**Advanced surface-wave analysis for 3D ocean bottom cable data to detect localized
heterogeneity in shallow geological formation of a CO₂ storage site**

Tatsunori Ikeda¹ and Takeshi Tsuji^{1*}

1. International Institute for Carbon-Neutral Energy Research (WPI-I2CNER), Kyushu University,
Motooka, Nishi-ku, Fukuoka 819-0395, Japan.

*Corresponding author. E-mail: tsuji@i2cner.kyushu-u.ac.jp

International Institute for Carbon-Neutral Energy Research (WPI-I2CNER), Kyushu University,
Motooka, Nishi-ku, Fukuoka 819-0395, Japan.

TEL: +81-92-802-6875

FAX: +81-92-802-6875

15

Highlights

16 1. We propose a surface-wave analysis using 3D data to estimate subsurface heterogeneity

17 2. Pseudo-3D S-wave velocity model is obtained by multimode inversion of Scholte-wave

18 3. Horizontal component data analysis can contribute to robust surface-wave analysis

19 4. The lithology boundary across the survey area is identified from our proposed analyses

20 5. The heterogeneity identified from our proposed approach is vital in safe CO₂ storage

21

22

23

24

25 **Keywords:** surface-wave analysis; CO₂ storage; Scholte waves; 3D Vs structure; 3D attenuation
26 structure; geological heterogeneity

27

28

29

Abstract

30 To characterize heterogeneous geological formation with high-resolution, we propose a
31 surface-wave analysis method for estimation of the spatial variation of S-wave velocity and
32 attenuation coefficients. We applied the proposed method to the 3D ocean bottom cable (OBC)
33 seismic data acquired at the Tomakomai CO₂ storage site in Japan. By inversion of the observed
34 dispersion curves of surface-wave (Scholte-wave in offshore survey), we successfully estimated the
35 high-resolution pseudo-3D S-wave velocity structure for the shallow geological formation. We
36 observed the dominant higher modes of the surface-wave, generated by the complex velocity
37 structure such as inclusions of low-velocity layers. The interpretation based on the dispersion curves
38 (or inverted S-wave velocities) and attenuation coefficients identified the lithological boundary in
39 shallow formation across the 3D seismic survey area, but no fractures were detected. Our proposed
40 method to estimate pseudo-3D S-wave velocity and attenuation structure can provide vital
41 information for the CO₂ geological storage, as it provides useful data for evaluation of CO₂ leakage
42 paths and estimation of permeability heterogeneity used in reservoir simulation. Moreover, it
43 improves the accuracy of reflection seismic profile for deep structural analysis, and enhances
44 estimations of lithology strength used for geomechanical modeling.

45

46

47 **1. Introduction**

48 Surface-wave analysis is a useful tool for characterizing subsurface structure, mainly S-wave
49 velocity structures in shallow formations (e.g., Socco et al., 2010; Xia et al., 2009; Foti et al., 2011;
50 Foti et al., 2014). In surface-wave analysis, experimental dispersion curves (i.e., relationship
51 between phase-velocity and frequency) of surface-wave are extracted from observed seismic data. In
52 particular, S-wave velocity profiles can be well constrained in inversion of observed dispersion
53 curves, because surface waves are much more sensitive to S-wave velocities than P-wave velocities
54 and densities (e.g., Xia et al., 1999).

55 Surface-wave analysis generally investigates geologic formation down to depths that are
56 shallower than the depth range of the seismic reflection method. Although seismic reflection analysis
57 is one of the most effective techniques for deep seismic exploration, subsurface heterogeneity
58 derived from the shallow seismic methods influences the accuracy of extracting deep reflections.
59 Therefore, surface-wave analysis plays an important role in constructing shallow formation models
60 for static correction in reflection analysis (e.g., Roy et al., 2010; Strobbia et al., 2010). Robust
61 near-surface structures obtained from surface-wave analysis also provide useful information for
62 full-wave form inversion or PP-PS waveform matching (Boiero et al., 2013). The S-wave velocity is
63 related to the Poisson's ratio (ratio of P-wave velocity to S-wave velocity), which reflects the
64 lithology strength and pore pressure (Tsuji et al., 2011; Tsuji et al., 2014). In offshore seismic data

65 analysis, surface-wave analysis is particularly effective in reservoir characterization for the
66 following reasons (Foti et al., 2014). The P-wave velocity derived from conventional seismic
67 surveys is not sensitive to the shallow geological structures because the pore compressibility is a
68 dominant factor in the P-wave velocity of unconsolidated sediments. Furthermore, because a source
69 signal close to the sea surface (i.e., an airgun) cannot generate direct S-waves, the S-wave reflection
70 and refraction analyses from offshore data is usually challenging. Using the surface-wave
71 sufficiently generated by artificial sources in offshore environment, S-wave velocity variation in the
72 shallow sediments can be identified (e.g., Foti et al., 2014).

73 Although less attention has been paid to attenuation of surface waves, some studies used
74 attenuation coefficients of surface waves to estimate quality factors of S-waves (e.g., Xia et al.,
75 2002; Lai et al., 2002; Misbah and Strobbia, 2014). In strongly heterogeneous media, however,
76 estimation of the attenuation coefficients would be influenced by energy reflection at the interfaces
77 (Foti, 2004; Bergamo and Socco, 2014). Recently, Bergamo and Socco (2014) demonstrated that
78 sharp changes in the lateral variation of the attenuation coefficients are highly correlated with the
79 locations of lateral discontinuities. Ikeda and Tsuji (submitted) further demonstrated that lateral
80 variation of the attenuation coefficients is more sensitive to localized fractures than lateral variation
81 of phase velocities. They proposed an approach to distinguish between structural boundaries and
82 localized fractures by parallel interpretation of phase velocity with surface-wave attenuation data.

