Earth, Planets and Space

FULL PAPER

Open Access

Estimation of phase velocity using array observation of microtremors with arbitrary shape

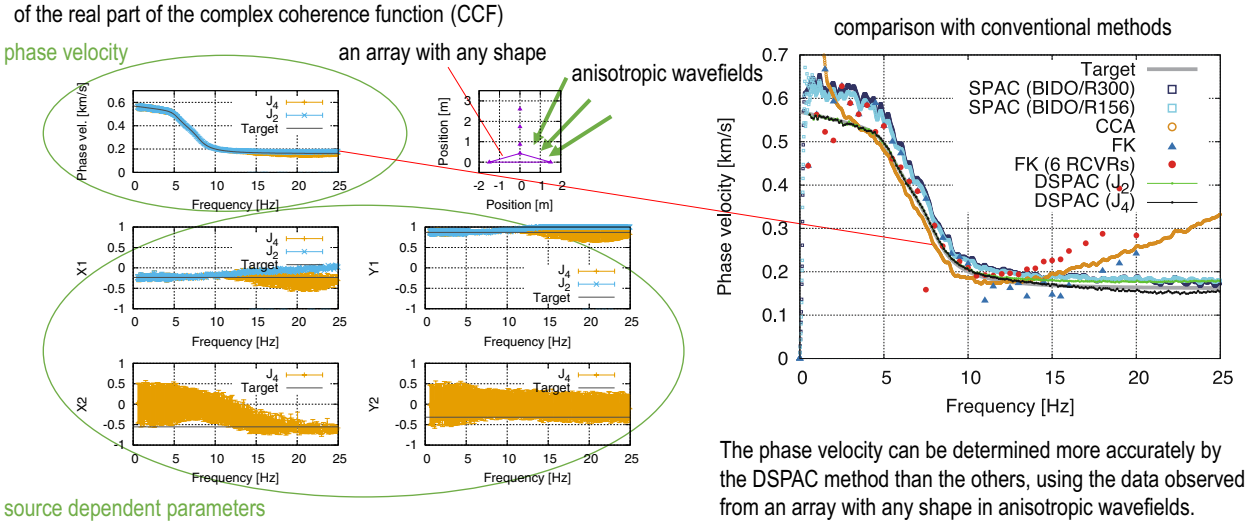
Harusato Kimura[1], Hitoshi Morikawa[1*], Haruka Tomobe[1] and Kahori Iiyama[2]

Abstract To estimate the phase velocity using the array observations of microtremors, some algorithms for the estimation include constraints on the array shape, such as equilateral triangles or the placement of receivers on a circle, in order to reduce the estimation error of the phase velocity. In the present study, a direct estimation technique is introduced for the phase velocity using records obtained through an array with an arbitrary shape based on a complex coherency function (CCF), where CCF is defined as the normalized cross spectrum of the microtremor records observed simultaneously by two receivers. The particle swarm optimization (PSO) method, one of metaheuristic optimization methods, is applied and optimal values are provided for the phase velocity and other unknown parameters. Approximate representations of the stochastic properties for the unknown variables are analytically derived based on the discrete representation of the CCF, for a case where the arrival directions of microtremors are treated as random variables following a uniform distribution. Furthermore, the validity of the proposed method is confirmed using numerical simulations and actual observation records.

Keywords Microtremor survey, Arbitrary array shape, Array observation, Phase velocity, Numerical simulation, Complex coherency function (CCF), Direct spatial auto-correlation (DSPAC) method, Spatial auto-correlation (SPAC), Particle swarm optimization (PSO)

*Correspondence:
Hitoshi Morikawa
morikawa.h.aa@m.titech.ac.jp
Full list of author information is available at the end of the article

Graphical Abstract



Introduction

Methods for estimating the velocity structures of grounds, which explain the phase velocity obtained through array observations of microtremors, are often applied to evaluate earthquake ground motions and to assess the damage caused by earthquakes. Many studies have estimated the velocity structures of grounds using microtremor array observations and have found relationships among the velocity structures, earthquake ground motions, and anomalies of the damage distributions (e.g., Komazawa et al. 2002; Goto et al. 2018). This means that the array observations of microtremors play an important role in the estimation of ground structures. In fact, ever since methods to analyze array observation records were proposed around 1960, these methods have been applied to practical issues. Parallel to this, the theoretical background has been clarified and new techniques for observations and analyses have been developed.

Methods to estimate the phase velocity can be divided into two main approaches: methods based on the frequency–wavenumber (FK) spectrum (Lacoss et al. 1969; Capon 1969; Horike 1985, hereafter called "FK methods") and methods based on the complex coherency function (CCF), which is the cross spectrum of waveforms obtained at two different sites normalized by the power spectra at each site (Aki 1957; Okada 2003, hereafter called "CCF methods").

Receivers can be freely set for FK methods and correct values for the phase velocity can be obtained unless the array shape is close to being a linear distribution of receivers. Since the purpose of FK methods is to obtain the three dimensional Fourier spectrum of a time-space wave field, they can provide the arrival directions of waves directly from the wavenumber vector in addition to the phase velocity.

To obtain the velocity structure of soft sediments, it is necessary to estimate the phase velocity of the surface waves with short wavelengths; the receivers should be situated densely so as to prevent the spatial aliasing of the FK spectrum. Therefore, a large number of receivers must be employed and this requires a great amount of resources for the observations. Even if a large number of receivers is installed for the observations, it is noted that the spatial resolution of the records will generally be much lower than the temporal resolution; thus, the results of the analysis may be unstable.

On the other hand, CCF methods, typified by the spatial auto-correlation (SPAC) method, can estimate the phase velocity over a wider frequency range with a relatively smaller number of sensors than FK methods, and they are thought to be more efficient in terms of the observations (Okada 2003, 2006). In principle, however, the sensors must be evenly distributed around the same or concentric circles. This is because, for example, the SPAC method is based on the azimuthal integral of the

CCF, and the weights for the summation are expected to be distributed equally in cases where the integral is discretized at the positions of the receivers. Of course, by evaluating the weights appropriately, it is possible to estimate the correct phase velocity even with an unequal arrangement of receiver positions in the azimuthal direction.

The SPAC method is based on an algorithm that cancels out many terms caused by bias in the arrival direction of microtremors by the azimuthal average (Shiraishi and Matsuoka 2005), and it is worth noting that it is a very robust algorithm for estimating the phase velocity despite its simple computation. Hereafter, the terms caused by bias in the arrival direction are referred to as "wavefield anisotropic terms". To be more exact, not all of the "wavefield anisotropic terms" are canceled out, but the remaining terms are negligibly small if the argument of the Bessel function of the first kind is sufficiently small, and this property can be used to accurately estimate the phase velocity (Shiraishi and Matsuoka 2005). In addition, Yokoi and Margaryan (2008) have shown that this approximation can be expressed using the theory of seismic interferometry.

In cases where the microtremors arrive isotropically, all of the "wavefield anisotropic terms" become zero (Henstridge 1979; Cho et al. 2006, also discussed in the following section of this article), so the SPAC method can be used to estimate the phase velocity without error. Not only can the SPAC method be used to estimate the phase velocity without error, but not even the azimuthal integral is necessary.

Although the CCF methods, represented by the SPAC method, are very useful techniques for estimating the phase velocity using microtremor records from array observations, they have a restriction. The important point is that the SPAC method works either if the wavefield is isotropic or the number of receivers in the array gives sufficient azimuthal averaging of the CCFs. However, it is difficult to know whether the wavefield is isotropic or not before the observations, which requires the use of equilateral triangular arrays. For pragmatic reality, this restriction are lifted (Morikawa et al. 2004; Chávez-García et al. 2005, 2006) without enough confirmation.

CCF methods were generalized by Cho et al. (2006) and the SPAC method is positioned as one special case among the CCF methods. They also proposed the centerless circular array (CCA) method (Cho et al. 2004) as another special case.

In the conventional CCF methods, the "wavefield anisotropic terms" have been treated as objects of evaluation of the estimation error of the phase velocity. The estimation error has been discussed quantitatively by, for example, Cho et al. (2006) and Asten (2006). Konno (2000) and Shiraishi and Matsuoka (2005) showed explicitly, using the discrete series expansion of CCFs, that many of the "wavefield anisotropic terms" disappeared through azimuthal averaging, which is a discretized azimuthal integration. However, they also showed that some terms do not disappear, but remain and cause an estimation error in phase velocities in the case of anisotropic wave fields.

Based on the above discussion, it is important to ensure the accuracy of the estimation of the phase velocities by eliminating the "wavefield anisotropic terms" using an azimuthal integration in the case of anisotropic wave fields. In other words, CCF methods are inextricably related to the constraints of the array shape and the error brought about by the "wavefield anisotropic terms". The constraints of the array shape comprise a major obstacle in observation practices, and the relaxing of these constraints is an important engineering issue.

The CCA method proposed by Cho et al. (2006) can estimate the phase velocities using only receivers placed on the circumference, so that even if the array shape is not an equilateral triangle, but a triangle of arbitrary shape, the phase velocities can be estimated by considering the three receivers as an array constituting a circle. This means that their CCA method succeeds in avoiding part of the first disadvantage of many CCF methods, which are the constraints of the array shape. However, in the case of four or more receivers, they must all be on the circumference of the same circle or concentric circles, so arrays with arbitrary shapes are not allowed.

As a method that can use for asymmetric arrays, krSPAC is also known (Asten and Hayashi 2018; Asten et al. 2019). This method wisely converts data from the frequency domain to the dimensionless kr domain, which enables to use different separations of receivers together. This method is applied by many researchers, and some successful results are presented for field observations (Asten et al. 2019; Stephenson et al. 2019; Hayashi et al. 2022; Asten et al. 2023). However, the mathematical background of azimuthal averaging is not sufficiently discussed.

The only way to estimate the phase velocity using an array with an arbitrary shape is to directly determine the unknown parameters, including the phase velocity, in the real part of the CCFs which are obtained through each pair of receivers of the array. The unknown parameters can be determined using nonlinear optimization techniques to satisfy the CCFs obtained from the observations. Shiraishi et al. (2005) and Shiraishi et al. (2006) proposed a direct estimation method (DEM), which determines all the unknown parameters of the truncated series-expansion of the CCFs in Shiraishi and Matsuoka (2005) to satisfy the observed values. In this method, the CCFs are averaged over a combination of records at some

specific azimuths to cancel out the lower-order terms that depend strongly on the azimuth, and the unknown parameters are reduced to solve the problem. Although this is a sophisticated idea, it is obviously not effective for increasing the efficiency of the observations, as it entails more restrictions on the positions of the receivers.

Shiraishi and Asanuma (2009) proposed a method to estimate all the unknown parameters literally and directly without azimuthal averaging using records taken from an arbitrary array. The problem to be solved was a highly nonlinear optimization problem involving Bessel functions of the first kind. As will be described later, the unknown parameter included in the nonlinear function is only the phase velocity; the other unknowns form linear equations. According to the method, Shiraishi and Asanuma (2009) firstly give a specific value to the phase velocity, and the optimal solution to the linear equation part is obtained by the least squares method. Then, the value of the phase velocity is varied and the corresponding optimal solutions for the other unknown parameters are determined. Finally, a search is made for a combination of values for the unknowns in order to minimize the difference between the observed and the analytical values of the real part of the CCF. Their algorithm adopts a combination of a grid search for the phase velocity and the least squares method for the other unknowns. In this way, the computational cost is significantly reduced.

In the case of a lack of computer power, this kind of contrivance is important and useful. The number of unknown parameters is limited by the number of receivers because the algorithm proposed by Shiraishi and Asanuma (2009) requires a larger number of observation equations than unknowns for the least squares method. Furthermore, since the least squares method is an algorithm for obtaining solutions deterministically, it has no index to objectively evaluate the adequacy of the solutions. It has another problem in that it cannot determine whether the setting of the unknown parameters and the obtained solutions are appropriate or not.

The metaheuristic optimization method (particle swarm optimization (PSO), proposed by Kennedy and Eberhart (1995)), is presented here to determine the unknowns in order to remove all the constraints on the array shape for the estimation of the phase velocity. PSO can be used to evaluate the reliability of the solutions using multiple sets of random numbers to search for optimal solutions. The computational cost is large and increases rapidly with the number of receivers, which makes it difficult to solve the problem for arrays with many receivers. In principle, however, PSO can be applied to arrays of any shape without error. As for the metaheuristic methods, simulated annealing (SA) and genetic algorithm (GA) are available in addition to the PSO, but these methods require appropriate tuning depending on the target problem (Burke and Kendall 2013). Since the characteristics of the problem targeted in the proposed method varies depending on the array geometry used, the velocity structure of the ground under study, and the degree of noise, the PSO, which does not require detailed parameter tuning, was selected in this paper.

The method proposed in this paper is novel in that the CCF is not solved by reducing the unknowns by azimuthal averaging, but explicitly identify the unknowns using the metaheuristic optimization method. Furthermore, it attempts to clarify the mathematical properties of the unknown parameters under the assumption of stochastic properties for the arrival directions of microtremors. For the sake of simplicity, our proposed method is referred to as the Direct SPAC (DSPAC) method.

Formulation of direct SPAC (DSPAC) method

Definitions

Firstly, some technical terms are defined. Letting t and $\kappa_j(t)$ $(j = 0, 1, \ldots)$ be time and any stochastic process, respectively, and ω and $K_j(\omega)$ be the circular frequency and Fourier transform of $\kappa_j(t)$, the power spectra, $S_{jj}(\omega)$ $(j = 0, 1, \ldots)$, of $\kappa_j(t)$ and the cross spectra, $S_{0j}(\omega)$ $(j = 1, 2, \ldots)$, between $\kappa_0(t)$ and $\kappa_j(t)$ are represented as:

$$\begin{aligned} S_{jj}(\omega) &= E[K_j^*(\omega)K_j(\omega)] \ (j = 0, 1, 2, \ldots), \\ S_{0j}(\omega) &= E[K_0^*(\omega)K_j(\omega)] \ (j = 1, 2, \ldots), \end{aligned} \tag{1}$$

where $E[\,\cdot\,]$ and * denote the expectation and complex conjugates, respectively. The complex coherency function (CCF), $\gamma_{0j}(\omega)$, can be defined using the power and cross spectra as:

$$\gamma_{0j}(\omega) = \frac{S_{0j}(\omega)}{\sqrt{S_{00}(\omega)S_{jj}(\omega)}}. \tag{2}$$

It is considered that $\kappa_0(t)$ and $\kappa_j(t)$ $(j = 1, 2, \ldots, N)$ are waveforms of the microtremors observed at the center of the array and on the circumference, respectively, for a circular array with radius r. Then,

$$\eta(\omega) = \frac{1}{N}\sum_{j=1}^{N}\gamma_{0j}(\omega) \tag{3}$$

is called the SPAC spectrum (Asten 2006; Cho 2020a), where $\eta(\omega) \in \mathbb{C}$, and especially, CCF and SPAC spectrum are identical as $N = 1$. The real part of the SPAC spectrum is also called the SPAC coefficient, $\rho(\omega)$, as follows:

$$\rho(\omega) = \Re[\eta(\omega)], \tag{4}$$

where $\Re[\,\cdot\,]$ stands for the real part of a complex number.

The SPAC method determines the phase velocity using the property whereby the SPAC coefficient is represented

by a zeroth order Bessel function of the first kind. The optimal value of the phase velocity, $\hat{c}(\omega)$, of $c(\omega)$ is identified to satisfy an observed realization, $\tilde{\rho}(\omega)$, of $\rho(\omega)$, using the following relationship:

$$\rho(\omega) \approx J_0(kr) = J_0\left(\frac{r\omega}{c(\omega)}\right), \tag{5}$$

where k is the wavenumber. In cases where the value of kr is less than π and N is set to 3, approximation errors of Eq. (5) induced by the directional aliasing are sufficiently small (Shiraishi and Matsuoka 2005; Shiraishi et al. 2006).

