Spatially averaged coherencies (krSPAC) and Rayleigh effective-mode modeling of microtremor data from asymmetric arrays

Michael W. Asten¹, William J. Stephenson², and Stephen Hartzell²

ABSTRACT

The spatial autocorrelation and spatially averaged coherency (SPAC) methods of processing Rayleigh-wave microtremor noise observations for estimation of S-wave velocity profiles traditionally require a circular or triangular array symmetry to allow spatial (azimuthal) averaging of interstation coherencies over a constant station separation. Common processing methods allow for station separations to vary by typically ±10% in the azimuthal averaging before degradation of the SPAC spectrum is excessive. Transformation of a set of frequency-coherency spectra to wavenumber-coherency spectra (kr spectra) allows spatial averaging of spectra from multiple pairs of sensors irrespective of differences in spatial separation of the pairs. The method is called krSPAC and is implemented by iterative direct fitting of observed and model kr spectra to determine an optimal layered-earth S-wave velocity model. The observed kr spectra are updated with each iteration of the velocity model because the wavenumber is a function of model phase velocity that varies with each iteration of the modeling process. The method proves applicable when modeling either with the assumption of fundamental mode Rayleigh-wave propagation or with a summation of fundamental and higher modes. The method proves robust when compared with alternative methodologies using symmetric and asymmetric arrays on a sample of synthetic data and on field data in which station spacings vary from 70 to 800 m side lengths.

INTRODUCTION

Array-based microtremor methods make use of the fact that background seismic noise is dominated by surface waves; if only vertical components are measured, then it is the Rayleigh-wave fraction of energy that is observed. The use of an array allows measurement of the dispersive property of Rayleigh waves (phase velocities vary with frequency) from which the S-wave velocity $V_S$ profile (variation with depth) may be obtained for the location by model fitting or inversion.

The choice of method for extracting dispersion curves (DCs) from passive seismic or microtremor energy depends strongly on assumptions as to the nature of the waveform and on the instrumental and logistical constraints on array complexity used for field observations. Concepts of surface-wave propagation and how they are used to extract information on layering of the earth are reviewed by Levshin et al. (2018) with an emphasis on applications to regional and crustal studies of the earth. Foti et al. (2015) review surface-wave methodology applied to the upper tens and hundreds of meters of the earth. In this paper, we limit ourselves to a discussion of Rayleigh-type surface waves only (because they are detectable using vertical-component seismometers) and we limit our processing methodology to the use of small arrays (typically four to seven seismometers) processed using variants of the spatial autocorrelation or spatially averaged coherency (SPAC) method. The basic SPAC method as originated by Aki (1957) contains two important assumptions, one of an azimuthal distribution of wave sources and another of summation of fundamental and higher modes. The method proves robust when compared with alternative methodologies using symmetric and asymmetric arrays on a sample of synthetic data and on field data in which station spacings vary from 70 to 800 m side lengths.
symmetric. Foti et al. (2018) and Asten and Hayashi (2018) provide reviews and comparisons of these methods. Seismic interferometry is a more recent alternative methodology (Mularia and Castellaro, 2013; Cheng et al., 2016; Xu et al., 2017). This may be considered as a time-domain method analogous to the frequency-domain method represented by SPAC (e.g., Tsai and Moschetti, 2010), possessing an advantage of not requiring array symmetry, but it is mathematically more cumbersome (Asten and Hayashi, 2018). The objective of this paper is to develop a variant of the SPAC method maintaining the advantages of frequency-domain computation and application to small arrays, but releasing the user from the requirement of array symmetry.

A fourth approach to microtremor observation and analysis used either as an adjunct to or alternative to array studies is to record 3C data on at least one sensor. This allows the horizontal-vertical spectral ratio (HVSR) method to be used as a further constraint on the S-wave velocity profile. It is generally observed that the HVSR method allows the use of lower frequencies than does the SPAC method alone (see, e.g., Hayashi et al., 2011; Asten et al., 2014; Asten and Hayashi, 2018). Formal combined inversion of DC data and HVSR data is now in common use (e.g., Ikeda et al., 2013; Cheng et al., 2016; Xu et al., 2017). This may be considered as a time-domain method analogous to the frequency-domain method the name krSPAC as described by Asten et al. (2013, 2015).

We begin with the multiple-mode SPAC (MMSPAC) method described by Asten et al. (2004) and Asten (2006a). The method builds on the basic spatial autocorrelation (SPAC) method described by Aki (1957) and Okada (2003), but it differs from most other implementations (such as Picozzi and Albarelli, 2007; Di Giulio et al., 2012) by making use of direct fitting of averaged interstation coherency spectra for layered-earth models (also used by Wathelet et al., 2005), rather than the more common approach of extracting phase-velocity DCs followed by inversion of the dispersion data for a layered-earth model.