83 In this study, we apply surface-wave analysis to offshore seismic data acquired in the CO₂
84 geological storage site located in Tomakomai, Japan. In the CO₂ geological storage, understanding
85 the heterogeneous structures in shallow formation contributes in the following ways: (1) Improving
86 the accuracy of the deep reflection analysis by providing better static corrections; (2) evaluating the
87 stability of the CO₂ storage site; (3) providing S-wave velocity models to estimate the lithology
88 strength for geomechanical modeling of the formation stability; and (4) predicting permeability for
89 reservoir simulation to estimate future CO₂ distribution. In the CO₂ storage site in Ketzin, Germany,
90 the 2D subsurface S-wave velocity structure was estimated by surface-wave analysis using ambient
91 noise data (Xu et al., 2012). Because a 3D seismic survey was performed in our study area, we
92 extend the surface-wave analysis to the analysis for 3D seismic data. To effectively estimate the
93 heterogeneous structure of the CO₂ storage field, we extract both the dispersion curves and
94 attenuation coefficients at each local point. The pseudo-3D S-wave velocity structure is constructed
95 by inversion of the estimated dispersion curves. Because multicomponent seismic data were
96 acquired at this site, we further investigate the possibility of using the horizontal component data and
97 vertical component data in a joint analysis.
98

99 **2. Method: surface-wave extraction from 3D seismic data**

100 We extend the advanced 2D surface-wave analysis to a 3D seismic data analysis. For the 2D
101 surface-wave analysis, we apply the common mid-point crosscorrelation (CMPCC) analysis for the
102 local phase velocity estimation and the amplitude ratio analysis for the local attenuation coefficient
103 estimation. Note that 2D or 3D surface-wave analysis in this study is a collection of 1D
104 surface-wave analysis at each local point.

105

106 **2.1 CMPCC method**

107 Details of the data processing procedure of the CMPCC method for 2D seismic data can be found
108 in Hayashi and Suzuki (2004) and Ikeda et al. (2013). Possible receiver pairs with the same midpoint
109 (i.e., CMP in surface-wave analysis) are collected from observed shot gathers. Crosscorrelations
110 between two receivers are grouped and defined as a CMPCC gather. Because the CMPCC gather has
111 information of the phase difference between receiver pairs, the local phase velocity near the CMP
112 can be extracted from the CMPCC gather. In this study, the multichannel analysis of surface waves
113 (MASW; Park et al., 1999) using the phase shift method (Park et al., 1998) is used to estimate the
114 dispersion curves from the CMPCC gathers. By repeating the same procedure for other CMPs, the
115 dispersion curves at each local point (i.e., CMP) can be estimated.

116

117 **2.2 Amplitude ratio analysis method**

118 To estimate the local attenuation coefficient with high lateral resolution, we proposed an
119 amplitude ratio analysis method based on the CMPCC analysis (Ikeda and Tsuji, submitted; Ikeda
120 and Tsuji, 2015). In the amplitude ratio analysis, we focus on amplitude decay between receiver
121 pairs on the CMP. After correcting for geometrical spreading of the surface waves, the amplitude
122 ratio data between two receivers are grouped and defined as an amplitude ratio gather. Similar to
123 local phase velocity estimation in CMPCC analysis, the local attenuation coefficient can be
124 estimated from the amplitude ratio gather. To estimate the attenuation coefficients from the
125 amplitude ratio gathers, we perform linear regression of the receiver spacing versus amplitude ratio
126 data in a natural logarithmic scale (e.g., Xia et al., 2002). Thus, both the local phase velocity and
127 attenuation coefficient can be estimated at the same CMPs.

128 If N receivers are employed in the data acquisition, $N(N-1)/2$ receiver pairs can be generated from
129 each shot gather. For the 3D seismic data analysis, therefore, a large number of receiver pairs can be
130 collected at each CMP. If we collect receiver pair data in the time domain and extract surface waves
131 from the gathered data, large amounts of memory space and huge computational time are required.
132 To overcome this problem, we generate the gathered data in the frequency domain. In the frequency
133 domain, the number of useful data samples for surface-wave extraction is much smaller than that in
134 the time domain (Ikeda and Tsuji, submitted).

135

136 **2.3 Evaluation of error caused by the plane wave assumption in 3D data analysis**

137 For simplicity, we apply the CMPCC analysis and amplitude ratio analysis to the 3D seismic data
138 assuming that the observed seismic data are composed of surface waves propagating as plane waves
139 for the CMPs. However, the source and receiver positions are not aligned in a straight line in the 3D
140 data acquisition. Therefore, the plane wave assumption will generate some errors in the phase
141 velocity and attenuation coefficient estimations because the surface waves have a cylindrical
142 wavefront (Fig. 1). Here, we quantitatively evaluate the errors caused by the plane wave assumption
143 and propose an effective criteria for selection of source-receiver pairs in 3D data analysis. The CMP
144 in surface-wave analysis can be defined at the midpoint between two receivers (Fig. 1). Considering
145 the cylindrical wavefront of the surface waves, the phases of the Fourier transform at receivers 1 and
146 2 in Fig. 1b (R_1 and R_2 , respectively) can be expressed as

$$\begin{aligned}\phi_1 &= \phi_0 + \frac{2\pi f}{c} r_{c1} \\ \phi_2 &= \phi_0 + \frac{2\pi f}{c} r_{c2}\end{aligned}, \tag{1}$$

147 where ϕ_1 and ϕ_2 are the phases of R_1 and R_2 , respectively. ϕ_0 , f , and c are the initial phase, the
148 frequency, and the phase velocity, respectively, and

$$r_{c1} = \sqrt{\left(\frac{r}{2}\sin\theta\right)^2 + \left(d + \frac{r}{2}\cos\theta\right)^2}$$

$$r_{c2} = \sqrt{\left(\frac{r}{2}\sin\theta\right)^2 + \left(d - \frac{r}{2}\cos\theta\right)^2}, \quad (2)$$