Basic formulation

Shiraishi and Matsuoka (2005) used the far-field analytical solution for the vertical component of the Rayleigh wave for Lamb's problem to discretely represent the CCF of the vertical waveform of the microtremors observed by two receivers, p and q, separated by r_{pq}, as given in Fig. 1a, and showed that the real part is represented by the series expansion, as follows:

$$\Re[\gamma_{pq}(\omega)] = J_0(kr_{pq}) + 2\sum_{n=1}^{\infty}\left\{(-1)^n J_{2n}(kr_{pq})X_n\right\}, \tag{6}$$

where $\gamma_{pq}(\omega)$ is obtained by substituting p and q into 0 and j of Equation (2), and X_n is defined by Eq. (8).

As shown in Fig. 1(b), if receiver s is added at a distance of r_{ps} from line pq in the direction of ψ, counterclockwise from the center of p, the real part of the CCF for line ps is represented as follows (Shiraishi et al. 2005, 2006; Shiraishi and Asanuma 2009):

$$\begin{aligned}\Re[\gamma_{ps}(\omega)] =& J_0(kr_{ps}) + 2\sum_{n=1}^{\infty}\left\{(-1)^n J_{2n}(kr_{ps})\right.\\ &\left.(X_n\cos 2n\psi + Y_n \sin 2n\psi)\right\},\end{aligned} \tag{7}$$

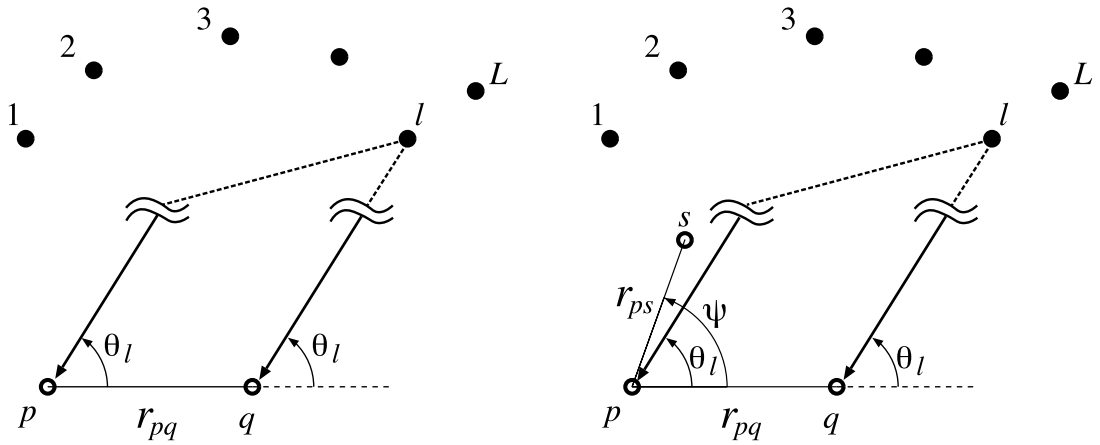
where

$$\begin{aligned}X_n \equiv E[\xi_n] &= E\left[\sum_{\ell=1}^{L}\alpha_\ell \cos 2n\theta_\ell\right]\\ Y_n \equiv E[\zeta_n] &= E\left[\sum_{\ell=1}^{L}\alpha_\ell \sin 2n\theta_\ell\right].\end{aligned} \tag{8}$$

L and α_ℓ denote the total number of microtremor sources and the contribution of amplitude of the ℓth source, respectively, and $\sum_{\ell=1}^{L}\alpha_\ell = 1$. X_n and Y_n relate to the arrival direction θ_ℓ and its amplitude α_ℓ, as shown in Equation (8). However, X_n and Y_n can not represent directly a kind of simple physical meanings, because they are defined as the sum of θ_ℓ and α_ℓ. The physical meanings of X_n and Y_n should be important to understand the properties of CCF, though this is not easy problem to answer. Thus, we leave this problem for our future researches, we treat X_n and Y_n as just parameters to consist of the CCF, and discuss their mathematical properties.

Equation (7) is obtained by replacing r_{pq} by r_{ps} and θ_ℓ by $\theta_\ell - \psi$ in Eq. (6). Equation (7) uses the assumption that the microtremor sources are sufficiently far from the receivers, and that the azimuths to the ℓth source from any of the receivers, p, q, and s, are equal and can be represented as θ_ℓ.

Equation (7) is a general representation of the real part of the CCF for a line of arbitrary azimuth with



(a) Basic setting of receivers and microtremor sources. (b) Setting including an additional receiver.

Fig. 1 Settings of problems and parameters. The open circles, p, q, and s, are receivers. The closed circles, 1, 2, $\cdots$, L, are microtremor sources, where L is the total number of sources. r_{pq} and r_{ps} are the distances between two receivers of p–q, and p–s, respectively. The sources are far enough from the receivers as shown in wiggly lines, and the lines from the receivers to a source are parallel with an azimuth, θ_ℓ. ψ is the angle formed by spq

an arbitrary distance between two receivers. Adding receiver s, line sq can also be used for the analysis. The angle between sqp is set as ψ', and the real part of the CCF, $\Re[\gamma_{sq}(\omega)]$, for line sq can be obtained by replacing ψ with $-\psi'$ in Equation (7).

Comparing Bessel functions $J_6(kr)$ with $J_4(kr)$, the value of $J_6(kr)$ is about 1/10 of $J_4(kr)$ at $kr = 3.2$. In other words, even if the series expansion, with respect to n of Eq. (7), is terminated at $n = 2$ in the range of $kr < 3.2$, the error will be 10% at most. In the following discussion, for the purpose of convenience, the phase velocity is estimated in the range of $kr_{\max} < \pi$, and the series is terminated at $n = 2$, where $r_{\max}$ is the maximum value among the distances of any pairs of receivers constituting an array. The unknowns are five parameters, such as k, X_n, and Y_n $(n = 1, 2)$. X_n and Y_n are variables that depend only on the unknown microtremor sources and not on the number of receivers. m is the number of receivers and $m(m-1)/2$ is the number of lines for the measurements. Thus, if there are three receivers $(m = 3)$, there are five unknowns and three equations, and if there are four receivers $(m = 4)$, there are still five unknowns, but six equations. It is seen that the number of constraints increases and the determinacy of the unknowns increases as the number of receivers increases.

The realization of the left-hand side of Eq. (7), $\tilde{\gamma}_{pq}(\omega)$, can be obtained from observations; thus, the problem of searching for the five unknowns that satisfy these equations can be solved. However, this involves solving an optimization problem with strong nonlinearity.

Shiraishi and Asanuma (2009) solved this problem by applying the least squares method for determining X_n, and Y_n $(n = 1, 2, \ldots, \nu)$ and the grid search technique for the phase velocity. The maximum value for ν is determined automatically by providing the number of receivers m. Although there is no way to validate the value of ν, the computational cost does not increase greatly even though m is large and the solution is uniquely determined.

On the other hand, a nonlinear optimization problem with five unknowns can be determined simultaneously with our method. For this purpose, the particle swarm optimization (PSO) method (Kennedy and Eberhart 1995), one of the metaheuristic optimization techniques, is applied. For the direct optimization, it is expected that the optimal solution can be obtained for arrays with any number of receivers m and arbitrary shapes. However, the following points should be noted: the convergence of the solution depends on the control parameters of the PSO, the optimal solution depends on the initial value and may fall into the local solution, and the computational cost is very high as m is large.

The first problem can only be solved by fine tuning the parameters through trial and error. The second problem can be confirmed with different initial value sets (initial positions of particles for the PSO) to check the reproducibility of the solution. Although this would further increase the computational cost, it is probably possible to solve this problem together with the third problem by using the recently developed high-performance computers.

Stochastic properties of ξ_n and ζ_n ($n \in \mathbb{N}$) in isotropic wave fields

It is assumed that the microtremor sources are both temporally and spatially piecewise stationary and that θ_ℓ is an independent stochastic parameter with respect to ℓ and follows a uniform distribution at the interval of $[0, 2\pi)$. In this case, $\cos(2n\theta_\ell)$ and $\sin(2n\theta_\ell)$ $(n \in \mathbb{N})$ follow the arc sine distribution and their expectations and variances are determined as follows, respectively:

$$\begin{aligned} \mathrm{E}[\cos(2\mathrm{n}\theta_\ell)] &= \mathrm{E}[\sin(2\mathrm{n}\theta_\ell)] = 0 \\ \mathrm{Var}[\cos(2\mathrm{n}\theta_\ell)] &= \mathrm{Var}[\sin(2\mathrm{n}\theta_\ell)] = \frac{1}{2}. \end{aligned} \tag{9}$$

The magnitude, $\hat{\alpha}_\ell$, of the ℓth source is assumed to be a stochastic variable that follows a uniform distribution and is independent in terms of ℓ. Then, the contribution of amplitude from the ℓth source, α_ℓ, can be represented by $\alpha_\ell = \frac{\hat{\alpha}_\ell}{\sum_{\ell=1}^{L} \hat{\alpha}_\ell}$. Since it is quite difficult to obtain the probability density function (PDF) of α_ℓ in closed form, its histogram is numerically calculated for various values of L using the Monte Carlo method, under the assumption that $\hat{\alpha}_\ell$ follows a uniform distribution in the interval [0,1]. α_ℓ is independent of ℓ; thus, its histogram does not depend on ℓ. The histogram of α_ℓ obtained after 100 thousand trials, which is normalized by the value of L, is shown in Fig. 2.

From Fig. 2, as L is large enough, the stochastic properties for α_ℓ are approximated by a uniform distribution and its PDF, $f_{\alpha_\ell}(\alpha_\ell)$, is represented by

$$f_{\alpha_\ell}(\alpha_\ell) \approx \begin{cases} \dfrac{L}{2} & \left(0 \le \alpha_\ell \le \dfrac{2}{L}\right) \\ 0 & (\text{otherwise}) \end{cases}. \tag{10}$$

Under this approximation, the expected value and variance of α_ℓ are approximately expressed as follows:

$$\begin{aligned} \mathrm{E}[\alpha_\ell] &\approx \frac{1}{L} \\ \mathrm{Var}[\alpha_\ell] &\approx \frac{1}{3L^2}. \end{aligned} \tag{11}$$

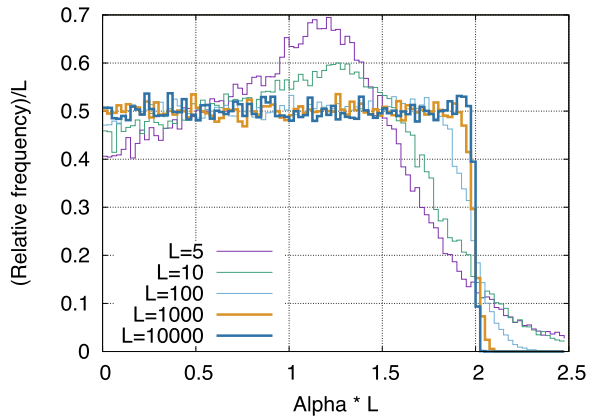


Fig. 2 Histogram of α_ℓ for various values of total number of microtremor sources, L. The histograms seem to converge to a uniform distribution as L increases

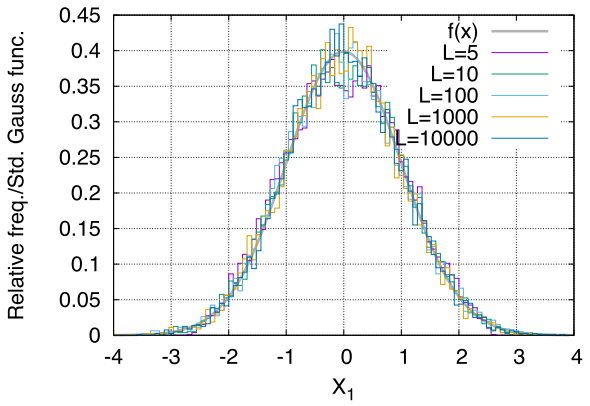


Fig. 3 Histogram of ξ_1 for various values of total number of microtremor sources, L. The histograms are converted to the standard Gaussian distribution using the analytical values of expectations and variances obtained through Eq. (13). The histograms agree with the standard Gaussian distribution drawn by the thick gray line. Notations X_1 and $f(x)$ mean ξ_1 and the standard Gaussian function, respectively

From the above, assuming that α_ℓ and θ_ℓ are mutually independent, the expected value and variance of $\alpha_\ell \cos(2n\theta_\ell)$ yield

$$\begin{aligned} \mathrm{E}[\alpha_\ell \cos(2\mathrm{n}\theta_\ell)] &= \mathrm{E}[\alpha_\ell]\mathrm{E}[\cos(2\mathrm{n}\theta_\ell)]] = 0 \\ \mathrm{Var}[\alpha_\ell \cos(2\mathrm{n}\theta_\ell)] &= \mathrm{E}[\{\alpha_\ell \cos(2\mathrm{n}\theta_\ell)\}^2] - \{\mathrm{E}[\alpha_\ell \cos(2\mathrm{n}\theta_\ell)]\}^2 \approx \frac{2}{3}\frac{1}{\mathrm{L}^2}. \end{aligned} \tag{12}$$

The same results are obtained for $\alpha_\ell \sin(2n\theta_\ell)$.

For $\xi_n = \sum_{\ell=1}^{L} \alpha_\ell \cos(2n\theta_\ell)$ of Eq. (8), the terms on the right-hand side are independent of each other with respect to ℓ; thus, ξ_n is Gaussian as $L \to \infty$ because of the central limit theorem. The expectation and variance of ξ_n are obtained as:

$$\begin{aligned} X_n = \mathrm{E}[\xi_\mathrm{n}] &= E\left[\sum_{\ell=1}^{L} \alpha_\ell \cos(2n\theta_\ell)\right] = 0, \\ \mathrm{Var}[\xi_\mathrm{n}] &= \mathrm{Var}\left[\sum_{\ell=1}^{\mathrm{L}} \alpha_\ell \cos(2\mathrm{n}\theta_\ell)\right] \approx \frac{2}{3}\frac{1}{\mathrm{L}}. \end{aligned} \tag{13}$$

From Eq. (13), it is observed that ξ_n converges to 0 because $Var[\xi_n] \to 0$ as $L \to \infty$. The same results can be obtained for ζ_n.

This indicates that the terms of Eq. (7) become zero except for the first term, $J_0(kr)$, if the microtremor field is isotropic and the number of sources are sufficiently large. For isotropic wavefields, this means that the azimuthal average in the SPAC method is not required and only the real part of the CCF (i.e., $\Re[\gamma_{pq}(\omega)]$) for the records obtained from two receivers (e.g., p and q in Fig. 1) can be applied to estimate the phase velocity. This is consistent with the previously known properties (e.g., Cho et al. 2006).

Figure 3 compares a histogram for ξ_1 obtained by the Monte Carlo method, for various values of L, with the theoretical Gaussian distribution. In this figure, the realized values for ξ_1 are converted to the standard Gaussian distribution using the approximations of expectations and variances for ξ_1 shown in Eq. (13), and are compared with the standard Gaussian distribution function. Table 1 shows the expectations and standard deviations of ξ_1 obtained through the Monte Carlo method and their approximations obtained through Eq. (13). From Fig. 3 and Table 1, it is observed that the stochastic properties of ξ_1 can be approximated by Eq. (13), although the convergence is not as fast as $L \to \infty$.

