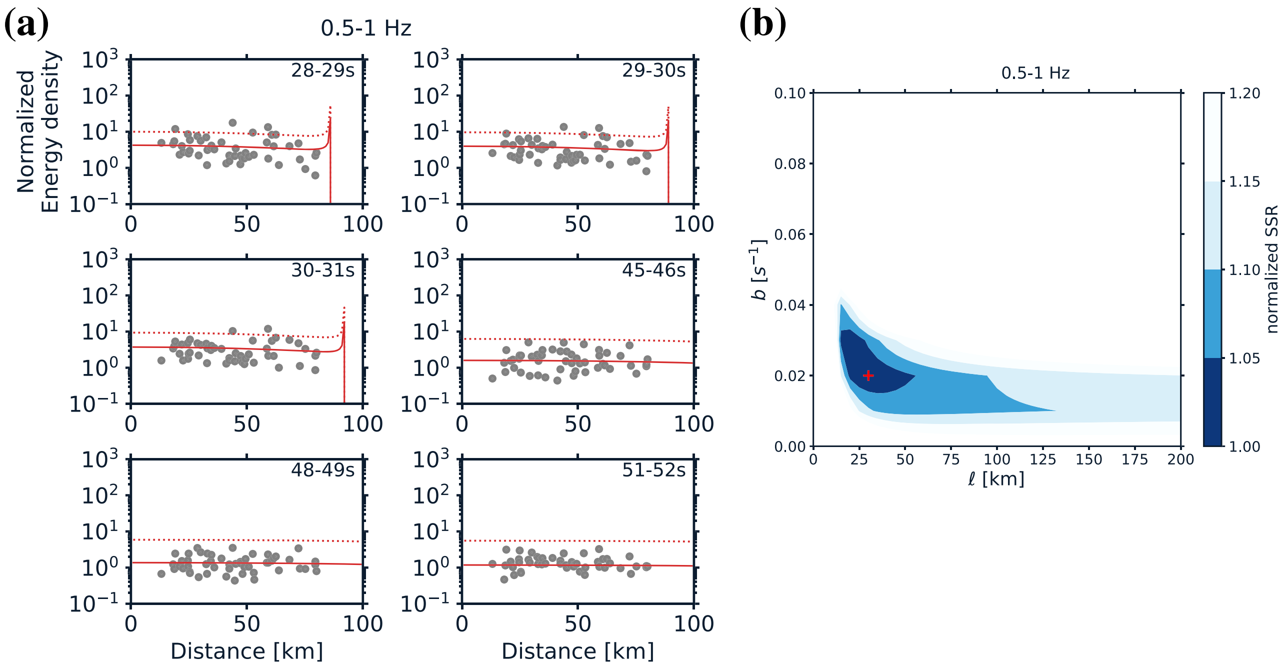

Figure A1. (a) Spatial distributions of energy densities for six time windows. The gray circles are observations that are calculated using the envelopes of the seismic ambient noise CCFs. The red solid curves represent synthesized values that are calculated by using the 2-D radiative transfer equation (Equation 4 multiplied by e^{-bt}) and the best fit values of the scattering mean-free path and intrinsic absorption parameter. The red dashed curves represent synthesized ones without intrinsic absorption. (b) Distribution of SSR between the observed and synthesized energy densities at the 0.5–1 Hz band. The horizontal and vertical axes represent the scattering mean-free path and the intrinsic absorption parameter, respectively. The red cross indicates the minimum SSR (best fit), by which SSRs on the grid points are normalized.

calculated the sum of the squared residuals (SSR) between the observed and synthesized energy densities and determined the best fit values. Figure A1a shows space-time distributions of observed (gray circles) and synthesized (red solid lines) energy densities. Figure A1b shows the result of grid search. The best fit scattering mean-free path ℓ is estimated to be 30 km (total scattering coefficient $g_0 = 0.033$ km^{-1}), and best fit intrinsic absorption parameter b ($b = Q_i^{-1}\omega$) is 0.02 s^{-1}.