149 are the distances between the source and R_1 and R_2 , respectively (Fig. 1b). In Eq. (2), d , r , and θ are
 150 the distance between the source and the CMP, the receiver spacing between R_1 and R_2 , and the angle
 151 between the R_1 - R_2 line and the line between the source and the CMP, respectively. The phase
 152 difference $\Delta\phi$ between R_1 and R_2 can be expressed as

$$\Delta\phi = \phi_1 - \phi_2$$

$$= \frac{2\pi f}{c}(r_{c1} - r_{c2}) \quad (3)$$

153 As described previously, we assume that the surface waves propagate as plane waves across the
 154 CMP (Fig. 1c). The propagation distance (receiver spacing) of the plane wave between R_1 and R_2
 155 can be expressed as

$$\Delta r = r\cos\theta. \quad (4)$$

156 The phase velocity estimated with the plane wave assumption can be expressed as

$$c_{est} = \frac{2\pi f}{\Delta\phi}\Delta r. \quad (5)$$

157 Substituting Eqs. (2) - (4) into Eq. (5), we obtain

$$c_{est} = \beta c, \quad (6)$$

158 where

$$\beta = \frac{1}{4} \left(\sqrt{\left(\frac{r}{d} \sin \theta\right)^2 + \left(2 + \frac{r}{d} \cos \theta\right)^2} + \sqrt{\left(\frac{r}{d} \sin \theta\right)^2 + \left(2 - \frac{r}{d} \cos \theta\right)^2} \right). \quad (7)$$

Equations (6) and (7) indicate that the deviation of the estimated phase velocity from the true phase velocity is dependent only on $\frac{d}{r}$ and θ . By applying a similar approach for the attenuation coefficient α , we obtain

$$\alpha_{est} = \frac{1}{\beta} \alpha, \quad (8)$$

where α_{est} is the estimated attenuation coefficient with the plane wave assumption.

Based on Eqs. (6) and (8), we define the absolute value of $\beta - 1$ as the error in the phase velocity estimation and $\frac{1}{\beta} - 1$ as the error in the attenuation coefficient estimation. The errors become smaller as θ becomes smaller and $\frac{d}{r}$ becomes larger (Fig. 2). When $\theta = 30^\circ$ and $\frac{d}{r} = 1$, the corresponding errors are 3.72 % for the phase velocity estimation and 3.58 % for the attenuation coefficient estimation. To minimize the errors at this level, we select the seismic traces satisfying the following relationship

$$\theta < 30^\circ \text{ and } \frac{d}{r} > 1. \quad (9)$$

170 **3. Data and Results**

171 **3.1 Ocean Bottom Cable (OBC) Data**

172 We apply the proposed surface-wave analyses to the 3D marine seismic data acquired in the
173 Tomakomai CCS project (Fig. 3; Abe et al., 2013; Tanase et al., 2013). From 2016 to 2018, CO₂ over
174 100,000 tonnes/year will be injected into subseabed reservoirs close to shoreline of the Tomakomai
175 city, Japan (Tanase et al., 2013). The seafloor depth in the seismic survey area is 7-40 m. The seismic
176 data were recorded by ocean bottom cable (OBC) with 3C accelerometers and hydrophones (Sercel
177 SeaRay). Seismic waves were generated by an airgun array located 2 m below the sea surface at a
178 seafloor depth of < 10 m, or 6 m below the sea surface at a seafloor depth of > 10 m. The record
179 length was 6.5 sec with a sampling rate of 2 ms. Eight receiver lines were aligned in the N-S
180 direction and each line consisted of 120 or 125 receivers (Fig. 3). The receiver intervals for the N-S
181 and E-W directions were ~ 25 and ~ 300 m, respectively.

182 We carefully analyze the observed seismic data for the surface-wave analyses because the data
183 acquisition was originally designed for a seismic reflection survey. The surface-wave trains that have
184 longer offsets are cut because of the limited recording time (i.e., 6.5 s). To include the surface-wave
185 trains in the observed seismic data, we analyze the seismic data up to 500 m offset. The CMP
186 positions are defined at every 62.5 m in the N-S directions for all lines (blue circles in Fig. 3). To
187 extract the phase velocities and attenuation coefficients, the CMPCC gather and amplitude ratio

gather are computed from the receiver pair data at each CMP with bin size of 12.5×12.5 m. As described in the previous section, only seismic data satisfying Eq. (9) is used in this study.

The use of receiver pairs in the E-W direction is limited because the receiver distribution in the E-W direction was sparse (~ 300 m interval). Our analysis is therefore limited to receiver pairs aligned in an almost N-S direction. Considering Eq. (9), the sources used for our analysis are mostly located at the north side or south side of the CMPs. By separately extracting surface waves depending on the source directions relative to the CMPs, we can investigate the source direction-dependent characteristics of the surface waves in the N-S direction. Indeed, the lateral variation of the attenuation coefficients depending on the source directions is a useful tool for identifying heterogeneous structures such as sharp lateral discontinuities (Bergamo and Socco, 2014; Ikeda and Tsuji, submitted). In this study, therefore, we perform attenuation analysis depending on source directions. We define positive- and negative-offset data for sources located at the south side and north side of the CMPs, respectively.

3.2. Dispersion curve

We demonstrate the application of the proposed surface-wave analyses using the vertical components of the observed seismic data (OBC data). We compute the dispersion spectrum for the phase velocities from 100 to 1000 m/s (step size 1 m/s). By picking the phase velocities with the

206 maximum amplitude at each frequency, we obtain the dispersion curves for the dominant mode at
207 each CMP.

208 We can extract the dispersion curves in the frequency range from ~1 to 6 Hz. The high-frequency
209 limit would be caused by strong attenuation of the high-frequency surface waves, the effect of spatial
210 aliasing, and the limited recording time. In the south side of the study area, the fundamental mode of
211 the surface-wave is mostly dominant (e.g., Fig. 4a). In the north side, however, higher modes of
212 surface waves are dominant at some frequencies (e.g., Fig. 4b). On the map view of the frequency
213 slice of the phase velocity at 2.0 Hz (Fig. 5b), we observe clear lateral variations of the phase
214 velocity in the NW-SE direction and E-W direction. The boundary in NW-SE direction can also be
215 observed on the map view of the higher frequencies slice (or probably shallower slice) (e.g., 2.9 Hz;
216 Fig. 5c). The high phase velocity area aligned in the E-W direction shifts from the north to the
217 middle of our study area as the frequency increases (Fig. 5). Such high phase velocity areas are
218 caused by the dominance of the higher modes. The dominance of the higher mode surface waves can
219 be induced by complex underground structures such as a low-velocity layer inclusion (e.g., Gucunski
220 and Woods, 1992; Tokimatsu et al., 1992; Tsuji et al., 2012). Thus, spatial heterogeneity in a shallow
221 formation can be observed as the spatial variation of the phase velocities. We further construct a
222 pseudo-3D S-wave velocity model by inverting the observed dispersion curves, as described in
223 section 3.4.