Table 1 Expectations and standard deviations of ξ_1 obtained through Monte Carlo simulation and Equation (13) for various values of L

	$\mathrm{E}[\xi_1]$		$\sqrt{\mathrm{Var}[\xi_1]}$	
L	Monte Carlo	Eq. (13)	Monte Carlo	Eq. (13)
1	−0.0001	0	0.7014	0.8165
3	0.0037	0	0.4641	0.4714
5	−0.0129	0	0.3640	0.3651
10	−0.0002	0	0.2590	0.2582
100	−0.0010	0	0.0832	0.0816
1000	−0.0002	0	0.0257	0.0258
10000	0.00001	0	0.0081	0.0082

Stochastic properties of ξ_n and ζ_n ($n \in \mathbb{N}$) in anisotropic wave fields

As shown in the previous section, parameters X_n and Y_n ($n \in \mathbb{N}$), defined by Eq. (8), are zero in the cases of the isotropic wave fields of microtremors. In cases where there is a bias for the arrival directions of microtremors, the values of the parameters are determined by the arrival direction, θ_ℓ. The stochastic properties of ξ_n and ζ_n of Eq. (8) will be discussed for the case of anisotropic wavefields under simple conditions.

For the analytical discussion in this section, it is assumed that the total number of microtremor sources is L and the arrival directions of microtremors, θ_ℓ, is limited in the range of $[\phi_0, \phi_0 + \Delta\phi]$ and θ_ℓ follows a uniform distribution in the range. The PDF of θ_ℓ has a constant value for $1/\Delta\phi$ in the same range and zero otherwise. Letting $\phi_1 \equiv \phi_0 + \Delta\phi$ and recalling that α_ℓ and θ_ℓ are independent of each other, the PDF of θ_ℓ is a constant value, $1/\Delta\phi$, in the range of $[\phi_0, \phi_1]$, and the PDF of α_ℓ is approximated by Eq. (10). The expectation of $\alpha_\ell \cos(2n\theta_\ell)$ can be represented as follows:

$$\mathrm{E}[\alpha_\ell \cos(2\mathrm{n}\theta_\ell)] \approx \frac{1}{\mathrm{L}} \frac{\sin(2\mathrm{n}\phi_1) - \sin(2\mathrm{n}\phi_0)}{2\mathrm{n}\Delta\phi}. \tag{14}$$

In cases where the arrival directions of microtremors are limited to the range of $[\phi_0, \phi_0 + \Delta\phi]$, parameter ξ_n is rewritten as $\xi_n(\phi_0; \Delta\phi)$ and the approximated representation for its expectation is obtained as

$$X_n = \mathrm{E}[\xi_\mathrm{n}(\phi_0; \Delta\phi)] = \sum_{\ell=1}^{\mathrm{L}} \mathrm{E}[\alpha_\ell \cos(2\mathrm{n}\theta_\ell)] \approx \frac{\sin(2n\phi_1) - \sin(2n\phi_0)}{2n\Delta\phi}. \tag{15}$$

Similarly for Y_n, the following equation is obtained:

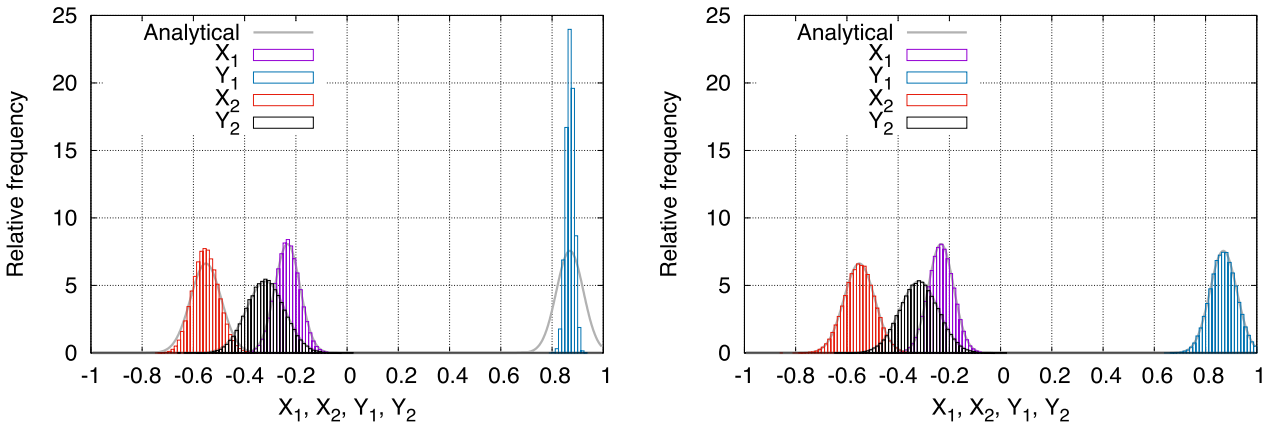
$$Y_n = \mathrm{E}[\zeta_\mathrm{n}(\phi_0; \Delta\phi)] \approx -\frac{\cos(2\mathrm{n}\phi_1) - \cos(2\mathrm{n}\phi_0)}{2\mathrm{n}\Delta\phi}. \tag{16}$$

Under the same conditions as those above, approximations for variances of $\xi_n(\phi_0; \Delta\phi)$ and $\zeta_n(\phi_0; \Delta\phi)$ can be obtained from Appendix A, as follows:

$$Var[\xi_n(\phi_0; \Delta\phi)] \approx \frac{4}{3L}\left\{\frac{1}{2} + \frac{\sin(4n\phi_1) - \sin(4n\phi_0)}{8n\Delta\phi}\right\} - \frac{E[\xi_n(\phi_0; \Delta\phi)]^2}{L} \tag{17}$$

$$Var[\zeta_n(\phi_0; \Delta\phi)] \approx \frac{4}{3L}\left\{\frac{1}{2} - \frac{\sin(4n\phi_1) - \sin(4n\phi_0)}{8n\Delta\phi}\right\} - \frac{E[\zeta_n(\phi_0; \Delta\phi)]^2}{L}. \tag{18}$$

$\xi_n(\phi_0; \Delta\phi)$ and $\zeta_n(\phi_0; \Delta\phi)$ are the summations of independent stochastic variables with respect to ℓ; thus, they are Gaussian variables as L is large enough because of the central limit theorem. The histograms of $\xi_n(\phi_0; \Delta\phi)$ and $\zeta_n(\phi_0; \Delta\phi)$ ($n = 1, 2$) are calculated using 2^{17} samples obtained through the Monte Carlo method for the parameters $L = 100$, $\phi_0 = \pi/6$, and $\Delta\phi = \pi/4$. The histograms are compared with the Gaussian distributions



(a) Histograms based on definition of α_ℓ (b) Histograms based on Equation (10)

Fig. 4 Histograms of ξ_n and ζ_n ($n = 1, 2$) and Gaussian distributions approximated by Eqs. (15) to (18) as drawn by thick gray lines. The parameters are set at $L = 100$, $\phi_0 = \pi/6$, and $\Delta\phi = \pi/4$. Notations X_n and Y_n ($n = 1, 2$) mean $\xi_n(\phi_0; \Delta\phi)$ and $\zeta_n(\phi_0; \Delta\phi)$, respectively

using the expectations and variances obtained from the above approximations in Fig. 4a. This figure shows that the expectations are approximated correctly and that the PDF can be represented by Gaussian distributions. However, the variances of the histograms deviate from the approximation as the absolute values of the expectations increase.

Figure 4b shows the histograms for $\xi_n(\phi_0;\Delta\phi)$ and $\zeta_n(\phi_0;\Delta\phi)$ $(n=1,2)$ which are calculated by applying the realizations obtained from the random numbers following the uniform distribution shown in Eq. (10). From this figure, it is observed that the histograms agree well with the Gaussian functions with variances obtained through Eqs. (17) and (18). This suggests that the worse approximations for variances of $\xi_n(\phi_0;\Delta\phi)$ and $\zeta_n(\phi_0;\Delta\phi)$ are caused by the worse approximation of Eq. (10) as the larger absolute values of the expectations for $\xi_n(\phi_0;\Delta\phi)$ and $\zeta_n(\phi_0;\Delta\phi)$.

Table 2 lists three cases: (i) the expectations and standard deviations obtained analytically using Eqs. (15) to (18), shown in 1st column, (ii) those obtained through the Monte Carlo method, according to the definition of α_ℓ in the 2nd column, and (iii) those obtained through realizations of Eqn. (10) in the 3rd column. Since Eqs. (15) to (18) are based on Eq. (10), the results based on the 1st and 3rd cases are very consistent, but it is seen that the 1st and 3rd cases may not be consistent with the 2nd case, which we want to approximate and know exactly.

From the above discussion, the expectations of $\xi_n(\phi_0;\Delta\phi)$ and $\zeta_n(\phi_0;\Delta\phi)$ $(n=1,2)$, based on Eq. (10), are approximated well, although good approximations for their variances are limited in the range where the absolute values of their expectations are close to zero. Moreover, since the accuracy of the approximations for the variances does not improve as L increases, Eqs. (17) and (18) for the variances should be applied with care.

Table 2 Expectations and standard deviations of $\xi_n(\phi_0;\Delta\phi)$ and $\zeta_n(\phi_0;\Delta\phi)$ $(n=1,2;\phi_0=\pi/6,\Delta\phi=\pi/4,L=100)$

	Eqs.(15) to (18)	Using Monte Carlo	
		$\alpha_\ell \equiv \frac{\hat{\alpha_\ell}}{\sum\hat{\alpha_\ell}}$	Eq.(10)
$E[\xi_1]$	− 0.2328	− 0.2329	− 0.2329
$\sqrt{Var[\xi_1]}$	0.04947	0.04778	0.04958
$E[\zeta_1]$	0.8697	0.8697	0.8697
$\sqrt{Var[\zeta_1]}$	0.05277	0.01612	0.05278
$E[\xi_2]$	− 0.5513	− 0.5515	− 0.5515
$\sqrt{Var[\xi_2]}$	0.06022	0.05126	0.06039
$E[\zeta_2]$	− 0.3183	− 0.3182	− 0.3181
$\sqrt{Var[\zeta_2]}$	0.07519	0.07315	0.07534

Estimation of phase velocity and other unknown parameters

Particle swarm optimization (PSO)

Using the data observed through a microtremor array, the realization of the left-hand side of Eq. (7) can be obtained. For example, three receivers, p, q and s, are set as shown in Fig. 1(b). Optimal solutions are sought for the phase velocity, $c(\omega)=\omega/k$, and other unknowns, X_n and Y_n $(n=1,2,\ldots)$, by satisfying the three equations for $\Re[\tilde{\gamma}_{pq}(\omega)]$, $\Re[\tilde{\gamma}_{qs}(\omega)]$, and $\Re[\tilde{\gamma}_{sp}(\omega)]$. The objective function of the optimization problem is formulated as a problem to minimize the sum of the squares of the differences between the observed values (the left-hand side of Eq. (7)) and the analytical values (the right-hand side of Eq. (7)) for all the possible pairs of receivers. As discussed previously, since the value for $J_6(kr)$ is very close to zero in $0<kr<\pi$, the series expansion on the right-hand side of Eq. (7) is terminated at $n=2$, and then the optimal values for the five unknowns: $c(\omega)$, X_n and Y_n $(n=1,2)$, are found.

Various techniques have been proposed to solve this type of nonlinear optimization problem. If one is not concerned about the computational cost, metaheuristic optimization methods are useful because they are good for avoiding the fall into local solutions and for finding global solutions. The particle swarm optimization (PSO), proposed by Kennedy and Eberhart (1995), is applied to solve the above problem.

The details of PSO are left to Kennedy and Eberhart (1995) and many other literatures (e.g., Clerc 2006), and only the parameter settings necessary for our calculations are described below. PSO can be controlled for its behavior of optimizations using three parameters: inertia control of particles, w, weight of personal best value, C_p, and weight of global best value, C_g. Any values around one are usually applied for C_p and C_g and any value in the interval [0,1] is applied for w, although the most appropriate values depend on the target optimization problem. There is no way to determine the values of these parameters to control PSO. Therefore, several analyses have been conducted and the values have been determined subjectively by trial and error. The results show that the solutions converge well when the values for w and C_g are small and the value for C_p is large. Thus, the values of $w=0.2$, $C_p=1.4$, and $C_g=0.7$ are applied to the following analyses.

The number of particles must also be given. A problem with several different numbers of particles is solved; the less likely it is for a small number of particles to fall into a local solution and the more likely it is for a large number of particles to yield a correct solution. However, the

solution is not improved when the number of particles is above 10,000, and 10,000 particles are used below.

Trials of optimization using PSO technique

To confirm whether PSO can solve Equation (7), a simple array configuration is set up, simulating the observed values for $\Re[\gamma_{uv}]$ (uv is a combination of p, q, and s) by the forward calculation of Equation (7), and performing an optimization to identify the unknowns from the observed values. This trial, namely, can be said as a kind of blind tests. Firstly, giving values to phase velocity $c(f)$, and parameters X_n and Y_n ($n = 1, \ldots, 10$) at a frequency of $f = \omega/(2\pi)$, they are substituted into the right-hand side of Equation (7). The values for X_n and Y_n ($n = 1, \ldots, 10$) are calculated from realizations for $\hat{\alpha}_\ell$ and θ_ℓ on the right-hand side of Equation (8) as uniform random numbers. The obtained values for $\Re[\gamma_{uv}]$ are then considered as observed values and $c(f)$, X_n, and Y_n ($n = 1, 2$) are identified using PSO. The validity of the solutions is examined by comparing the identified values with the values initially given.

As shown in Fig. 5, five isosceles triangular array configurations with R6-R7 are considered as the base and R1 to R5 as the vertices. In particular, "R4-R6-R7" is an equilateral triangle and R2 is its centroid.

The values of the parameters given for the forward problem are arbitrary, but the following values are used for the analysis: frequency $f = 10$ [Hz], phase velocity $c(f) = 0.165$ [km/s], and the number of microtremor sources $L = 1000$. As the most basic problem, isotropic wave fields of microtremors are assumed, the realizations of θ_ℓ and $\hat{\alpha}_\ell$ are generated from the random numbers with uniform distributions in the ranges of $[0, 2\pi]$ and [0,1], respectively, and the contribution of source ℓ, α_ℓ, is calculated from $\alpha_\ell = \hat{\alpha}_\ell / \sum_{\ell=1}^{L} \hat{\alpha}_\ell$. These realizations are substituted into Eq. (8) to calculate X_n and Y_n ($n = 1, \ldots, 10$), and they are substituted into Eq. (7) with $k = 2\pi f / c(f)$ to obtain the realization of the left-hand side, $\Re[\tilde{\gamma}(f)]$.

In a case of the three-receiver array (e.g., R1-R6-R7 in Fig. 5), $\Re[\tilde{\gamma}(f)]$ is considered as the observed value for three pairs (e.g., R1-R6, R6-R7, and R7-R1), and $c(f)$, X_n, and Y_n ($n = 1, 2$), which satisfy Eq. (7), are searched using PSO. The target values of the parameters obtained through the above procedure are shown in the caption given for Fig. 6.

To investigate the dependency of the solution estimated by PSO on the initial values, 1000 data sets with different initial values (sets of initial positions of particles) are created, and the results obtained from the search with each initial value are shown in Fig. 6. The figure shows the results of the blind test using 1000 initial data sets for the PSO. The results are shown by the mean values and the range of standard deviations of the estimation results for each triangular array with a different color. Variables with large error bars indicate a strong dependency of the solution on the initial values.