We also estimate scattering mean-free path and intrinsic absorption parameter at the 0.125–0.25 and 0.25–0.5 Hz bands. The intrinsic absorption parameters are estimated to be 0.03–0.04 s^{-1} at the 0.125–0.25 Hz band and 0.02–0.03 s^{-1} at the 0.25–0.5 Hz band, respectively. In contrast, the scattering mean-free paths are not constrained in these frequency bands. The scattering mean-free path is mainly determined by spatial gradients of the energy densities in early lapse times which include direct and early coda parts of CCFs. At those frequency bands, the spatial gradients of the energy densities are almost flat within the seismic network, and this may lead to the weak constraint of the scattering mean-free paths. Accordingly, it is necessary to use a larger seismic network to estimate scattering mean-free paths at the 0.125–0.25 and 0.25–0.5 Hz bands in the case of the focal area of the Iwate-Miyagi Nairiku earthquake.

Acknowledgments
We thank the Editor, Yehuda Ben-Zion, and two anonymous reviewers for constructive comments. T. H. was partially supported by the International Joint Graduate Program in Earth and Environmental Science (GP-EES) of Tohoku University and the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant 18J10232. We also acknowledge the support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement 742335, F-IMAGE). Further support was received from JSPS KAKENHI Grant JP6K05528 and also from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan under its Earthquake and Volcano Hazards Observation and Research Program.

Data Availability Statement

We are grateful to the National Research Institute for Earth Science and Disaster Resilience for providing us with continuous seismic data. We downloaded the data from the Data Management Center of the National Research Institute for Earth Science and Disaster Resilience (http://www.hinet.bosai.go.jp)

References

Aoi, S., Kunugi, T., & Fujiwara, H. (2008). Trampoline effect in extreme ground motions. *Science*, *322*, 727–730. https://doi.org/10.1126/science.1163113

Beck, A., & Teboulle, M. (2009). A fast iterative Shrinkage-Thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, *2*(1), 183–202.

Beck, A., & Teboulle, M. (2010). *Gradient-based algorithms with applications to signal recovery problems, convex optimization in signal processing and communications*. Cambridge, UK: Cambridge Univ. Press.

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., et al. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, *169*, 1239–1260.

Brenguier, F., Campillo, M., Takeda, T., Aoki, Y., Shapiro, N. M., Briand, X., et al. (2014). Mapping pressurized volcanic fluids from induced crustal seismic velocity drops. *Science*, *345*, 80–82. https://doi.org/10.1126/science.1254073

Brenguier, F., Shapiro, N. M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O., & Nercessian, A.(2008). Towards forecasting volcanic eruptions using seismic noise. *Nature Geoscience*, *1*, 126–130.

Claerbout, J. F., & Muir, F. (1973). Robust modeling with erratic data. *Geophysics*, *38*(5), 826–844. https://doi.org/10.1190/1.1440378

Curtis, A., Gerstoft, P., Sato, H., Snieder, R., & Wapenaar, K. (2006). Seismic interferometry-turning noise into signal. *The Leading Edge*, *25*, 1082–1092. https://doi.org/10.1016/j.jvolgeores.2006.04.003

Evans, E. L., & Meade, B. J. (2012). Geodetic imaging of coseismic slip and postseismic afterslip: Sparsity promoting methods applied to the great Tohoku earthquake. *Geophysical Research Letters*, *39*, L11314. https://doi.org/10.1029/2012GL051990

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning*. New York: Springer New York Inc. https://doi.org/10.1007/978-0-387-84858-7

Hillers, G., Husen, S., Obermann, A., Planès, T., Larose, E., & Campillo, M. (2015). Noise-based monitoring and imaging of aseismic transient deformation induced by the 2006 Basel reservoir stimulation. *Geophysics*, *80*(4), KS51–KS68.