The MMSPAC direct-fitting method has been applied in urban areas (Roberts and Asten, 2007; Stephenson et al., 2009), suburban areas (Boore and Asten, 2008; Clapood and Asten, 2009; Asten et al., 2013), and remote areas (Schramm et al., 2012; Smith et al., 2013). It has particular advantages in allowing the use of higher frequencies than can generally be accessed when using the more common approach of inversion of the phase velocity DC to a layered-earth model of S-wave velocities (Asten, 2006b). The algorithm used in the earlier references is augmented here by modeling the “effective Rayleigh mode,” which is the summation of multiple modes based on the theoretical energy partition between modes resulting from ideal vertical impacts on a layered earth (Arai and Tokimatsu, 2004, 2005; Ikeda et al., 2012, 2013).

The SPAC method depends for its effectiveness on azimuthal averaging of interstation coherencies either by an azimuthal distribution of wave-propagation directions or by the use of multiple pairs of stations distributed in an azimuth in an array, or both. Thus, an array of seismometers arranged in one or more equilateral triangles is a common approach. Asten et al. (2014) give examples of how this may be achieved in areas of limited access. Bettig et al. (2001) develop an extended methodology called modified SPAC (MSPAC) allowing azimuthal averaging using arrays containing some departures from symmetry, a method implemented in the public domain geopsy software (Geopsy Team, 2019).

In this study, we extend the direct-fitting MMSPAC method to allow its use with arbitrarily asymmetric arrays. We have given the method the name krSPAC as described by Asten et al. (2013, 2015). We show that the krSPAC formulation allows SPAC processing to be applied in a precise sense to asymmetric arrays thus releasing the interpreter (and the field practitioner) from the need to use averaging or integration approximations, such as that used by Bettig et al. (2001), on interstation distances within an array.

AZIMUTHAL AVERAGING WITH IRREGULAR STATION SPACING USING krSPAC

The concept of azimuthal averaging of coherencies for a plane wave passing multiple pairs of stations can be expressed, following Aki (1957), Okada (2003), and Asten (2006a), in the form

$$\rho(f) = \int e^{ikr \cos \theta} d\theta = J_0(kr),$$

(1)

where \(\rho(f)\) is the coherence spectrum, \(f\) is the frequency, \(k = 2\pi f / C(f)\) is the wavenumber, \(C(f)\) is the phase velocity DC, \(r\) is a constant station separation, \(\theta\) is the azimuthal angle of a station pair relative to the plane-wave vector, and \(J_0\) is the Bessel function of zero order. When performing interpretation via direct fitting of SPAC spectra, we use equation 1 with a known value of \(r\) and a forward model DC \(C(f)\) as obtained by using the routine sdisp96 from Herrmann (2013). Thus, we compute a model SPAC spectra that are fitted to the observed azimuthally averaged SPAC spectrum (Figure 1). The direct fitting is achieved by iterative forward modeling; the quality of the fit over a given bandwidth is measured objectively by the standard deviation of the fit, and variables in a layered earth model (S-wave velocity \(V_s\) and thickness \(h\)) are varied to achieve a best fit via a least-squares criterion. Further details and examples are given by Asten et al. (2014) and Asten and Hayashi (2017, 2018).

For an array of four stations in a centered equilateral triangle (such as triangle ABC of Figure 2), the azimuthally averaged observed SPAC spectrum becomes
\[ \rho(f) = \frac{\rho_1(f) + \rho_2(f) + \rho_3(f)}{3}, \]

(2)

where \(\rho_i\) are the coherency spectra for individual pairs of stations. If the wavefield is omnidirectional, each of the three terms in equation 3 approximates a Bessel function and equation 2 may be written as

\[ \rho(kr) = \frac{\rho_1(kr_1) + \rho_2(kr_2) + \rho_3(kr_3)}{3}, \]

(3)

\[ \approx [J_0(kr_1) + J_0(kr_2) + J_0(kr_3)]/3, \]

(4)

where \(r_i\) is equal and can be either the radius or the side length of the triangle. If the triangle is slightly irregular in shape, then the three \(J_0\) terms are no longer additive at higher wavenumbers and the shape of the averaged coherency curve cannot be solved for the wavenumber and phase velocity. Figure 1 shows an example in which the minimum to maximum station separations vary by 30% (a common scenario when using the MSPAC approach).

For an assumed DC \(C(f)\) and a single average \(r\)-value for an array, we can use equation 3 to sum observed SPAC spectra for multiple pairs of stations (the MMSPAC approach). For an asymmetric array, the \(r_i\) is not equal but equation 3 remains valid as a method to sum the observed SPAC spectra. As a detail of implementation, the abscissa for the three complex spectra represented in equation 3 have different sampling due to the different values of \(r_i\), and it is necessary to resample the spectra using a cubic spline to perform the addition. The averaged observed spectrum \(\rho(kr)\) changes with each iteration involving an updated model dispersion, but this presents no problem in implementing the iterative inversion process previously described.