Figure 6 shows that the solutions are close to the target values regardless of the array shape. The estimated values for the phase velocity are less affected by the initial values

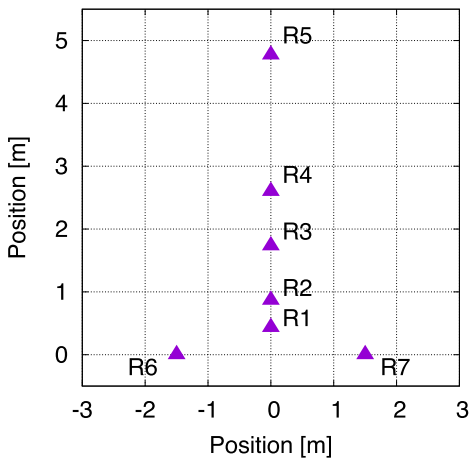


Fig. 5 Collocation of receivers. "R4-R6-R7" is an equilateral triangle and R2 is its centroid

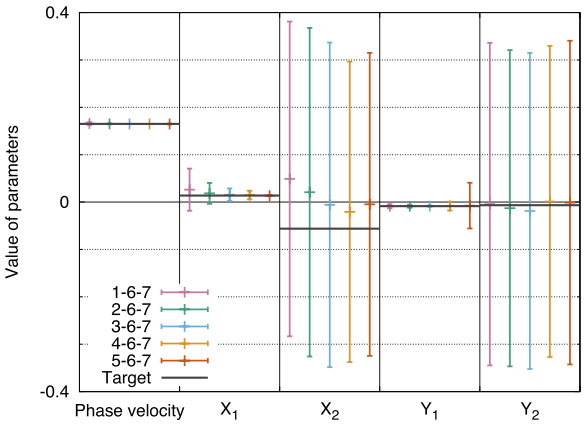


Fig. 6 Estimated values of unknown parameters for various shapes of triangles. The different color corresponds to the different shape of triangles. The estimated value is shown with its error bar, which denotes the mean value and plus-minus one standard deviation. The fluctuations in the estimated value are caused by the initial data sets of PSO. Short error bars show that PSO can identify stably unknown parameters, though large error bars show unstable identifications. The parameters are set at $f = 10$ [Hz], $c(f) = 0.165$ [km/s], $X_1 = 0.01378$, $Y_1 = -0.008617$, $X_2 = -0.05611$, and $Y_2 = 0.006514$. The horizontal gray lines are the target values for each unknown parameter. The three-digit numbers in the legend represent the receiver ID and the array shape (e.g., "4–6–7" is identical to R4-R6-R7 and corresponds to an equilateral triangle)

(small standard deviation), and the results are highly reliable and accurate. The estimated values for X_1 and Y_1 show almost the same properties as the phase velocity, although their standard deviations are slightly larger than those of the phase velocities. On the other hand, for X_2 and Y_2, the mean values of the estimated values are not only slightly different from the target values, but also show very large standard deviations. This indicates that the solutions strongly depend on the initial values, namely, the values for X_2 and Y_2 have little effect on the value for $\Re[\gamma(f)]$ in Equation (7).

Figure 6 shows that the standard deviation of X_1 is larger for flat triangles, while the standard deviation of Y_1 is larger only for the tallest triangle, R5-R6-R7. The more the shape is distorted from an equilateral triangle, the more the solution of X_1 or Y_1 will be dependent on the initial value, that is, the stability of the solution depends on the array shape.

In Fig. 6, the estimated phase velocity, which is the main target of the estimation, seems to be consistent regardless of the array shape. However, for a closer look, Fig. 7 shows a histogram of the phase velocities obtained through 1000 sets of initial values. It can be observed that the equilateral triangle, R4-R6-R7, has the smallest variance and a flattened shape; R1-R6-R7 has the largest variance. In this figure, R3-R6-R7 has the smallest standard deviation after the equilateral triangle, and R2-R6-R7 and R5-R6-R7 are almost equal to each other with even larger variances than R3-R6-R7.

Figures 6 and 7 show that, although it is possible to estimate the phase velocity for any array shapes, the initial value dependency of the estimated phase velocity is smaller for an array shape close to an equilateral triangle and larger for an array shape that is flat. For an array shape close to an equilateral triangle, the exact phase velocity can be obtained no matter what initial value is given to it, so there is almost no need to search for solutions using various initial values. On the other hand, it should be noted that the flatter the triangular shape, the more the phase velocity must be evaluated from the searches using many sets of initial values; otherwise, there is a risk of misinterpreting completely different values as the solution.

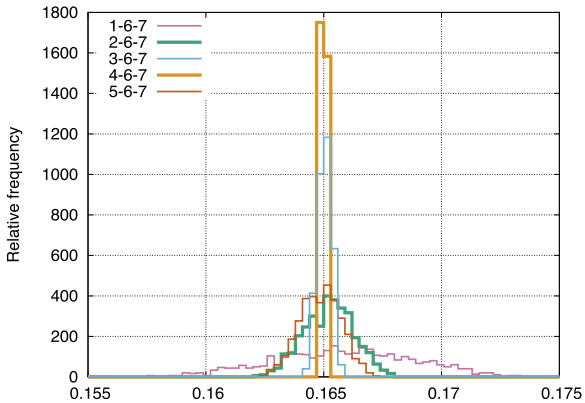
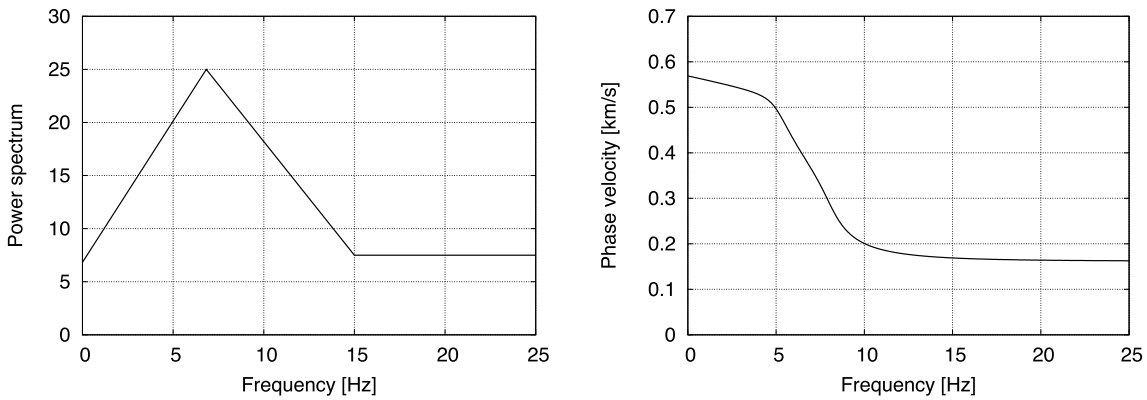


Fig. 7 Histogram of phase velocities determined by PSO for various shapes of triangular arrays. The equilateral triangle (R4-R6-R7) shows the smallest variance and the most-flattened triangle (R1-R6-R7) shows the largest variance. The legend is the same as that in Fig. 6

Numerical examples of DSPAC method

Simulating microtremor wave fields

Microtremors are simulated by realizations of the stochastic process. The data are created with appropriate power spectrum and a uniformly distributed phase angle of $[0, 2\pi]$ in the time domain. They are generated at 60 samples per second for 2^{16} samples. In addition, an



(a) Power spectrum (b) Dispersion curve

Fig. 8 Power spectrum density function and dispersion curve for numerical analyses. The vertical axis of the power spectrum has no physical meaning

appropriate dispersion curve is also given and the differences in phase angles between the different receivers are given using the phase velocities for each frequency to simulate the wave propagation of a plane wave. Propagation directions of the plane waves are considered as random variables, and L realizations of the stochastic process are generated using different sets of random numbers, where L is the total number of microtremor sources. These are superimposed to obtain the observed waveform at a receiver.

Figure 8 shows the power spectrum and dispersion curve used in the following analyses. The vertical axis of the poser spectrum is a number with no particular physical meaning, and the absolute value of the amplitude of the time histories as the realizations has no physical meaning either.

In the following analysis, $L = 100$ and the configuration of the receivers is the same as that seen in Fig. 5. For determining the phase velocity and unknowns X_n and Y_n, using the proposed method, the DSPAC method, the constraints $kr \leq \pi$, i.e., $c(f) \leq 2r_{\max}f$ are introduced as long as $J_6(kr)$ is sufficiently small, and the series of Equation (7) is terminated up to $n = 2$. The unknown parameters are solved under the constraints that $|X_n|, |Y_n| \leq 1$ $(n = 1, 2)$.

As the preliminary study in the previous section showed, the optimal solution of the unknowns by PSO may depend on the initial values. Therefore, 200 sets of initial values (initial positions of particles) are created using random numbers, and the unknowns are identified from each initial data set. The means and standard deviations are also calculated for the 200 sets of the optimal solutions. The results shown in this section, that is, Figs. 9, 10, 11, 12, 13, 14, 15, to 16 are obtained on the basis of the synthetic data generated by the above procedure.

Isotropic wave fields

Case of noise-free data

From the discussion in the previous section, it is expected that the equilateral triangle of R4-R6-R7 provides the best estimation, while the most flattened triangle of R1-R6-R7 is the least accurate. In the following, the discussion will be focused on these two cases.

The identified unknowns are shown in orange of Fig. 9, in which the series expansion is terminated up to $n = 2$ of Eq. (7). Since the microtremors arrive isotropically, X_n and Y_n $(n = 1, 2)$ should be zero as shown in Eq. (13), namely, the means of the identified values are close to zero. The phase velocity is also almost the same as the target dispersion curve, and all the unknowns seem to have been correctly identified.

However, the standard deviations of X_2 and Y_2 are much larger than those of the other parameters. This indicates that the search for the optimal solution depends highly on the initial values and suggests that X_2 and Y_2 hardly affect the results of the real part of the CCF, as described in the previous section.

If X_2 and Y_2 do not contribute to the forward calculation of the real part of the CCF, there is no need to identify these parameters. Thus, the series expansion of Equation (7) is terminated at n=1, reducing the number of unknowns to three, namely, $c(f)$, X_1, and Y_1. The three unknowns identified from the same waveform data are shown in blue in Fig. 9.

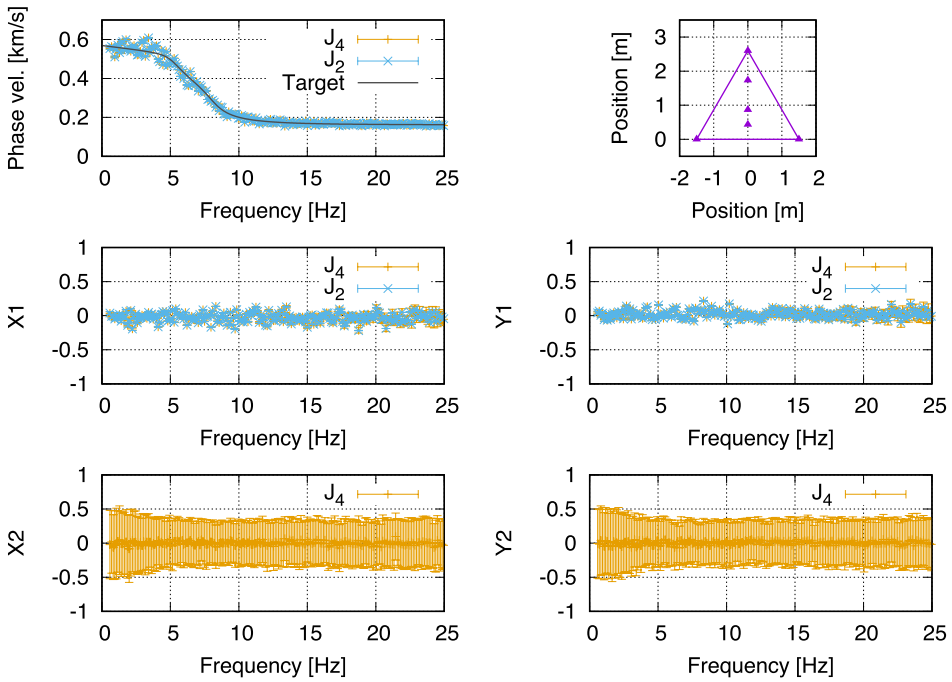
Comparing the orange and blue points in Fig. 9, the results are almost identical for $n = 1$ as for the $n = 2$ series termination. This means that X_2 and Y_2 are hardly sensitive parameters. It should be noted, however, that this is the case for isotropic wave fields of microtremors.

Case of data contaminated by incoherent noise

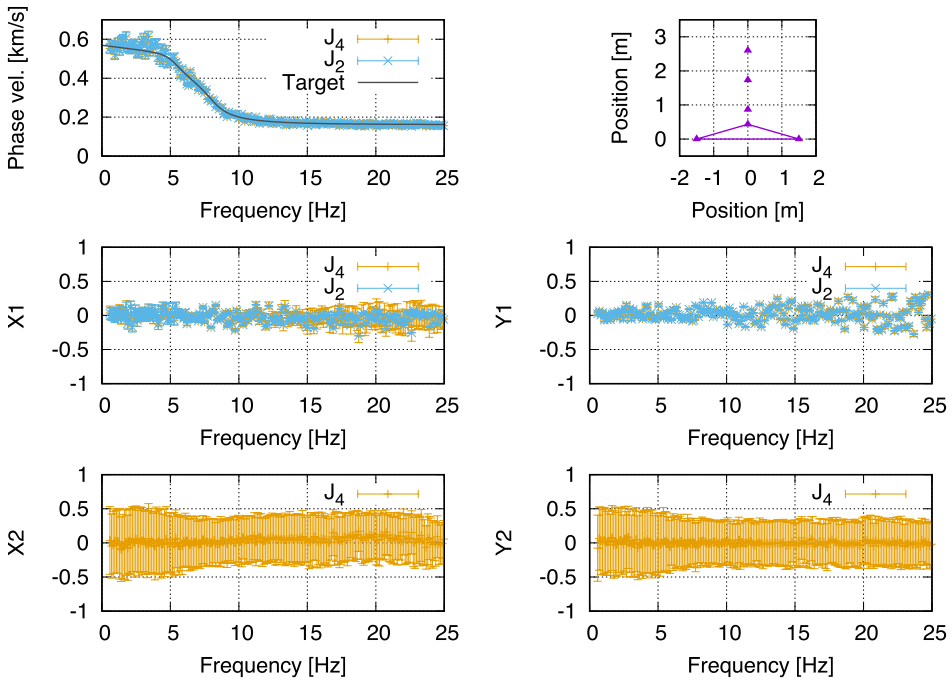
It is known that the SPAC method is affected by incoherent noise, and that the accuracy of the estimation for the phase velocity becomes worse at lower frequencies (Cho 2019; Cho and Iwata 2021). To investigate the effect of incoherent noise, the DSPAC method is applied to data to which mutually independent white noise is appended to the waveform at each receiver.

Adding white noise to a signal, which is the waveform of microtremors, incoherent noise is simulated, where the one side amplitude of white noise is set as β% of the RMS of the signal. That is, the RMS of the signal of microtremor s_{rms}, and the white noise generated from the uniformly distributed random number in $[-\beta s_{\mathrm{rms}}/100, \beta s_{\mathrm{rms}}/100]$ is added to the signal. The RMS ratio of noise to signal is $\beta/(100\sqrt{3})$ based on the standard deviation of the uniform distribution. For example, in the case of $\beta = 10\%$, the RMS ratio of noise to microtremor is about 0.058. For the case of $\beta = 10\%$, the unknowns identified for the equilateral triangle (R4-R6-R7) and the flattened triangle (R1-R6-R7) are shown in Fig. 10. Note that the signal-to-noise ratio (SNR) at a particular frequency depends on the shape of the power spectrum density.

Since the DSPAC method is based on CCF, the effect of noise is similar to the results shown by Cho and Iwata (2021), resulting in significantly poorer identification accuracy of the unknowns in the low-frequency range. The flattened triangle (R1-R6-R7) is more affected by noise, probably because it contains shorter pairs of observation points than the equilateral triangle (R4-R6-R7), which is more affected by noise in the low-frequency range.