224

225 3.3. Attenuation coefficients

226 We estimate the attenuation coefficients from the amplitude ratio gathers at each CMP. Two
227 attenuation coefficient curves are estimated at each CMP from the positive-offset data and
228 negative-offset data. To extract a trend in the data, we compute the mean values of the amplitude
229 ratio data in a natural logarithmic scale within each 25 m receiver spacing bin for the specified
230 receiver spacing, discarding bins with fewer than 5 data samples. The values of the attenuation
231 coefficients are estimated by linear regression using the mean values of the amplitude ratio data
232 (Ikeda and Tsuji, submitted; Ikeda and Tsuji, 2015).

233 To emphasize the location of large lateral heterogeneity from the source direction dependence of
234 the estimated attenuation coefficients, we introduce the following indicator

$$\gamma = \alpha_+ - \alpha_-, \quad (10)$$

235 where α_+ and α_- are the attenuation coefficients estimated from the positive-offset data and
236 negative-offset data, respectively. When the difference between α_+ and α_- is large, the absolute
237 γ value becomes larger, indicating the existence of sharp lateral discontinuities (Bergamo and
238 Socco, 2014).

239 We observe spatial variation of the estimated attenuation coefficients, as shown in Fig. 6. As the
240 frequency increases, the estimated attenuation coefficients become larger at most local points,

241 probably reflecting shallower sequences with low quality factor. At 2.0 Hz, the lateral variation
242 (frequency slice) of the estimated attenuation coefficients shows a lateral boundary in the NW-SE
243 direction (Figs. 6a and 6c). Near the NW-SE boundary, however, the increasing trend in the
244 attenuation coefficients from the positive-offset data (Fig. 6a) is opposite to that from the
245 negative-offset data (Fig. 6c). The opposite trend anomaly in the estimated attenuation coefficients
246 can be more clearly identified as the large absolute γ values (Fig. 7a). The locations with large
247 absolute γ values indicate the existence of large heterogeneities in shallower depth (Fig. 7b). On the
248 frequency slice of 2.9 Hz, the NW-SE anomaly is not clear in the lateral variation of the estimated
249 attenuation coefficients (Figs. 6b and 6d).

250

251 **3.4. Pseudo-3D S-wave velocity model by inversion**

252 The pseudo-3D S-wave velocity model is constructed by interpolating 1D velocity structures
253 inverted from the dispersion curves at the CMPs. As the inversion method, we employ a genetic
254 algorithm (GA) inversion (Goldberg, 1989; Sen and Stoffa, 1995) with dynamic mutation and elite
255 selection (Yamanaka and Ishida, 1996). We carry out a multimode inversion by using the amplitude
256 response of surface waves (Harkrider, 1964, 1970) without identifying observed mode number. In
257 the multimode inversion, we theoretically calculate relative amplitude for each mode, and the
258 theoretical phase velocities are defined as the velocities for the mode with the maximum amplitude

259 for a given frequency (e.g., Hayashi and Saito, 2004; Lu and Zhang, 2006; Tsuji et al., 2012; Ikeda et
260 al., 2013). We construct reference S-wave velocity models with six layers for geological formation
261 by applying 1/3 wavelength transformation to the observed dispersion curves (e.g., Heisey et al.,
262 1982; Hayashi, 2008). The thickness of each layer in the reference models is 20 m. The search range
263 of the S-wave velocity and thickness is $\pm 40\%$ for the reference models. The P-wave velocity and
264 density are determined by the empirical equations of Kitsunezaki et al. (1990) and Ludwig et al.
265 (1970), respectively. Although the above inversion procedure was originally developed for
266 Rayleigh-wave inversion, we apply the method for Scholte-wave inversion. Scholte waves are
267 modeled by including an additional water layer over the solid layered model (e.g., Foti et al., 2014).
268 Note that when the P-wave velocities determined by the empirical equation are < 1500 m/s, they are
269 replaced by 1500 m/s (i.e., sound speed within water) because the survey area is in the offshore
270 region. The compound matrix method (Saito, 1988; Saito and Kabasawa, 1993) is used to compute
271 the theoretical Scholte-wave dispersion curves and amplitude response. In the GA inversion, both the
272 number of generations and the population is 200. Twenty trials are performed with the random seeds
273 of an initial population. By averaging the 20 inverted velocity models, the smoothed S-wave velocity
274 structures are estimated at each CMP (e.g., red lines in Fig. 8). Before the inversion, we remove the
275 low-quality phase velocity data based on the values of the dispersion spectrum (see details in
276 Appendix A). If the observed mode branches are not well retrieved by the multimode inversion,

277 those phase velocities are manually removed and we repeat the inversion.

278 When the fundamental mode of the dispersion curves is dominant, the observed dispersion curves
279 are good agreement with the fundamental mode of the theoretical dispersion curves for the inverted
280 models (e.g., Fig. 8a). The observed dispersion curves with dominant higher modes are also well
281 retrieved by the multimode inversion (e.g., Fig. 8b). The pseudo-3D S-wave velocity structure (Fig.
282 9) is constructed by interpolating the inverted velocity structures. In the interpolation, we use 1D
283 S-wave velocity structures in the depth range (below seafloor) from 1/3 of the minimum wavelength
284 to 1/3 of the maximum wavelength of the observed dispersion curves at each CMP. We interpolate
285 the 1D S-wave velocities every 0.5 m in the vertical direction and every 50 m in the N-S and E-W
286 directions. The low S-wave velocity interval can be observed at ~100 m in depth in the north part of
287 the inverted model (Fig. 9a). This interbedded low velocity zone causes the dominant higher mode in
288 the dispersion curves. The boundary in the NW-SE direction can be observed in the inverted S-wave
289 velocity model (Fig. 9b) as observed in the estimated dispersion curves and attenuation coefficients
290 (Figs. 5-7).