Hirose, T., Nakahara, H., & Nishimura, T. (2019). A passive estimation method of scattering and intrinsic absorption parameters from envelopes of seismic ambient noise cross-correlation functions. *Geophysical Research Letters*, *46*, 3634–3642. https://doi.org/10.1029/2018GL080553

Hobiger, M., Wegler, U., Shiomi, K., & Nakahara, H. (2012). Coseismic and postseismic elastic wave velocity variations caused by the 2008 Iwate-Miyagi Nairiku earthquake, Japan. *Journal of Geophysical Research*, *117*, B09313. https://doi.org/10.1029/2012JB009402

Hobiger, M., Wegler, U., Shiomi, K., & Nakahara, H. (2014). Single-station cross-correlation analysis of ambient seismic noise: Application to stations in the surroundings of the 2008 Iwate-Miyagi Nairiku earthquake. *Geophysical Journal International*, *198*, 90–109. https://doi.org/10.1093/gji/ggu115

Iinuma, T., Ohzono, M., Ohta, Y., Miura, S., Kasahara, M., Takahashi, H., et al. (2009). Aseismic slow slip on an inland active fault triggered by a nearby shallow event, the 2008 Iwate-Miyagi Nairiku earthquake (Mw6.8). *Geophysical Research Letters*, *36*, L20308. https://doi.org/10.1029/2009GL040063

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Larose, E., Obermann, A., Digulescu, A., Planès, T., Chaix, J. F., Mazerolle, F., & Moreau, G. (2015). Locating and characterizing a crack in concrete with diffuse ultrasound: A four-point bending test. *Journal of Acoustical Society of America*, *138*(1), 232–241.

Machacca-Puma, R., Lesage, P., Larose, E., Lacroix, P., & Anccasi-Figueroa, R. M. (2019). Detection of pre-eruptive seismic velocity variations at an andesitic volcano using ambient noise correlation on 3-component stations: Ubinas volcano, Peru, 2014. *Journal of Volcanology and Geothermal Research*, *381*, 83–100. https://doi.org/10.1016/j.jvolgeores.2019.05.014

Margerin, L., Planès, T., Mayor, J., & Calvet, M. (2016). Sensitivity kernels for coda-wave interferometry and scattering tomography: Theory and numerical evaluation in two-dimensional anisotropically scattering media. *Geophysical Journal International*, *204*, 650–666. https://doi.org/10.1093/gji/ggv470

Nakata, R., Kuwatani, T., Okada, M., & Hori, T. (2016). Geodetic inversion for spatial distribution of slip under smoothness, discontinuity, and sparsity constraints. *Earth, Planets and Space*, *68*, 20. https://doi.org/10.1186/s40623-016-0396-2

National Research Institute for Earth Science and Disaster Resilience (2019). NIED Hi-net. National Research Institute for Earth Science and Disaster Resilience, https://doi.org/10.17598/nied.0003

Nesterov, Y. (2007). Gradient methods for minimizing composite objective function, No 2007076 CORE Discussion Papers. Université catholique de Louvain, Center for Operations Research and Econometrics (CORE, https://EconPapers.repec.org/RePEc:cor:louvco:2007076

Nimiya, H., Ikeda, T., & Tsuji, T. (2017). Spatial and temporal seismic velocity changes on Kyushu Island during the 2016 Kumamoto earthquake. *Science Advances*, *3*(11), e1700813. https://doi.org/10.1126/sciadv.1700813

Nishida, K., Kawakatsu, H., & Obara, K. (2008). Three-dimensional crustal S wave velocity structure in Japan using microseismic data recorded by H-net tiltmeters. *Journal of Geophysical Research*, *113*, B10302. https://doi.org/10.1029/2007JB005395

Obermann, A., Froment, B., Campillo, M., Larose, E., Planès, T., Valette, B., et al. (2014). Seismic noise correlations to image structural and mechanical changes associated with the Mw7.9 2008 Wenchuan earthquake. *Journal of Geophysical Research: Solid Earth*, *119*, 3155–3168. https://doi.org/10.1002/2013JB010932

Obermann, A., Planès, T., Hadziioannou, C., & Campillo, M. (2016). Lapse-time-dependent coda-wave depth sensitivity to local velocity perturbations in 3-D heterogeneous elastic media. *Geophysical Journal International*, *207*(1), 59–66.

Obermann, A., Planès, T., Larose, E., & Campillo, M. (2013). Imaging preeruptive and coeruptive structural and mechanical changes of a volcano with ambient seismic noise. *Journal of Geophysical Research: Solid Earth*, *118*, 6285–6294. https://doi.org/10.1002/2013JB010399

Ohta, Y., Ohzono, M., Miura, S., Iinuma, T., Tachibana, K., Takatsuka, K., et al. (2008). Coseismic fault model of the 2008 Iwate-Miyagi Nairiku earthquake deduced by a dense GPS network. *Earth, Planets and Space*, *60*, 1197–1201.