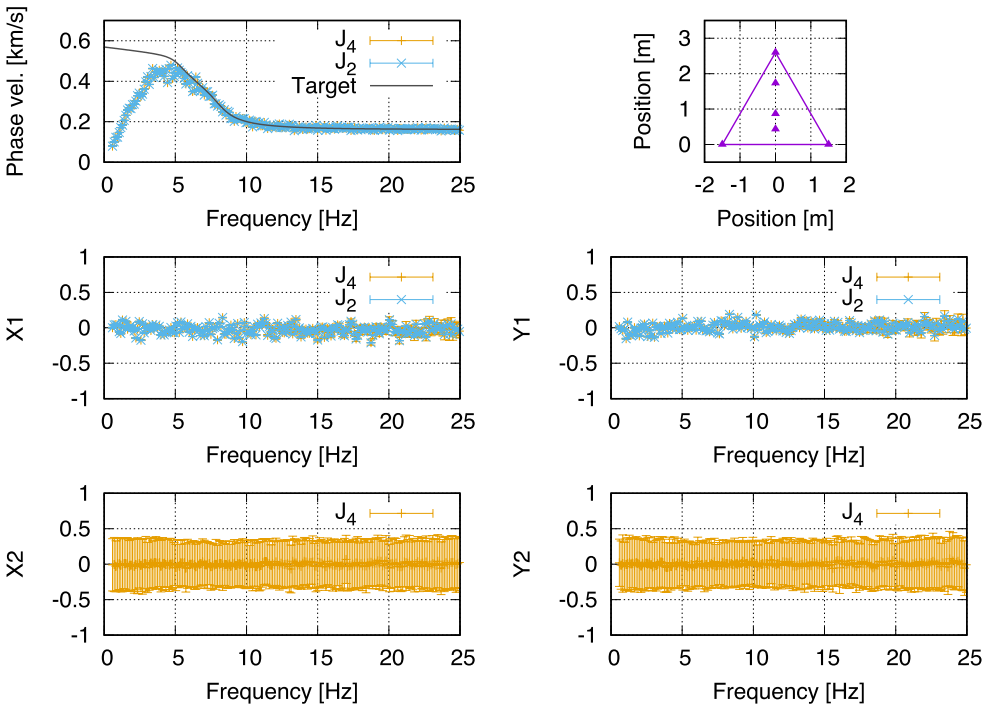


(a) Equilateral triangle (R4-R6-R7)

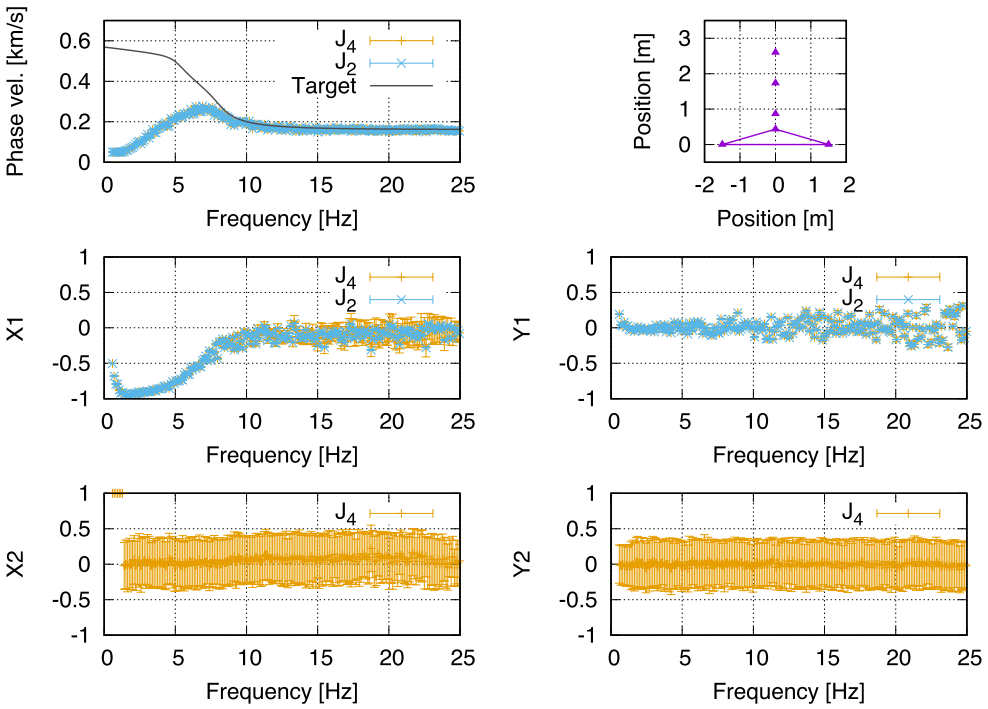


(b) Most-flattened triangle (R1-R6-R7)

Fig. 9 Estimated values of unknown parameters for equilateral and most-flattened triangles (synthetic data). The top left panel shows the phase velocity and the other panels show the other unknown parameters, such as X_n and Y_n ($n = 1, 2$). The solid gray line is the target dispersion curve. The orange points show the results which are used in the series expansions of Eq. (7) until $n = 2$, and the blue points show those until $n = 1$. The points show the averaged values with error bars of the range of $\pm$ one standard deviation

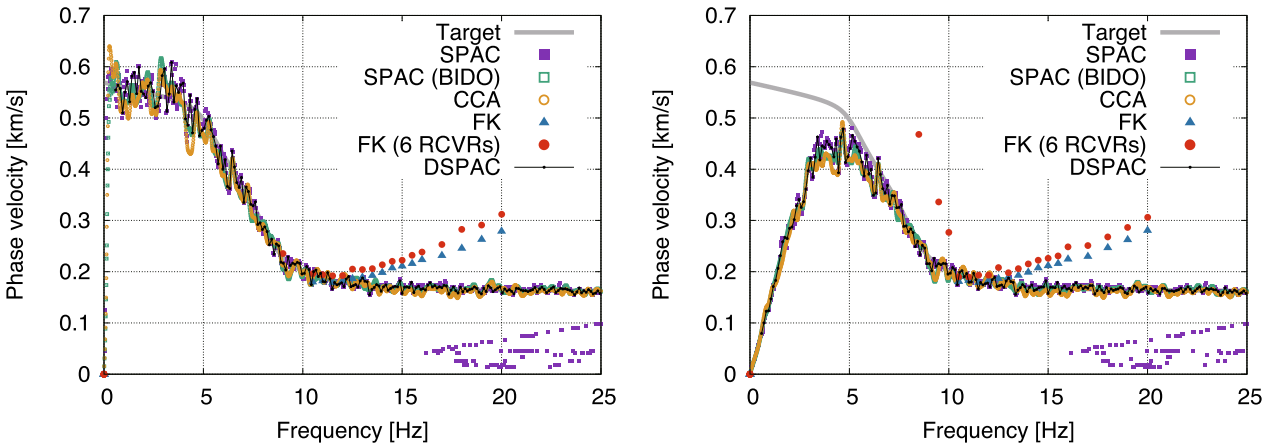


(a) Equilateral triangle (R4-R6-R7)



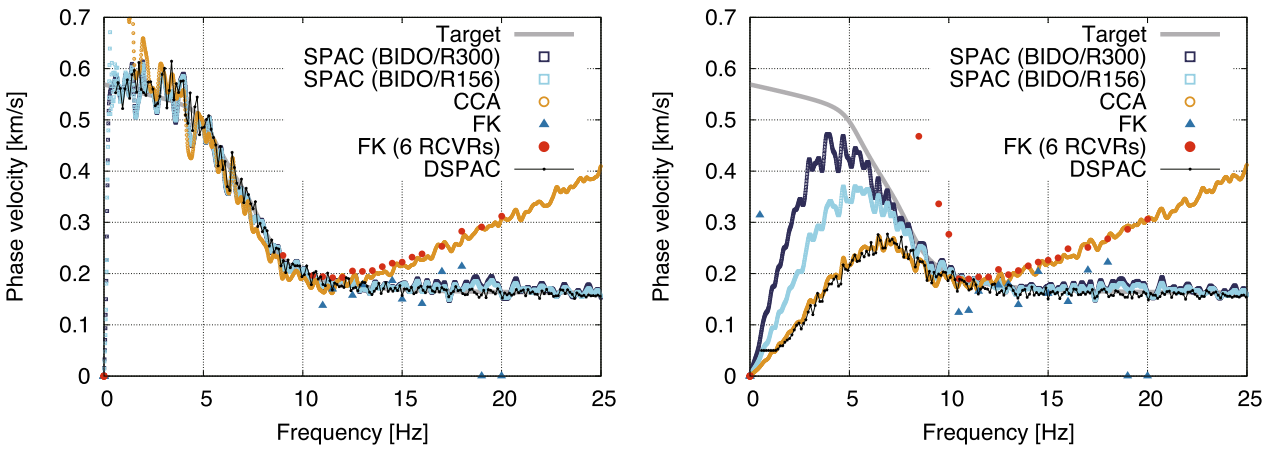
(b) Most-flattened triangle (R1-R6-R7)

Fig. 10 Estimated values of unknown parameters for equilateral and most-flattened triangles in a case including incoherent noise (synthetic data). The legend is the same as that in Fig. 9



(a) Noise-free data (b) Data including incoherent noise

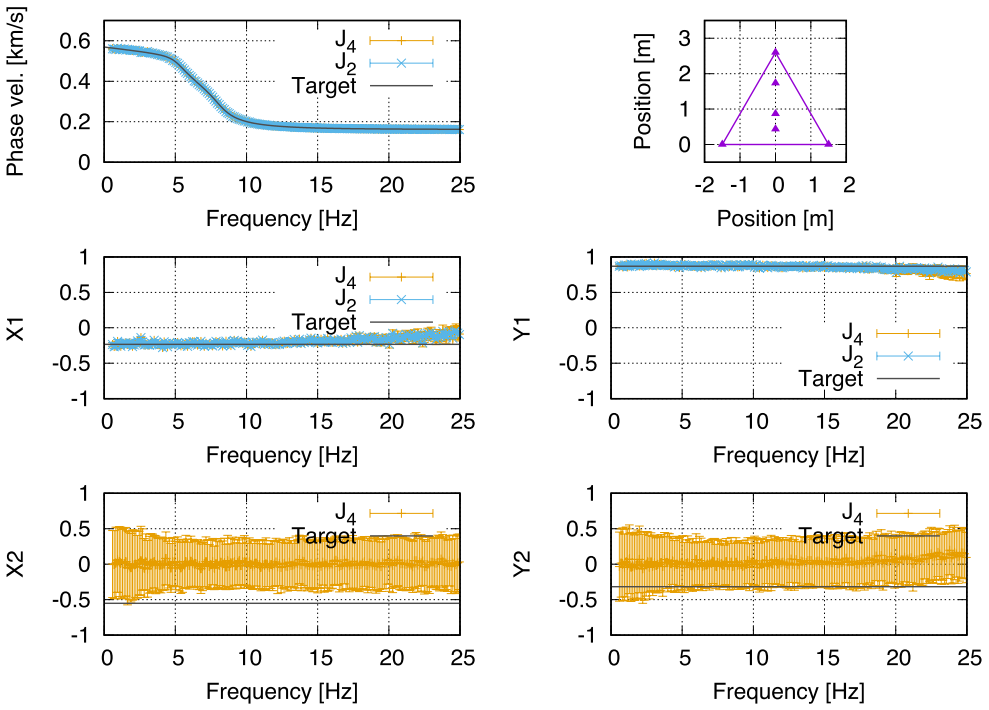
Fig. 11 Comparisons of estimated phase velocities obtained through DSPAC method with those obtained through conventional methods (synthetic data). The array shape is an equilateral triangle (R4-R6-R7), except for "FK (6 RCVRs)". "BIDO" denotes the results when the program code, BIDO 3.0, developed by Cho (2020b), is applied, and RCVR is the receiver. The FK method (6 RCVRs) provides reasonable results in the very narrow frequency range of 9 to 12 Hz. On the other hand, other methods work very well, and provide almost identical results; however, they are strongly affected by noise. To avoid spatial aliasing in the analysis of the FK method, we searched over a limited range of phase velocities, which resulted in the misalignment with the target value in the frequency range above 12 Hz



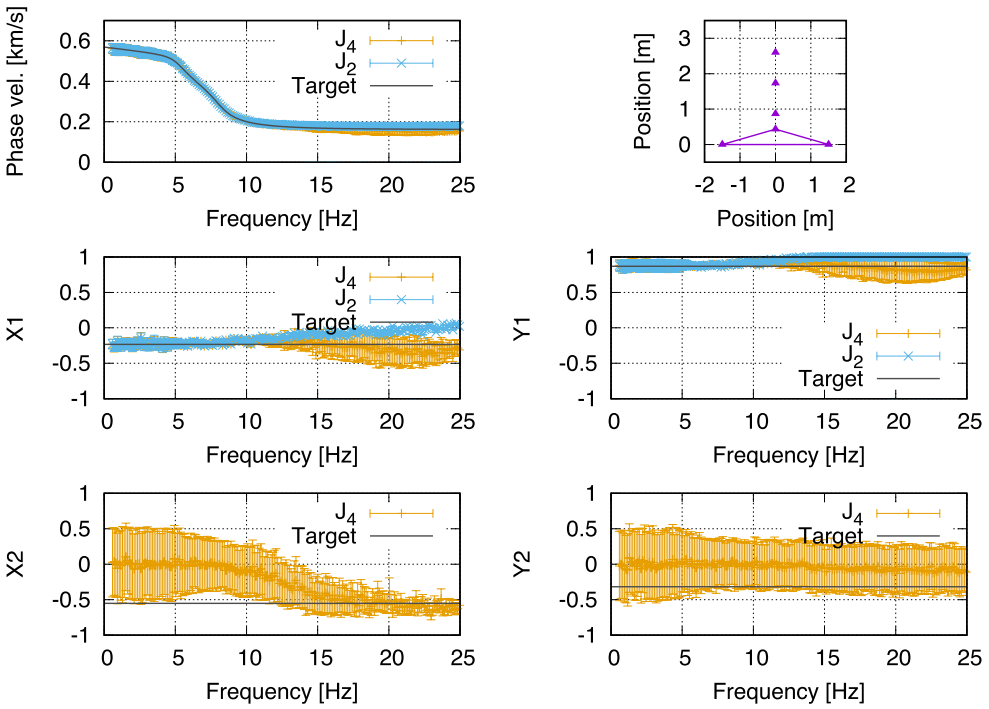
(a) Noise-free data (b) Data including incoherent noise

Fig. 12 Comparisons of estimated phase velocities obtained through DSPAC method with those through conventional methods (synthetic data). The array shape is the most-flattened triangle (R1-R6-R7), except for "FK (6 RCVRs)". Although the SPAC method is not applicable to this array shape, two cases of "SPAC" are applied without or with the limited conditions of azimuthal averaging. The FK and CCA methods do not work so well in almost any frequency range and higher ranges, respectively. The incoherent noise affects the results in the lower frequency range. As a result, reasonable results are obtained for a very limited range of frequency by the CCA method. To avoid spatial aliasing in the analysis of the FK method, we searched over a limited range of phase velocities, which resulted in the misalignment with the target value in the frequency range above 12 Hz.

The high dependency of X_2 and Y_2 on the initial values, regardless of the presence of noise, suggests that these parameters do not contribute to the search for a solution, as expected in the noise-free case. Figure 10 shows that the results obtained by terminating the series in Eq. (7) up to $n = 2$ (orange points) are almost the same as those obtained by censoring up to $n = 1$ (blue points), confirming that the prediction is reasonable.