291

292 **3.5. Horizontal component data analysis**

293 In the sections 3.2-3.4, we demonstrate the proposed surface-wave analyses only using the vertical
294 component data. Because the horizontal component data were also acquired in the Tomakomai site,

295 we investigate the possibility of the extracting surface-wave data included in the horizontal
296 component records. To estimate the Scholte and Love waves from the horizontal component data, we
297 extract the radial and transverse components from two horizontal components in the N-S and E-W
298 directions. The Scholte- and Love-wave dispersion curves and attenuation coefficients at each CMP
299 are extracted from the horizontal component data by the same procedure used for the vertical
300 component data analysis.

301 The Love-wave dispersion curves estimated from the transverse component data at point C in Fig.
302 3 are not clearly observed in the dispersion spectrum (Fig. 10c), compared with the Scholte-wave
303 dispersion curves estimated from the vertical or radial component data (Figs. 10a and 10b). This is
304 probably because explosive sources do not generate SH waves. At point C in Fig. 3, two modes are
305 observed in the dispersion curve of the radial component data (Fig. 10b), whereas only the
306 fundamental mode is observed in the vertical component data (Figs. 10a and 10d). The use of
307 additional mode information extracted from the radial component data would improve the accuracy
308 of the S-wave velocity estimation as discussed in the Rayleigh-wave analysis (Ikeda et al., 2015).
309 Here, we demonstrate the joint inversion of Scholte waves estimated from vertical and radial
310 component data. When only the fundamental mode estimated from the vertical component data is
311 used in the inversion, the first higher mode of the inverted models is not constrained (Fig. 11a). On
312 the other hand, when we include the first higher mode extracted from the radial component data,

313 both observed modes are well fitted to the theoretical dispersion curves of each mode (Fig. 11b). The
314 resulting S-wave velocity structure is more accurately estimated, and differs from the result derived
315 only from the fundamental mode. Note that in the joint inversion in Fig. 11b, we perform a different
316 type of multimode inversion from that described above. In the inversion, we explicitly assign the
317 observed two modes as the fundamental mode and first higher mode because they are clearly
318 identified at point C in the multicomponent data (Fig. 10).

319 The spatial variation of the phase velocities estimated from the horizontal component data is
320 different from that estimated from the vertical component data (Fig. 12). The Scholte-wave phase
321 velocities from the radial component data (Figs. 12a-12c) seem to be higher than those from the
322 vertical component data (Fig. 5). This is because the higher-order Scholte-wave modes are extracted
323 from the radial component data. The lateral boundary aligned in the NW-SE direction can be
324 observed in both results derived from the Scholte waves using horizontal component data and Love
325 waves (Fig. 12).

326 The spatial variation of the γ values estimated from the horizontal component data also differs
327 from that estimated from the vertical component data (Fig. 13). This difference can be caused by the
328 different sensitivity of the higher-order Scholte waves or Love waves to geological heterogeneity.
329 Although large heterogeneity is observed in the result derived from the horizontal component data,
330 the NW-SE boundary is still interpreted as the high absolute γ value.

331

332 **4. Interpretation and Discussion**

333 In this study we propose a method to identify spatial variation in a shallow geological formation.
334 The lateral boundary in the NW-SE direction can be detected from both the dispersion curves and
335 attenuation coefficients (Figs. 5 and 6). Along this boundary, we observe the opposite trend of the
336 estimated attenuation coefficients depending on the source directions (Figs. 6 and 7). As discussed
337 by Ikeda and Tsuji (submitted), lateral variations of both the dispersion curves and attenuation
338 coefficients can be observed near lithological boundaries, reflecting lithological difference. The
339 lateral variation of the source direction-dependent attenuation coefficients can be also observed
340 because of the scattering at the lithological boundary. Therefore, the boundary in the NW-SE
341 direction extracted from our analysis can be characterized as a lithological boundary. The S-wave
342 velocity model (Fig. 9) and the results of horizontal data analyses (Figs. 12 and 13) further support
343 the existence of a lithological boundary in the NW-SE direction.

344 The spatial variation of the estimated dispersion curves differs from that of the estimated
345 attenuation coefficients. This can be caused by sensitivity differences between the attenuation and
346 phase velocity to structural heterogeneities as demonstrated by Ikeda and Tsuji (submitted). They
347 showed that the attenuation coefficients can be increased near localized fractures and the lateral
348 variation is more sensitive to fractures than the phase velocity. At high frequencies (i.e., shallow
349 structures), the spatial variation of the estimated attenuation coefficients and γ shows a more

350 heterogeneous pattern than the phase velocity result (e.g., Figs. 6 and 7). However, anomalies with
351 increase of attenuation coefficients are not continuously observed in both positive- and
352 negative-offset data over the analyzed frequency range. This result indicates that the probability of
353 the existence of a continuous fracture in the shallow formation is low. This information is extremely
354 helpful for detecting localized fractures, which could be leakage paths for injected CO₂ in CCS
355 projects. Because the heterogeneity of the permeability structure has a strong influence on the CO₂
356 flow, the high-resolution S-wave velocity data related to permeability is useful information in
357 geological model construction for reservoir simulation.

358 In our dataset, the seismic receivers were densely deployed only in the N-S direction. If 3D
359 seismic data were densely acquired in a wide variety of directions, the proposed 3D surface-wave
360 analyses can be applied to estimate azimuthal-dependent surface waves. As demonstrated by
361 Alvizuri and Tanimoto (2011) and Ikeda and Tsuji (2014) using teleseismic data, azimuthal
362 anisotropic phase velocity can be effective in estimating underground stress conditions. In CCS
363 projects, such anisotropic properties are important in evaluating fracture orientation and intensity,
364 which directly controls the injected CO₂ flow.

365

366 **5. Conclusions**

367 We propose a surface-wave analysis scheme using 3D OBC data to estimate localized
368 heterogeneity in shallow formation. The proposed method is applied to the 3D marine seismic data
369 acquired at the CO₂ storage site in Tomakomai, Japan.