Parikh, N., & Boyd, S. P. (2013). Proximal algorithms. *Foundations and Trends in Optimization*, *1*(3), 123–231.

Santosa, F., & Symes, W. W. (1986). Linear inversion of band-limited reflection seismograms. *SIAM Journal of Scientific and Statistical Computing*, *7*(4), 1307–1330. https://doi.org/10.1137/0907087

Sato, H. (1993). Energy transportation in one- and two-dimensional scattering media: Analytic solutions of the multiple isotropic scattering model. *Geophysical Journal International*, *112*, 141–146. https://doi.org/10.1111/j.1365-246X.1993.tb01443.x

Sato, H., Fehler, M. C., & Maeda, T. (2012). *Seismic wave propagation and scattering in the heterogeneous Earth* (Vol. 496). Berlin: Springer.

Shang, T., & Gao, L. (1988). Transportation theory of multiple scattering and its application to seismic coda waves of impulsive source. *Scientia Sinica (series B China)*, *31*, 1503–1514.

Shapiro, N. M. (2004). Campillo Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letters*, *31*, L07614. https://doi.org/10.1029/2004GL019491

Shapiro, N. M., Ritzwoller, M. H., & Bensen, G. D. (2006). Source location of the 26 sec microseism from cross-correlations of ambient seismic noise. *Geophysical Research Letters*, *33*, L18310. https://doi.org/10.1029/2006GL027010

Suzuki, W., Shin, A., & Sekiguchi, H. (2010). Rupture process of the 2008 Iwate-Miyagi Nairiku, Japan, earthquake derived from near-source strong-motion records. *Bulletin of the Seismological Society of America*, *100*(1), 256–266. https://doi.org/10.1785/0120090043

Takagi, R., Okada, T., Nakahara, H., Umino, N., & Hasegawa, A. (2012). Coseismic velocity change in and around the focal region of the 2008 Iwate-Miyagi Nairiku earthquake. *Journal of Geophysical Research*, *117*, B06315. https://doi.org/10.1029/2012JB009252

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B (Methodological)*, *58*(1), 267–288. http://www.jstor.org/stable/2346178

Wang, Q.-Y., Brenguier, F., Campillo, M., Lecointre, A., Takeda, T., & Aoki, Y. (2017). Seasonal crustal seismic velocity changes throughout Japan. *Journal of Geophysical Research: Solid Earth*, *122*, 7987–8002. https://doi.org/10.1002/2017JB014307

Wang, Q.-Y., Campillo, M., Brenguier, F., Lecointre, A., Takeda, T., & Hashima, A. (2019). Evidence of changes of seismic properties in the entire crust beneath Japan after the Mw 9.0, 2011 Tohoku-oki earthquake. *Journal of Geophysical Research: Solid Earth*, *124*, 8924–8941. https://doi.org/10.1029/2019JB017803

Wang, H., Li, B., & Leng, C. (2009). Shrinkage tuning parameter selection with a diverging number of parameters. *Journal of Royal Statistical Society Series B*, *71*(3), 671–683.

Weaver, R. L. (2011). On the amplitudes of correlations and the inference of attenuations, specific intensities and site factors from ambient noise. *Comptes Rendus Geoscience*, *343*(8–9), 615–622.

Yao, H., Gerstoft, P., Shearer, P. M., & Mecklenbräuker, C. (2011). Compressive sensing of the Tohoku-Oki Mw 9.0 earthquake: Frequency-dependent rupture modes. *Geophysical Research Letters*, *38*, L20310. https://doi.org/10.1029/2011GL049223

Yonezawa, C., Saito, G., Hori, M., & Watanabe, M. (2010). Continuous monitoring of landslides area caused by Iwate-Miyagi Nairiku earthquake in 2008 using ALOS images, International Archives of the Photogrammetry. *Remote Sensing and Spatial Information Science*, *XXXVIII*(8), 44–47.