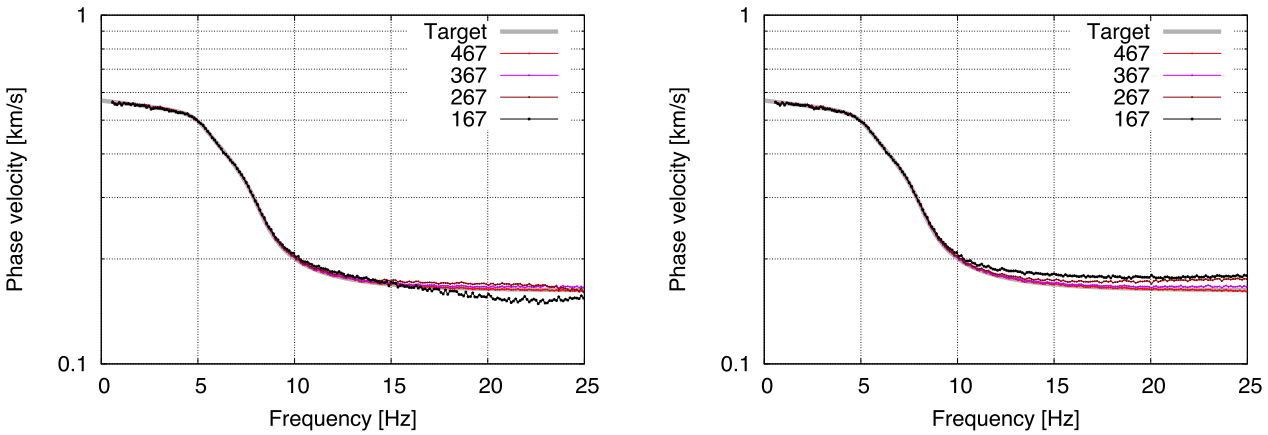


(a) Equilateral triangle (R4-R6-R7)



(b) Most-flattened triangle (R1-R6-R7)

Fig. 13 Estimated values of unknown parameters for equilateral and most-flattened triangles in a case of synthetic anisotropic wave fields. The legend is the same as that in Fig. 9



(a) Series expansions until $n = 2$. (b) Series expansions until $n = 1$.

Fig. 14 Comparison of estimated phase velocities for different shapes of triangle arrays (synthetic data). It is noted that the vertical axes are on a logarithmic scale to emphasize the differences in the estimated values. In the case where the series of Eq. (7) is terminated at $n = 2$, the estimated phase velocities are identical and agree with the target except for the most-flattened triangle (R1-R6-R7). In the case where the series is terminated at $n = 1$, the flatter the array shape, the more different the estimated phase velocity is from the target

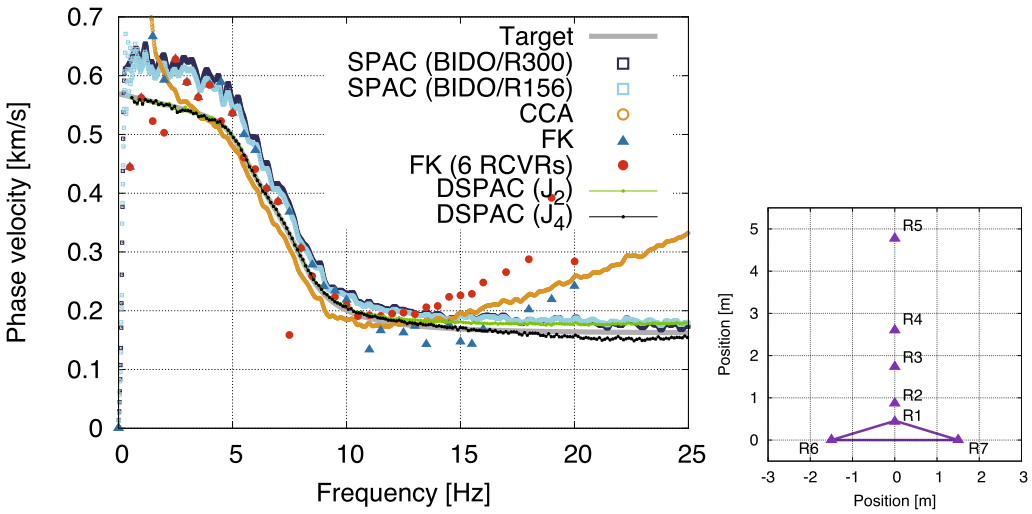


Fig. 15 Comparisons of estimated phase velocities obtained through DSPAC method with those obtained through conventional methods in case of noise-free synthetic data. The array shape is the most-flattened triangle (R1-R6-R7), except for "FK (6 RCVRs)". The legend is the same as that in Fig. 12. The DSPAC method is applied for two types of the termination of series expansion, such as terms until $n = 1$(shown as J_2) and $n = 2$ (shown as J_4). Only the DSPAC (J_4) method provides a good estimation in the wide frequency range. To avoid spatial aliasing in the analysis of the FK method, we searched over a limited range of phase velocities, which resulted in the misalignment with the target value in the frequency range above 12 Hz.

Comparisons with conventional techniques

Figures 11 and 12 compare the DSPAC method with conventional methods, such as the SPAC, FK, and CCA methods. A code developed by one of authors, hereafter referred to as just "SPAC", and another code, BIDO 3.0 (Tada et al. 2010; Cho 2020b), referred to as "SPAC (BIDO)", are used for the analyses of the SPAC method, BIDO 3.0 also for the CCA method, and a code developed by one of authors under the support of Dr. Miyakoshi (see Acknowledgments) for the FK method. Although "SPAC" and "SPAC (BIDO)" are mathematically equivalent, the implementation is different. It would be beneficial to present the results obtained from the different software when we discuss the detailed differences among the methods. Since applying the FK method to the records of three receivers is not likely to provide proper results, the results of the FK spectrum, estimated

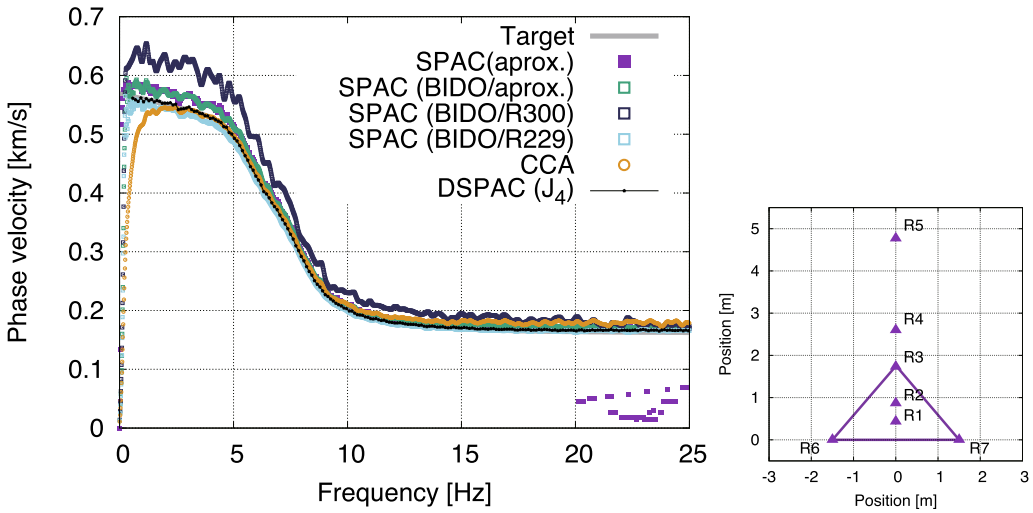


Fig. 16 Comparisons of estimated phase velocities obtained through DSPAC method with those obtained through conventional methods in case of noise-free synthetic data. The array shape is a slightly flattened triangle (R3-R6-R7). The legend is the same as that in Fig. 15 and "approx." denotes the averaged distances among the receivers used to take the azimuthal average. The approximated SPAC method includes bias from the target and the DSPAC method works well

using the records of six receivers (excluding R5) are also presented.

In the case of the equilateral triangle (R4-R6-R7) shown in Fig. 11, almost the same results are obtained, except with the FK method, which is in good agreement with the target. Except for the inability to estimate in the low-frequency range due to noise, the results are reasonable even for data containing noise. The FK method is limited to a very narrow range of 9–12 Hz for which phase velocities can be identified even with a set of six records, and the range is much narrower with noisy data than without noise.

The SPAC method is not applicable to the flattened triangle (R1-R6-R7) because it cannot be azimuthally averaged. For comparison, however, the results of phase velocities estimated from an imperfect SPAC method are also shown: one applies only the SPAC coefficients of the base (R6-R7) (BIDO/R300), and the other applies the average of the SPAC coefficients of the oblique sides (R1-R6 and R1-R7) (BIDO/R156). As discussed in the previous sections, for a case of isotropic wave fields of microtremors, the phase velocity can be identified properly without azimuthal averaging. However, it should be noted that it is not always a good choice to estimate the phase velocity by applying the SPAC method without azimuthal averaging. It can be confirmed whether microtremors arrive isotropically using the imaginary part of CCF, γ_{pq}, that is, $\Im[\gamma_{pq}]$ should vanish in a case of isotropic wavefield (Asten 2006; Cho 2020a). This properties, however, can be obtained after some data processings to the observed records. This means that it is impossible to know whether the wavefields are isotropic or not until the observation and analyses are completed.

Figure 12, for the case of the flattened triangle, shows that the identified parameters are consistent with the targets, except for the CCA and the FK methods, in a case of noise-free wave fields. The FK method is rarely proper for three receivers, and no appropriate solutions are obtained at all for the wave fields including noise. Furthermore, there are almost no solutions even if six receivers are applied for the FK method.

The CCA method can be applied to triangles of any shape, considering a circle passing through the three receivers. While the CCA method is, in principle, sensitive to low frequencies, it is less sensitive to high frequencies, resulting in large errors in the high-frequency range for the flattened array as shown in Fig. 12.

Incoherent noise, as discussed previously, prevents proper estimations in the low-frequency range. This is a common property of the SPAC, CCA, and DSPAC methods, which are algorithms based on CCF. As a result, in the case of the flattened triangular array and data with incoherent noise, the possible range of estimations is very narrow for the CCF method. Under the same situations of the incoherent noise, the DSPAC method can provide proper phase velocities in the same or a wider frequency range than the CCA method.

However, the dispersion curves that were obtained from the proposed method are not as accurate as those obtained from the conventional SPAC method in the low-frequency range. As presented by Cho and Iwata

(2021) and Cho (2022), normalized upper limit wavelength (NULW), the upper limit of the effective wavelength normalized by array radius, is determined mainly by the noise-to-signal ratio (NSR) in the conventional SPAC method. Since the flat triangle used in this study has a smaller average array radius, the upper limit wavelength (ULW) of the flat triangular array becomes smaller than that of the equilateral triangular array if similar properties can be assumed for the DSPAC method and the CCA method. This consideration is consistent with the fact that the SPAC (BIDO/R300), which only uses a base side with receiver spacing of 3 m, shows more accurate solutions than SPAC (BIDO/R156), which uses two oblique sides with receiver spacing of 1.56 m.

Anisotropic wave fields

To study the case of the anisotropic wave fields of microtremors, for example, the case of $\phi_0 = \pi/6$, $\Delta\phi = \pi/4$ is simulated, that is, the sources of microtremors are uniformly distributed in the range of $[\pi/6, 5\pi/12]$. Using the simulated wave fields, the unknown parameters such as phase velocity $c(f)$, X_n, and Y_n $(n = 1, 2)$ are identified. As in the previous section, the results are shown in Fig. 13 for the equilateral triangle (R4-R6-R7) and the most flattened triangle (R1-R6-R7). Incoherent noise is not considered for the analysis. The orange points show the results for the case where the series in Eq. (7) is terminated at $n = 2$, while the blue points show the results for the case where the series is terminated at $n = 1$.

As shown in Table 2, the values for X_1 and Y_1 are expected to be -0.2328 and 0.8697, respectively, while Fig. 13 shows that they are almost consistent with the target values regardless of the array shape in the frequency range below 15 Hz. On the other hand, although the target values for X_2 and Y_2 are -0.5513 and -0.3183, respectively, their estimated values are almost zero in the low frequency range. The estimation is not stable due to the high dependency on the initial values. These results support the previous predictions, namely, X_2 and Y_2 are not sensitive when determining the unknowns.

However, for the most flattened triangle (R1-R6-R7), the value of X_2 approaches -0.5513 above 15 Hz, and the dependency on the initial value becomes smaller. It is suggested that X_2 is sensitive in this frequency range and affects the estimation results. Comparing the orange and blue points in Fig. 13, the difference in the phase velocities is not so clear, although the values for X_1 and Y_1 deviate significantly from the target values for the case in which the series is terminated up to $n = 1$. This may indicate that X_2 cannot be ignored for the optimization.

Figure 14 shows the differences in the identified values of the phase velocity for different array configurations in the cases where the series is terminated at $n = 2$ and at $n = 1$. It should be noted that the vertical axis in this figure is logarithmic in order to emphasize the differences among the identified values. In the case of terminating the series at $n = 2$, all triangles except the most flattened one (R1-R6-R7) are in very good agreement with the target. This suggests that if the maximum value of the interior angles of a triangle is less than about $2\pi/3$, the DSPAC method can provide a proper estimation. However, X_1 and Y_1 are not always properly identified in these cases, and the closer the triangle is to an equilateral triangle, the better the identification results are.

The equilateral and slightly flattened triangles, R3-R6-R7, are almost identical and are in good agreement with the target even if terminated at $n = 1$. However, the flatter the triangle is, the larger the deviation from the target systematically becomes. This is generally consistent with the well-known properties that the accuracy of estimations for the phase velocity becomes worse for flat triangles.

From the above, it is seen that the contribution of X_2 and Y_2 depends not only on the array configuration but also on the frequency. Furthermore, although the results are not shown in this paper, the contribution of X_2 and Y_2 also depends on the arrival directions of the microtremors for different combinations of ϕ_0 and $\Delta\phi$.

At the end of this section, the phase velocities estimated from the waveforms of the most flattened triangular array using the DSPAC method are compared with those estimated using other conventional methods. Figure 15 shows the phase velocities identified from the SPAC, CCA, FK, and DSPAC methods. For the SPAC method, the phase velocities obtained from the imperfect calculations are shown: one is without azimuthal averaging from only one pair of CCFs at the bottom (BIDO/R300) and the other is from the average of two pairs of CCFs on the oblique sides (BIDO/R156). The results from the FK method are also shown using six receivers, excluding R5. The results by the DSPAC method are shown for two cases where the series of Eq. (7) is terminated up to $n = 1$ and $n = 2$.

Since the wave fields are not isotropic, the SPAC method without azimuthal averaging does not yield proper results. The phase velocity identified from the DSPAC method, terminated at $n = 1$, asymptotically approaches the results obtained from the SPAC method in the high-frequency range. This suggests that the error is due to terms including the higher-order Bessel functions. The results from the CCA method are close to the target values in the range of 4 to 9 Hz, although the proper solutions are not obtained, and they deviate significantly from the target values in other frequency ranges. Although the FK method provides no proper identifications from three receivers, the results from six

receivers are better than those from three. However, the identified phase velocities are slightly larger than the target velocities.

Figure 16 shows a comparison of the results from the DSPAC method with the SPAC method for a slightly irregular triangle. It is a common situation that the position of the vertex R4 of an equilateral triangle (R4-R6-R7) is slightly displaced to form an irregularly shaped triangle, such as R3-R6-R7. If the misalignment of the receivers is not very large, an approximation may be used to estimate the phase velocity using the azimuthal averaging of the SPAC coefficients for all the possible pairs of receivers and averaged distances of all the pairs.

The DSPAC method is expected to provide proper estimates in such cases. The DSPAC method is compared with (i) the SPAC method without azimuthal averaging, (ii) the SPAC method with an approximation using the average of the distances among the receivers, which has been used empirically, and (iii) the CCA method for the waveforms obtained from the slightly flattened triangle, R3-R6-R7. In this case, the distance of the base (R6-R7) is 3 m, the oblique sides (R3-R6 and R3-R7) are 2.29 m, and their average is 2.53 m.

In case (i), where only the base (R6-R7) is used, the identified phase velocity deviated significantly from the target velocity, because of the anisotropic wave fields. The averaged SPAC coefficients over the CCFs from the two oblique sides provide almost proper identification, although the results are very slightly underestimated in the low-frequency range and very slightly over-estimated in the high-frequency range. Although it is not intentional, the phase velocity can be estimated properly by averaging the two SPAC coefficients for the oblique sides because the arrival directions of the microtremors are almost parallel to one of the sides (R3-R6) and the arrival directions are approximately perpendicular to the other side (R3-R7) in the settings of the example. In other words, it is just fortuitous that the proper phase velocity is obtained. Therefore, it should be noted that the proper results are not always obtained from the two isosceles sides of any shape of isosceles triangles to find the phase velocity.