370 We successfully estimate the spatial variation of the subsurface structure by using vertical
371 component seismic data. The spatial variation of the estimated dispersion curves and attenuation
372 coefficients identifies the lithological boundary in the NW-SE direction. The NW-SE boundary can
373 also be estimated in the pseudo-3D S-wave velocity volume constructed by multimode inversion of
374 Scholte-wave dispersion curves. From the northern to the middle part of the study area, a high phase
375 velocity anomaly can be observed as a result of dominant higher modes. Multimode inversion results
376 suggest that the higher modes are derived from the reversal of S-wave velocities in the shallow
377 formation.

378 Scholte and Love waves are also extracted from the horizontal component data. We demonstrate
379 that the spatial variation of the phase velocities and attenuation coefficients estimated from the
380 horizontal component data differs from that estimated from the vertical component data. Joint
381 analysis using data of different components would contribute to the construction of a robust
382 subsurface heterogeneity model. We demonstrate the possibility of improving S-wave velocity

383 estimation by using the additionally estimated multimode Scholte waves derived from the radial
384 component data.

385 The estimated heterogeneity in a shallow formation can provide important information for CCS
386 projects. The parallel interpretation of the spatial variation of the estimated attenuation coefficients
387 with the phase velocities suggests a low probability of the existence of localized fractures that may
388 serve as leakage paths for injected CO₂. The constructed pseudo-3D S-wave velocity structure has
389 the potential to be used for improving the accuracy of the retrieval of deep reflection from the
390 reservoirs. The estimated subsurface properties can be further used for estimating the lithology
391 strength (for geomechanical modeling) or in CO₂ flow simulations.

392

393 **Acknowledgements**

394 We thank associate editor Charles Jenkins and two anonymous reviewers for constructive
395 comments that improved the manuscript. We gratefully acknowledge the Ministry of Economy,
396 Trade and Industry (METI) and the Japan CCS Co., Ltd. (JCCS); the seismic data used in this study
397 was acquired under the project of JCCS commissioned by METI. This work is partially supported by
398 and JST/JICA-SATREPS, and JSPS Grant-in-Aid for Scientific Research on Innovative Areas
399 (15H01143). This work is further supported by the International Institute for Carbon Neutral Energy
400 Research, sponsored by the World Premier International Research Center Initiative, MEXT, Japan.
401

402 **Appendix A**

403 Here, we describe how we remove the low-quality phase velocity data. The dispersion spectrum U
404 by the MASW method using the phase shift method (Park et al., 1998 and 1999) can be defined as

$$U(c, f) = \left| \int_{-\infty}^{+\infty} \frac{F(x, f)}{|F(x, f)|} e^{i2\pi f x / c} dx \right|, \quad (\text{A1})$$

405 where F is the CMPCC gather and x is the receiver spacing. As the receiver spacing in the
406 CMPCC gather is the finite and discrete value, Eq. (A1) should be expressed as

$$U(c, f) = \left| \sum_{i=1}^N \frac{F(x_i, f)}{|F(x_i, f)|} e^{i2\pi f x_i / c} \right|, \quad (\text{A2})$$

407 where N is the number of traces in the CMP gather. We define the normalized dispersion spectrum
408 U_{norm} as

$$U_{norm}(c, f) = U(c, f) / N. \quad (\text{A3})$$

409 U_{norm} is ranged from 0 to 1. In our inversion analysis, we remove phase velocity data with U_{norm}
410 < 0.4 .

411

412

References

- 413 Abe, M., Saito, S., Tanase, D., Sawada, Y., Hiram, Y., Motoyama, Y., 2013. CCS large-scale
414 demonstration in Japan, Energy Procedia 37, 6326-6334, doi:10.1016/j.egypro.2013.06.561.
- 415 Alvizuri, C., Tanimoto, T., 2011. Azimuthal anisotropy from array analysis of Rayleigh waves in
416 Southern California. Geophys. J. Int. 186(3), 1135-1151.
- 417 Boiero, D., Wiarda, E., Vermeer, P., 2013. Surface-and guided-wave inversion for near-surface
418 modeling in land and shallow marine seismic data. Leading Edge 32(6), 638-646.
- 419 Bergamo, P., Socco L. V., 2014. Detection of sharp lateral discontinuities through the analysis of
420 surface-wave propagation. Geophysics 79(4), EN77-EN90.
- 421 Foti, S., 2004. Using transfer function for estimating dissipative properties of soils from
422 surface-wave data. Near Surf. Geophys. 2(4), 231-240.
- 423 Foti, S., Parolai, S., Albarello, D., Picozzi, M., 2011. Application of surface-wave methods for
424 seismic site characterization. Surv. Geophys. 32(6), 777-825.
- 425 Foti, S., Lai, C. G., Rix, G. J., Strobbia, C. 2014. Surface Wave Methods for Near-Surface Site
426 Characterization. CRC Press.
- 427 Goldberg, D. E., 1989. Genetic Algorithms in Search, Optimization, and Machine Learning.
428 Addison-Wesley, Reading, MA.
- 429 Gucunski, N., Woods, R. D., 1992. Numerical simulation of the SASW test, Soil Dyn. Earthq. Eng.

430 11, 213-227. Harkrider, D.G., 1964. Surface waves in multilayered elastic media, Part 1. Rayleigh
431 and Love waves from buried sources in a multilayered elastic half-space. Bull. Seismol. Soc. Am.
432 54, 627-679.

433 Harkrider, D.G., 1970. Surface waves in multilayered elastic media, Part 2. Higher mode spectra
434 ratios from point sources in plane layered earth models. Bull. Seismol. Soc. Am. 60, 1937-1987.

435 Hayashi, K., (Ph.D. dissertation) 2008. Development of surface-wave methods and its application to
436 site investigations. Kyoto University.

437 Hayashi, K., Saito, M., 2004. The dispersion curve of P-SV wave propagation with a high velocity
438 top layer and its analysis. Proceedings of the 110th SEGJ Conference 11, 39-42.

439 Hayashi, K., Suzuki, H., 2004. CMP cross-correlation analysis of multi-channel surface-wave data.
440 Exploration Geophys. 35(1), 7-13.

441 Heisey, J. S., Stokoe, K. H., II, Hudson, W. R., Meyer, A. H., 1982. Determination of in situ shear
442 wave velocities from spectral analysis of surface waves. Res. Rep. 256-2, Ctr. for Transp.
443 Res., Univ. of Texas, Austin, TX.

444 Ikeda, T., Tsuji, T., Matsuoka, T., 2013. Window-controlled CMP crosscorrelation analysis for
445 surface waves in laterally heterogeneous media. Geophysics 78(6), EN95-EN105.

446 Ikeda, T., Tsuji, T., 2014. Azimuthal anisotropy of Rayleigh waves in the crust in southern Tohoku
447 area, Japan. J. Geophys. Res. 119 (12), 8964-8975.

448 Ikeda, T., Matsuoka, T., Tsuji, T., T., Nakayama, 2015. Characteristics of the horizontal component
 449 of Rayleigh waves in multimode analysis of surface waves. *Geophysics* 80(1), EN1-EN11.

450 Ikeda, T., Tsuji, T. Surface wave attenuation in shallow formation from multichannel-multishot
 451 seismic data: A new approach for detecting localized fracture and lithological discontinuity.
 452 submitted to *Earth Planets Space*.

453 Ikeda, T., and Tsuji, T, 2015. Detecting localized fracture and lithological boundary by the analysis
 454 of spatial variation of surface wave attenuation estimated from multichannel-multishot data. The
 455 Symposium on the Application of Geophysics to Engineering and Environmental Problems
 456 (SAGEEP 2015), Austin, USA.

457 Kitsunezaki, C., Goto, N., Kobayashi, Y., Ikawa, T., Horike, M., Saito, T., Kurota, T., Yamane, K.,
 458 Okuzumi, K., 1990. Estimation of P- and S-wave velocities in Deep Soil Deposits for Evaluating
 459 Ground Vibrations in Earthquake(in Japanese). *Sizen-saigai-kagaku* 9-3, 1-17.

460 Lai, C.G., Rix, G.J., Foti, S., Roma, V., 2002. Simultaneous measurement and inversion of surface
 461 wave dispersion and attenuation curves. *Soil Dyn. Earthq. Eng.* 22(9), 923-930.

462 Lu, L., Zhang, B., 2006. Inversion of Rayleigh waves using a genetic algorithm in the presence of a
 463 low-velocity layer. *Acoust. Phys.* 52(6), 701-712.

464 Ludwig, W. J., Nafe, J. E. Drake, C. L., 1970. *Seismic Refraction in the Sea*. Vol. 4, part 1,
 465 Wiley-Interscience, p. 74.

466 Misbah, A. S., Strobbia, C. L., 2014. Joint estimation of modal attenuation and velocity from
467 multichannel surface wave data. *Geophysics* 79(3), EN25-EN38.

468 Park, C. B., Miller, R. D., Xia, J., 1998. Imaging dispersion curves of surface waves on
469 multi-channel record. 68th Annual International Meeting, SEG, Expanded Abstracts, 1377–1380.

470 Park, C. B., Miller, R. D., Xia, J., 1999. Multichannel analysis of surface waves. *Geophysics* 64,
471 800–808.

472 Roy, S., Stewart, R. R., Al Dulaijan, K., 2010. S-wave velocity and statics from ground-roll
473 inversion. *Leading Edge* 29(10), 1250-1257.

474 Saito, M., 1988. DISPER80: a subroutine package for calculation of seismic normal-mode solution.
475 in *Seismological Algorithm*, ed. Doorbos, D.J., Academic Press, San Diego, CA.

476 Saito, M., Kabasawa, H., 1993. Computation of reflectivity and surface wave dispersion curves for
477 layered media; 2. Rayleigh wave calculations (in Japanese). *Butsuri-tansa* 46, 283–298.

478 Sen, M. K., Stoffa, P. L., 1995. *Global optimization methods in geophysical inversion*. Elsevier.

479 Socco, L. V., Foti, S., Boiero, D., 2010. Surface-wave analysis for building near-surface velocity
480 models - Established approaches and new perspectives. *Geophysics* 75(5), 75A83-75A102

481 Strobbia, C., El Emam, A., Al-Genai, J., Roth, J., 2010. Rayleigh wave inversion for the near-surface
482 characterization of shallow targets in a heavy oil field in Kuwait. *First Break* 28(5), 103-109.

483 Tanase, D., Sasaki, T., Yoshii, T., Motohashi, S., Sawada, Y., Aramaki, S., Yamanouchi, Y., Tanaka,

484 T., Ohkawa, S., Inowaki, R., 2013. Tomakomai CCS demonstration project in Japan, Energy
 485 Procedia 37, 6571-6578, doi:10.1016/j.egypro.2013.06.589.

486 Tokimatsu, K., Tamura, S., Kojima, H., 1992. Effects of multiple modes on Rayleigh wave
 487 dispersion characteristics, J. Geotech. Eng., ASCE 118, 1529-1543.

488 Tsuji, T., Dvorkin, J., Mavko, G., Nakata, N., Matsuoka, T., Nakanishi, A., Kodaira, S., Nishizawa,
 489 O., 2011. Vp/Vs ratio and shear-wave splitting in the Nankai Trough seismogenic zone: Insights
 490 into effective stress, pore pressure and sediment consolidation, Geophysics 76(3), WA71-WA82.

491 Tsuji, T., Johansen, T. A., Ruud, B. O., Ikeda, T., Matsuoka, T., 2012. Surface-wave analysis for
 492 identifying unfrozen zones in subglacial sediments. Geophysics 77(3), EN17-EN27.

493 Tsuji, T., Kamei, R., Pratt, G., 2014. Pore pressure distribution of a mega-splay fault system in the
 494 Nankai Trough subduction zone: Insight into up-dip extent of the seismogenic zone, Earth Planet.
 495 Sci. Lett. 396, 165-178, doi:10.1016/j.epsl.2014.04.011.