In case (ii) of the approximated SPAC method, that is azimuthally averaging and applying the averaged distance to estimate the phase velocity, the identified phase velocity is over-estimated. In particular, the error in the low-frequency range is conspicuous.

The CCA method, which is case (iii), provides a proper solution over a wide frequency range because the array shape is not an extremely flattened triangular, although it is over-estimated in the high-frequency range. In contrast, the DSPAC method provides proper results over the entire frequency range.

These results depend on the arrival directions of microtremors and the expanse of the source area; thus, it is not obvious what errors will be included if the proper solutions are not obtained. Moreover, there is no way to know whether the proper solution has been obtained or not for the data obtained from real fields. Therefore, the DSPAC method is a very useful algorithm for microtremor surveys under unknown sources, because it can, in principle, provide appropriate solutions regardless of the source distribution of microtremors.

Field test

To confirm the applicability of the DSPAC method, it is applied to actual data. Comparing the DSPAC method with the conventional methods, such as the SPAC, the CCA, and the FK methods, we validate the accuracy of the DSPAC method.

Summary of observations

Observations were carried out on the morning of July 20, 2022 on the paved road between the parking lot next to the J2 building and the tennis court of the Suzukake-dai Campus, Tokyo Institute of Technology, Japan. The site

Fig. 17 Observation site and settings of sensors for the field test. The site is located at the Suzukake-dai Campus, Tokyo Institute of Technology, Yokohama, Japan

was located on Pleistocene hills with cut and fill. The receivers were set as shown in Fig. 5.

Six integrated seismographs (CloverTech LA-352) were used for simultaneous observations: the sensor is a MEMS (micro-electro mechanical systems) accelerometer (Epson A352), made from a crystal oscillator, and the data logger is based on a board computer with time synchronization through the GPS (Global Positioning System). The data are recorded, 200 samples per second (sps), for 30 min. Figure 17 shows a distant view of the observation site and the settings of the sensors.

It is known from nearby borehole data that a surface layer a thickness of 5 to 20 m consists of fill, silt and clay whose shear wave velocities are around 200 m/s. An engineering base of mudstone with a shear wave velocity of 600 to 700 m/s exists below the surface layer. The ground water level is −1.0 to −2.0 m from the surface. However, details of the velocity structure at the site are unknown. For simplicity, a velocity structure that can accurately explain the phase velocity estimated from the SPAC method, is obtained, and its theoretical dispersion curve is shown for reference in the following figures. The assumed velocity structure is shown in Table 3. The theoretical dispersion curve for the fundamental mode of Rayleigh waves, which corresponds to the structure, is identical to the dispersion curve of the target used in the previous section and shown in Fig. 8b.

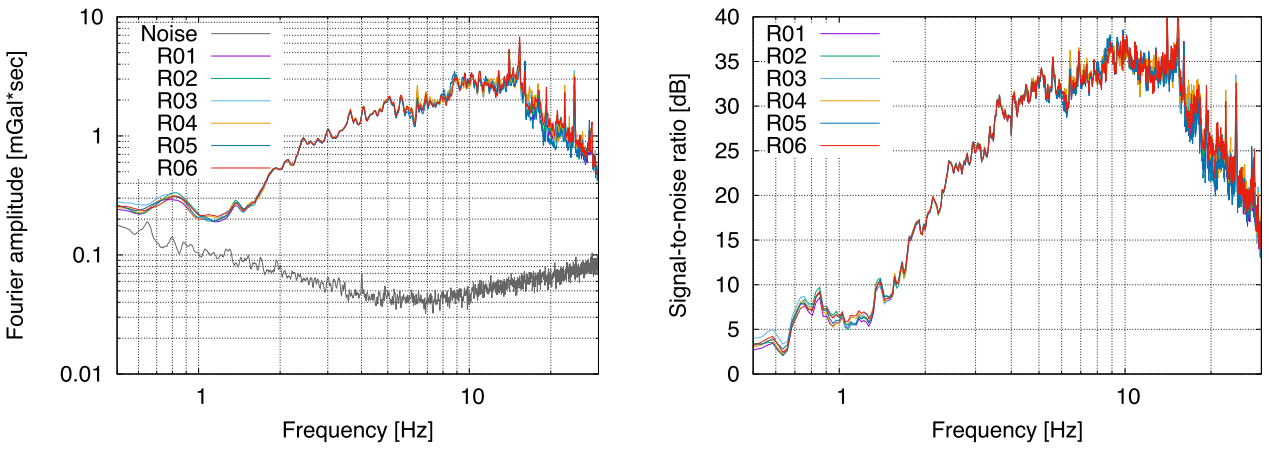
Results and discussion

Selecting 11 portions of 40.96-s length, to exclude disturbances caused by cars and pedestrians passing by the observation site, and applying the Fourier transform to each portion, the power spectrum at each receiver and the cross spectrum and CCF among the receivers are calculated from the average over the 11 portions.

The Fourier spectra of the observed records are shown in Fig. 18. To identify the noise level of the sensors, we obtained records at a bedrock site in a mountainous area, where the level of microtremors is known to be extremely low. Since the Fourier spectra of the observed records, shown as a gray solid line, is larger than the known microtremor level at the bedrock site, it can be regarded as the self-noise. The colored lines are the Fourier amplitudes of the records obtained through the simultaneous observation using six seismographs at the Suzukake-dai Campus. The shapes of the spectra are almost the same due to the small distance among the receivers, and it is unlikely that incoherent disturbances are contained only in a particular sensor.

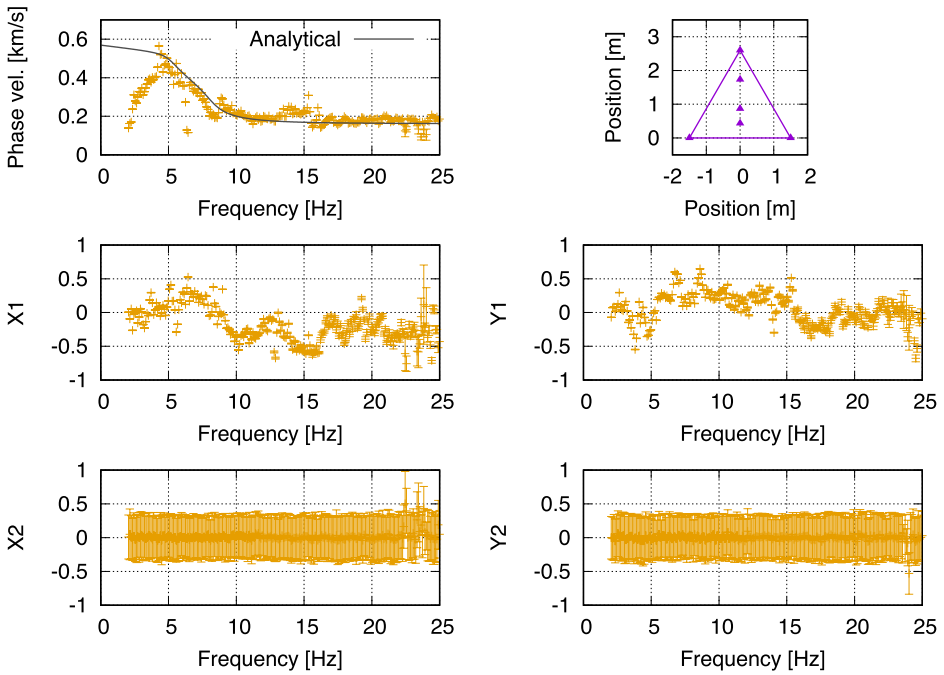
Figure 18 shows that the power ratio of the signal to noise is more than 20 dB (100 times power and about 10 times amplitude) in the range of 2.3 to 25 Hz and more than 30 dB (1000 times power and about 31 times amplitude) in the range of 4 to 16 Hz. Thus, it can be assumed that the effect of the noise is low enough for analyses in the range of 4 to 16 Hz.

Figure 19 shows identified phase velocity $c(f)$ and unknown parameters X_n and Y_n $(n = 1, 2)$ which depend on the arrival directions of microtremors, using the records of the equilateral triangle (R4-R6-R7) and the most flattened triangle (R1-R6-R7). The phase velocities are estimated in the range of 5 to 25 Hz, although the most flattened triangle shows poor estimation accuracy in the low-frequency range, which corresponds well with the properties observed in Fig. 10. Theoretically, the values of X_1 and Y_1 shown in Fig. 19a and those in Fig. 19b

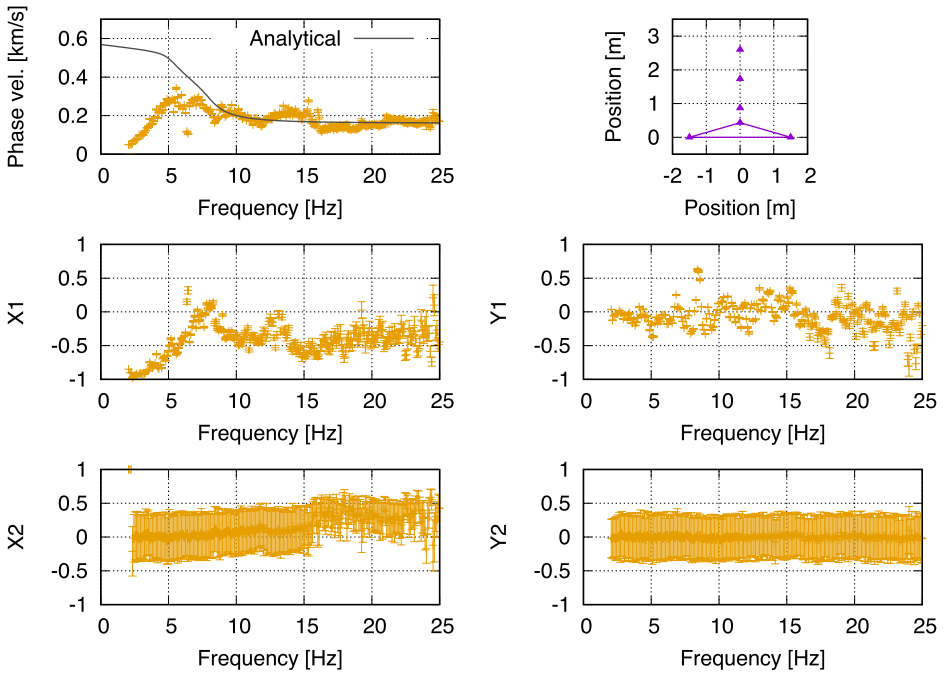


(a) Fourier amplitudes (b) Signal-to-noise ratio (SNR)

Fig. 18 Fourier amplitudes of observed microtremors at the site (solid colored lines) and noise floor of the sensor (solid gray line). The observed data have sufficient SNR in the frequency range of 4 to 16 Hz. Note that the SNR is shown as the power ratios which are calculated as colored lines divided by a gray line in the panel a

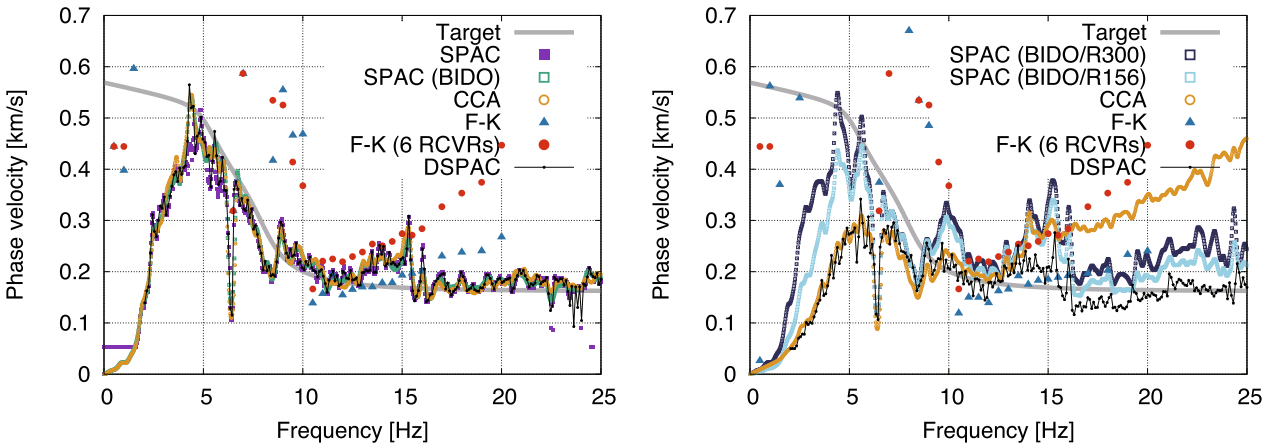


(a) Equilateral triangle (R4-R6-R7)



(b) Most-flattened triangle (R1-R6-R7)

Fig. 19 Estimated values of unknown parameters for equilateral and most-flattened triangles using the observed data. The legend is the same as that in Fig. 9. The phase velocities, X_1 and Y_1, are determined well, and the estimated values for X_2 and Y_2 depend on the initial data set and are not stable. The solid black lines show theoretical dispersion curves for the fundamental mode of Rayleigh waves propagating in a ground with the velocity structure shown in Table 3



(a) Equilateral triangle (R4-R6-R7) (b) Most-flattened triangle (R1-R6-R7)

Fig. 20 Comparisons of estimated phase velocities obtained through DSPAC method with those obtained through conventional methods using the observed data. The legend is the same as those in Figs. 11 and 12. The properties of each method agree well with the results of the numerical calculations. To avoid spatial aliasing in the analysis of the FK method, we searched over a limited range of phase velocities, which resulted in the misalignment with the target value in the frequency range above 12 Hz

Table 3 Estimated velocity structure at site for array observation

Thickness (km)	Density (t/m^3)	Vs (km/s)	Vp (km/s)
0.00800	1.40	0.170	1.000
0.00400	1.50	0.250	1.100
0.00300	1.65	0.350	1.400
∞	2.00	0.600	1.900

should be identical, respectively. However, since the phase velocity was not obtained correctly for the lower frequency range, the values of X_1 and Y_1 also show the same trend. These properties can be attributed to the fact that the parameters are identified simultaneously in the proposed method. The dependency of X_1 and Y_1 on the initial values is small, and nonzero values are stably estimated. This suggests that the microtremors do not arrive isotropically, and that the arrival directions of microtremors depend on the frequency as shown by the shapes of X_1 and X_2. As in the case of the numerical simulation in the previous section, X_2 and Y_2 have a great dependency on the initial values and cannot be estimated stably. However, for the flattened triangle, it can be seen that the estimated values change in the frequency range above 15 Hz for X_2. Although an unstable estimation, this behavior of the estimates is considered to be the same as that seen in Fig. 13b. This suggests that the sensitivity of X_2 to the real part of the CCF depends on the array configuration.

Figure 20 compares the phase velocity estimated by the DSPAC method with that estimated by the conventional methods using the observed data. It is impossible to verify which method is more accurate, as the correct solution is unknown. If the results are comparable to the phase velocities obtained by the conventional SPAC method, it is safe to conclude that the DSPAC method provides somewhat reliable results. This is considered acceptable based on previous numerical investigations.

In the case of the equilateral triangle (R4-R6-R7), almost the same results are obtained for all methods, except for the FK method, including fluctuations in the phase velocity. This leads us to believe that the estimated phase velocities are reasonable. The lack of phase velocity in the frequency range below 5 Hz is considered to be due to the effect of incoherent noise, corresponding to the frequency range where the signal-to-noise ratio is less than 30 dB.

On the other hand, in the case of the flattened triangle (R1-R6-R7), the CCA and DSPAC methods do not properly estimate the phase velocity in the low-frequency range. In addition, the results from the CCA method deviate significantly in the high-frequency range. As the SPAC method, which uses only the base (R6-R7), yields almost the same phase velocity as the equilateral triangle case, it is also expected that the microtremors arrive in a near-isotropic state. However, the results of Fig. 19 show that X_1 and Y_1 are not perfectly isotropic because they are not zero, which is not unrelated to the deviations of the imperfect SPAC

methods, such as R300 and R156, from the case of the equilateral triangle above 15 Hz.

From the above, it can be stated that the properties of each method correspond well to those observed in the numerical calculations.

Unfortunately, the FK method is not able to estimate phase velocities even though the records obtained from six receivers are used. The arrival directions of microtremors should be discussed using the results from the FK method and the values of X_1 and Y_1 from the DSPAC method. However, the results of the FK method are too poor to use in a discussion on the arrival directions; therefore, the issue of the arrival directions will be left for future works.

Conclusions

The direct spatial auto-correlation (DSPAC) method was proposed in this study. This method directly identifies sufficiently sensitive parameters included in the complex coherency functions (CCFs), which enables high accuracy estimation of the phase velocity without any restrictions on the array configurations. The conclusions reached in this study are listed as follows:

- The stochastic properties of the parameters emerge in the real part of CCFs, that have been treated as the terms of directional aliasing (Henstridge 1979; Cho 2020a), were analytically examined. Specifically, these are the parameters determined by the intensity distribution with respect to the arrival directions of the microtremors and take zero if the wave field under consideration is isotropic.
- The DSPAC method was shown to provide solutions as accurately as the conventional SPAC method as long as extremely flattened arrays are not used.
- The DSPAC method was shown to be able to estimate phase velocity more accurately than the centerless circular array (CCA) method in the higher frequency range.
- The DSPAC method was shown to be able to estimate phase velocity as accurately as the CCA method in the lower frequency range.
- Because the DSPAC method uses the particle swarm optimization (PSO) as its core algorithm, the method offers a way to identify the sufficiently sensitive parameters, through the comparison of the initial value dependency of obtained solutions.
- The applicability of the DSPAC method was confirmed through the analysis of the actual field data.

Further study is necessary on certain topics, which could not be sufficiently discussed in this study, for example, the sensitivity and accuracy of the estimation of the unknown parameters, and the quantitative relationships between the values of the unknown parameters and the arrival directions of microtremors, which are the physical meanings of the parameters.

Appendix A: derivation of equations (17) and (18)

In a case where the arrival directions of microtremors, θ_ℓ, is limited in the range of $[\phi_0, \phi_1]$ ($\phi_1 = \phi_0 + \Delta\phi$) and follows a uniform distribution in the range, the variance, $\mathrm{Var}[\xi_\mathrm{n}(\phi_0; \Delta\phi)]$, for parameter ξ_n can be obtained as follows:

The variance $\mathrm{Var}[\alpha_\ell \cos(2\mathrm{n}\theta_\ell)]$ is

$$
\begin{aligned}
\mathrm{Var}[\alpha_\ell \cos(2\mathrm{n}\theta_\ell)] &= \mathrm{E}[\{\alpha_\ell \cos(2\mathrm{n}\theta_\ell)\}^2] - \{\mathrm{E}[\alpha_\ell \cos(2\mathrm{n}\theta_\ell)]\}^2 \\
&= \mathrm{E}[\alpha_\ell^2]\mathrm{E}[\cos^2(2\mathrm{n}\theta_\ell)] - \left\{\frac{1}{\mathrm{L}}\frac{\sin(2\mathrm{n}\phi_1) - \sin(2\mathrm{n}\phi_0)}{2\mathrm{n}\Delta\phi}\right\}^2 \\
&= \{\mathrm{Var}[\alpha_\ell] + \mathrm{E}[\alpha_\ell]^2\} \int_{\phi_0}^{\phi_1} \cos^2(2\mathrm{n}\theta_\ell)\frac{1}{\Delta\phi}\mathrm{d}\theta_\ell - \left\{\frac{1}{\mathrm{L}}\mathrm{E}[\xi_\mathrm{n}(\phi_0; \Delta\phi)]\right\}^2 \\
&\approx \left(\frac{1}{3L^2} + \frac{1}{L^2}\right)\frac{1}{\Delta\phi}\int_{\phi_0}^{\phi_1}\frac{1+\cos(4n\theta_\ell)}{2}d\theta_\ell - \frac{\mathrm{E}[\xi_\mathrm{n}(\phi_0; \Delta\phi)]^2}{L^2} \\
&= \frac{4}{3L^2}\frac{1}{2\Delta\phi}\left\{\Delta\phi + \frac{1}{4n}[\sin(4n\theta_\ell)]_{\phi_0}^{\phi_1}\right\} - \frac{\mathrm{E}[\xi_\mathrm{n}(\phi_0; \Delta\phi)]^2}{L^2} \\
&= \frac{4}{3L^2}\left\{\frac{1}{2} + \frac{\sin(4n\phi_1) - \sin(4n\phi_0)}{8n\Delta\phi}\right\} - \frac{\mathrm{E}[\xi_\mathrm{n}(\phi_0; \Delta\phi)]^2}{L^2}.
\end{aligned}
\tag{A.1}
$$

Applying the relationship $\mathrm{Var}[\sum_{\ell=1}^{\mathrm{L}} \alpha_\ell \cos(2\mathrm{n}\theta_\ell)] = \sum_{\ell=1}^{L} \mathrm{Var}[\alpha_\ell \cos(2\mathrm{n}\theta_\ell)]$,

$$\begin{aligned}\mathrm{Var}[\xi_{\mathrm{n}}(\phi_0;\Delta\phi)] &= \mathrm{Var}[\textstyle\sum_{\ell=1}^{\mathrm{L}} \alpha_\ell \cos(2\mathrm{n}\theta_\ell)] = \sum_{\ell=1}^{\mathrm{L}} \mathrm{Var}[\alpha_\ell \cos(2\mathrm{n}\theta_\ell)] \\ &\approx \sum_{\ell=1}^{L}\left[\frac{4}{3L^2}\left\{\frac{1}{2}+\frac{\sin(4n\phi_1)-\sin(4n\phi_0)}{8n\Delta\phi}\right\}-\frac{\mathrm{E}[\xi_{\mathrm{n}}(\phi_0;\Delta\phi)]^2}{L^2}\right] \\ &= \frac{4}{3L}\left\{\frac{1}{2}+\frac{\sin(4n\phi_1)-\sin(4n\phi_0)}{8n\Delta\phi}\right\}-\frac{\mathrm{E}[\xi_{\mathrm{n}}(\phi_0;\Delta\phi)]^2}{L}\end{aligned} \tag{A.2}$$

is obtained.

Similarly for ζ_n, the following equation is obtained:

$$Var[\zeta_n(\phi_0;\Delta\phi)] \approx \frac{4}{3L}\left\{\frac{1}{2}-\frac{\sin(4n\phi_1)-\sin(4n\phi_0)}{8n\Delta\phi}\right\} - \frac{E[\zeta_n(\phi_0;\Delta\phi)]^2}{L}. \tag{A.3}$$

Abbreviations

CCA	Centerless circular array
CCF	Complex coherency function
DEM	Direct estimation method
DSPAC	Direct SPAC
FK	Frequency-wavenumber
GPS	Global Positioning System
MEMS	Micro-electro mechanical systems
NSR	Noise-to-signal ratio
PDF	Probability density function
PSO	Particle swarm optimization
RCVR	Receiver
RMS	Root mean square
SNR	Signal-to-noise ratio
SPAC	Spatial auto-correlation

Acknowledgements
The authors would like to thank Ms. Yoko Ohashi for her helpful support with the microtremor observations. Constructive comments from two reviewers significantly improved the manuscript. The authors used the BIDO 3.0 program code, which was developed by Tada et al. (2010) and downloaded from http://staff.aist.go.jp/ikuo-chou/bidodl_en.html (Cho 2020b), for the analyses of the CCA and a part of the SPAC methods. The authors also used Fortran codes for the FK method, which was developed by HM under the helpful advice and support of Dr. Ken Miyakoshi of Ohsaki Research Institute., INC., Japan.

Author contributions
HM designed the framework of the research, and HK and KI designed the research. HK and HM derived the theoretical formulations, carried out the observations of the microtremors, and wrote the article. HK, HM, and HT performed the numerical calculations. All the authors read and approved the final manuscript.

Funding
A part of this work was financially supported by the JSPS (Japan Society for the Promotion of Science) KAKENHI Grant Number JP19H02400.

Availability of data and materials
The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests
The authors have no competing interests.

Author details
[1]Department of Civil and Environmental Engineering, Tokyo Institute of Technology, Yokohama 226–8502, Japan. [2]Technical Research Institute, Kajima Corporation, Tokyo 182–0036, Japan.

Received: 31 October 2022 Accepted: 22 April 2023
Published online: 18 May 2023

References

Aki K (1957) Space and time spectra of stationary stochastic waves, with special reference to microtremors. Bull Earthq Res Inst Univ Tokyo 35:415–457
Asten MW (2006) On bias and noise in passive seismic data from finite circular array data processed using SPAC methods. Geophysics 71:V153–V162. https://doi.org/10.1190/1.2345054
Asten MW, Hayashi K (2018) Application of the spatial auto-correlation method for shear-wave velocity studies using ambient noise. Surv Geophys 39:633–659. https://doi.org/10.1007/s10712-018-9474-2
Asten MW, Stephenson WJ, Hartzell S (2019) Spatially averaged coherencies (krSPAC) and Rayleigh effective-mode modeling of microtremor data from asymmetric arrays. Geophysics 84:EN47–EN56
Asten MW, Askan A, Karimzadeh S (2023) Blind study site assessment of shear-wave velocity at Kumamoto city, Japan, using direct-fitting SPAC methods. Earth Planets Space 75:40. https://doi.org/10.1186/s40623-023-01801-y
Burke EK, Kendall K (2013) Search methodologies: Introductory tutorials in optimization and decision support techniques. Springer Science+Business Media, New York
Capon J (1969) High-resolution frequency-wavenumber spectrum analysis. Proc IEEE 57:1408–1418. https://doi.org/10.1109/PROC.1969.7278
Chávez-García FJ, Rodríguez M, Stephenson WR (2005) An alternative approach to the SPAC analysis of microtremors: exploiting stationarity of noise. Bull Seismol Soc Am 95:277–293. https://doi.org/10.1785/012003179
Chávez-García FJ, Rodríguez M, Stephenson WR (2006) Subsoil structure using SPAC measurements along a line. Bull Seismol Soc Am 96:729–736. https://doi.org/10.1785/0120050141
Cho I (2019) Compensating for the impact of incoherent noise in the spatial autocorrelation microtremor array method. Bull Seismol Soc Am 109:199–211. https://doi.org/10.1785/0120180153
Cho I (2020) Two-sensor microtremor SPAC method: potential utility of imaginary spectrum components. Geophys J Int 220:1735–1747. https://doi.org/10.1093/gji/ggz454
Cho I (2020) BIDO: Package of analysis codes to identify properties of surface waves using circular-array records of microtremors. https://staff.aist.go.jp/ikuo-chou/bidodl_en.html. Accessed 1 Oct 2022

Cho I (2022) Array-size dependency of the upper limit wavelength normalized by array radius for the standard spatial autocorrelation method. Earth Planets Space 74:75. https://doi.org/10.1186/s40623-022-01641-2
Cho I, Iwata T (2021) Limits and benefits of the spatial autocorrelation microtremor array method due to the incoherent noise, with special reference to the analysis of long wavelength ranges. J Geophys Res Solid Earth 126:e2020JB019850. https://doi.org/10.1029/2020JB019850
Cho I, Tada T, Shinozaki Y (2004) A new method to determine phase velocities of Rayleigh waves from microseisms. Geophysics 69:1535–1551. https://doi.org/10.1190/1.1836827
Cho I, Tada T, Shinozaki Y (2006) A generic formulation for microtremor exploration methods using three-component records from a circular array. Geophys J Int 165:236–258. https://doi.org/10.1111/j.1365-246X.2006.02880.x
Clerc M (2006) Particle swarm optimization, English. ISTE Ltd, London
Goto H, Mitsunaga H, Inatani M, Iiyama K, Hada K, Ikeda T, Takaya T, Kimura S, Akiyama R, Sawada S, Morikawa H (2018) Shallow subsurface structure estimated from dense aftershock records and microtremor observations in Furukawa district, Miyagi, Japan. Explor Geophys 48:16–27. https://doi.org/10.1071/EG16113
Hayashi K, Asten MW, Stephenson W, Cornou C, Hobiger H, Pilz M, Yamanaka H (2022) Microtremor array method using SPAC analysis of Rayleigh-wave data. J Seismol 26:601–627. https://doi.org/10.1007/s10950-021-10051-y
Henstridge JD (1979) A signal processing method for circular arrays. Geophysics 44:179–184. https://doi.org/10.1190/1.1440959
Horike M (1985) Inversion of phase velocity of long-period microtremors to the S-wave-velocity structure down to the basement in urbanized areas. J Phys Earth 33:59–96
Kennedy J, Eberhart R (1995) Particle swarm optimization. Proc of ICNN'95 – Int'l Conf on Neural Networks 1942–1948. https://doi.org/10.1109/ICNN.1995.488968
Komazawa M, Morikawa H, Nakamura K, Akamatsu J, Nishimura K, Sawada S, Erken A, Onalp A (2002) Bedrock Structure in Adapazari, Turkey – a possible cause of severe damage by the 1999 Kociaeli earthquake. Soil Dyn Earthq Eng 22:829–836. https://doi.org/10.1016/S0267-7261(02)00105-7
Konno K (2000) Theoretical study on the spatial auto-correlation method for estimation of underground structure using microtremor. Proc Jpn Soc Civil Eng 654:367–375
Lacoss RT, Kelly EJ, Toksöz MN (1969) Estimation of seismic noise structure using arrays. Geophysics 34:21–38. https://doi.org/10.1190/1.1439995
Morikawa H, Sawada S, Akamatsu J (2004) A method to estimate phase velocities of Rayleigh waves using microseisms simultaneously observed at two sites. Bull Seismol Soc Am 94:961–976. https://doi.org/10.1785/012003020
Okada H (2003) The microtremor survey method. Geophysical Monograph Series Vol 12 Society of Exploration Geophysicists, Tulsa, Oklahoma
Okada H (2006) Theory of efficient array observations of microtremors with special reference to the SPAC method. Explor Geophys 37:73–85. https://doi.org/10.1071/EG06073
Shiraishi H, Asanuma H (2009) Direct estimation method to extract phase celocities of Rayleigh waves in microtremors using arbitrary geometry array. BUTSURI-TANSA 62:339–350. https://doi.org/10.3124/segj.62.339
Shiraishi H, Matsuoka T (2005) Complex coherence function represented by discrete formula on the basis of Lamb's problem for Rayleigh waves: An application for new interpretation of the spatial autocorrelation method. BUTSURI-TANSA 58:137–146. https://doi.org/10.3124/segj.58.137
Shiraishi H, Matsuoka T, Asanuma H (2005) Direct estimation method to extract phase velocities of Rayleigh waves in microtremors using circular arrays. BUTSURI-TANSA 58:643–651. https://doi.org/10.3124/segj.58.643
Shiraishi H, Matsuoka T, Asanuma H (2006) Direct estimation of Rayleigh wave phase velocity in microtremors. Geophys Res Lett 33:L18307. https://doi.org/10.1029/2006GL026723
Stephenson WJ, Asten MW, Odum JK, Frankel AD (2019) Shear-wave velocity in the Seattle basin to 2 km depth characterized with the krSPAC microtremor array method: insights for urban basin-scale imaging. Seismol Res Lett 90(3):1230–1242. https://doi.org/10.1785/0220180194
Tada T, Cho I, Shinozaki Y (2010) New horizons in the utility of horizontal-motion microtremors. Proc 7th Int'l Conference on Urban Earthquake Engineering, Center for Urban Earthquake Engineering, Tokyo Institute of Technology, Tokyo, Japan
Yokoi T, Margaryan S (2008) Consistency of the spatial autocorrelation method with seismic interferometry and its consequence. Geophys Prospect 56:435–451. https://doi.org/10.1111/j.1365-2478.2008.00709.x

Publisher's Note