496 Xia, J., Miller, R. D., Ivanov, J., Zeng, C., 2009. High-frequency Rayleigh-wave method. J. Earth Sci.
 497 20(3), 563-579.

498 Xia, J., Miller, R. D., Park, C. B., 1999. Estimation of near-surface shear-wave velocity by inversion
 499 of Rayleigh waves. Geophysics 64(3), 691-700.

500 Xia, J., Miller, R. D., Park, C. B., Tian, G., 2002. Determining Q of near-surface materials from
 501 Rayleigh waves. J. Appl. Geophys. 51(2), 121-129.

502 Xu, Z., Juhlin, C., Gudmundsson, O., Zhang, F., Yang, C., Kashubin, A., Lüth, S., 2012.
503 Reconstruction of subsurface structure from ambient seismic noise: an example from Ketzin,
504 Germany. *Geophys. J. Int.* 189(2), 1085-1102.

505 Yamanaka, H., Ishida, H., 1996. Application of genetic algorithms to an inversion of surface-wave
506 dispersion data, *Bull. Seismol. Soc. Am.* 86, 436–444.

507

Figure Captions

508

509

510

511 **Figure 1.** The geometry among the source (S) and receiver pairs (R_1 and R_2). (b) Cylindrical

512 wavefront. (c) Plane wave assumption.

513

514 **Figure 2.** Errors caused by the plane wave assumption when estimating (a) phase velocity and (b)

515 attenuation coefficient. White circles are plotted at $\theta = 30^\circ$ and $\frac{d}{r} = 1$.

516

517 **Figure 3.** Location of the Tomakomai CO₂ storage site (star). Locations of source (green), receiver

518 (black) and CMP (blue). Magenta line indicates the location of the 2D profile displayed in

519 the following figures. Black dashed lines indicate the locations of the injection wells.

520

521 **Figure 4.** Dispersion spectrum at (a) point A and (b) point B in Fig. 3. The background color

522 represents the dispersion spectrum normalized by the number of receiver pairs. The red

523 dots represent the phase velocities picked for the inversion.

524

525 **Figure 5.** (a) Observed phase velocities along the receiver line (magenta in panels (b) and (c)). (b)

526 and (c) Map view (i.e., frequency slice) of the observed phase velocities at 2.0 Hz and 2.9

527 Hz. The background color is calculated by the linear interpolation of the observed phase
528 velocities at CMPs (black circles).

529

530 **Figure 6.** (a) and (c) Map view (i.e., frequency slice) of the estimated attenuation coefficients from
531 positive- and negative-offset data at 2.0 Hz. The background color is calculated by the
532 linear interpolation of the observed attenuation coefficients at CMPs (black circles). (b)
533 and (d) Same as panels (a) and (c) but for 2.9 Hz.

534

535 **Figure 7.** (a) and (b) Map view (i.e., frequency slice) of the observed γ values at 2.0 Hz and 2.9 Hz.
536 The background color is calculated by the linear interpolation of the observed γ values at
537 CMPs (black circles).

538

539 **Figure 8.** Inversion results at (a) point A and (b) point B in Fig. 3. Left panel represents the observed
540 dispersion curve (black) and theoretical dispersion curves of the inverted models for 20
541 trials (blue). Right panel represents the reference model of S-wave velocity (gray), the
542 inverted models for 20 trials (blue) and the smoothed inverted model (red).

543

544 **Figure 9.** (a) The inverted S-wave velocity model along the receiver line (magenta in Fig. 3). The
545 black line represents the seafloor. (b) Pesuedo-3D S-wave velocity structure constructed

546 by interpolation of the 1D inverted velocity models. Note that the S-wave velocity is
547 averaged over 1.5 and 150 m for the vertical and horizontal directions ($3 \times 3 \times 3$ grids),
548 respectively. The S-wave velocity model without averaging is shown in Fig. A.

549

550 **Figure 10.** (a)-(c) Dispersion spectrum estimated from vertical, radial and transverse component
551 data at point C in Fig. 3, respectively. The background color in the panels represents the
552 dispersion spectrum normalized by the number of receiver pairs. The red circles represent
553 the picked phase velocities in panel (d). (d) Scholte wave phase velocities picked from the
554 dispersion spectrum of the vertical and radial component data.

555

556 **Figure 11.** (a) Inversion results at point C in Fig. 3. Left panel represents the observed dispersion
557 curve (black) and theoretical dispersion curves of the fundamental mode and first higher
558 mode of the inverted models for 20 trials (blue). Right panel represents the reference
559 model of S-wave velocity (gray), the inverted models for 20 trials (blue) and the smoothed
560 inverted model (red). (b) Same as (a) but the first higher mode extracted from the radial
561 component data is included in inversion.

562

563 **Figure 12.** (a)-(c) Same as Fig. 5 but for Scholte waves from the radial (horizontal) component data.

564 (d)-(f) Same as Fig. 5 but for Love waves from the transverse (horizontal) component data.

565

566 **Figure 13.** (a) and (b) Same as Fig. 7 but for Scholte waves from the radial (horizontal) component

567 data. (c) and (d) Same as Fig. 7 but for Love waves from the transverse (horizontal)

568 component data.

569

570 **Figure A.** Same as Fig. 9 but without averaging.

571

Figure 1

Figure 2

(a) Phase Velocity

(b) Attenuation Coefficient

Figure 3

Figure 4

Figure 5

(b) 2.0 Hz

(c) 2.9 Hz

Figure 6

(a) 2.0 Hz for positive-offset data

(b) 2.9 Hz for positive-offset data

(c) 2.0 Hz for negative-offset data

(d) 2.9 Hz for negative-offset data

Figure 7

Figure 8

Figure 9

Figure 10

(a) Vertical Component

(b) Radial Component

(c) Transverse Component

(d)

Figure 11

Figure 12

Figure 13

(a) Sholte waves (radial) at 2.0 Hz

(b) Scholte waves (radial) at 2.9 Hz

(c) Love waves at 2.0 Hz

(d) Love waves at 2.9 Hz

Figure A