**MODELING WITH THE RAYLEIGH EFFECTIVE MODE \(R_e\)**

Although very useful results can be obtained from observations of microtremor noise by treating surface-wave energy as being solely of the fundamental mode, multiple studies have demonstrated the presence of multiple modes; see the examples in Asten and Hayashi (2018) and an extensive discussion in Appendix 5 of Foti et al. (2018). The partition of energy between multiple modes creates additional variables for interpretation, and one simple assumption in many cases reduces that uncertainty — that is, the definition of an “effective mode” \(R_e\). The assumption is that the energy partition between modes follows the theoretical partition, which applies for a vertical-impact point source on a layered earth, and we use a formulation from Tokimatsu et al. (1992) and Ikeda et al. (2012). If a layered model has one or more boundaries with a high acoustic impedance contrast, the DC for \(R_e\) typically lies close to the first or second higher mode for about an octave of frequencies (Foti et al., 2018).

Asten and Hayashi (2018) show examples of the MMSPAC method incorporating the effective mode \(R_e\), and it is apparent that although some sites demonstrate adherence to the “point-source” energy partition assumption, others do not. We expect that a range of cultural sources of microtremor energy will couple to the earth with nonvertical forces, and that some sites will be nonlayered thus providing a geologic departure from the \(R_e\) assumptions. Where a large number of geophones or seismometers is available, arrays of seven (preferably 10–15) sensors allow the use of frequency-wavenumber processing such as that described by Poggi et al. (2017).
Figure 4 graphically shows the perturbations applied to Figure 2d. The horizontal axis is in units of $kr$ (dimensionless). The thick horizontal black line shows the same frequency band calculated for standard deviation, converted to the $kr$ scale. Annotations show the position of 1, 3, 5, 8, and 10 Hz on the $kr$ scale, for $r = 146$ m. Colors and symbols as for Figure 2c.

Figure 3c shows the four-station equilateral triangle array used as part of our analysis. We continue to recognize that the $R_c$ concept has limitations as an approximation and it may not be applicable to the energy partition between modes at any given site.

**krSPAC COMPARED WITH MMSPAC FROM SYMMETRIC ARRAYS**

We first apply krSPAC to synthetic data generated for blind tests of microtremor array methods held in conjunction with the European Seismological Commission 2006 Symposium (Cornou et al., 2006). We select synthetic data set N104 because it has a shallow (25 m) and a deep (275 m) interface showing a strong $V_S$ contrast (Figure 2a), and we select a triangular array of radius 179 m from the data set. Synthetic microtremor data were generated using sources having random force orientation and amplitude (Moczo and Kristek, 2002), with the associated wavefield computed using the wavenumber-based method of Hisada (1994, 1995). Synthetic sources were randomly distributed over an approximately circular area of radius 500 m, thus including sources within, near, and far from the modeled array (Cornou et al., 2006, pp. 1131 and 1172–1173).

Figure 2 shows the layered earth model used to generate the synthetic data and the computed phase velocity Rayleigh modes. Figure 2 also shows the four-station equilateral triangle array used for this conventional MMSPAC processing and model fitting. It is evident that the useful fitting range of frequencies for (synthetic) observed and (known) model SPAC spectra is a wide band from 1 to 10 Hz, with fits extending to the 6th and 10th maxima for the station separations of 179 and 310 m, respectively.

Sensitivity studies were conducted by separately perturbing the thickness of layer 1 and layer 2 of the source model (Figure 2a), using the method described by Asten and Hayashi (2018). These sensitivities show that the frequency range 1–2 Hz defines the lower $V_S$ contrast (depth 275 m), whereas the frequency range 3–5 Hz defines the upper interface (depth 25 m). It is obvious that the frequency band 2–4 Hz is strongly influenced by higher modes and, hence, meaningful fitting of a model requires use of the Rayleigh-
The very close correspondence between the krSPAC representation of synthetic data and an effective-mode layered-earth fitted model provides validation of the krSPAC method developed here, where assumptions inherent in the formulation are valid. We note that although the krSPAC formulation presented and demonstrated here uses averaging over only three pairs of stations, there is no reason why the algorithm cannot be applied to an arbitrary number of pairs of variable spacing. For example, the three radial pairs and the three side-length pairs of the triangle could be combined, but we have chosen to keep the two groups separate to be able to compare fitted krSPAC spectra on two complementary plots. We now consider examples of the method applied to asymmetric arrays of real field data. A separate paper (Stephenson et al., 2019) demonstrates the krSPAC methodology applied to the twin challenges of a very deep Quaternary basin (to depths of order 2 km), and where cultural and access limitations caused arrays to become grossly asymmetric, with side lengths differing by a factor of two or more.

krSPAC COMPARED WITH MMSPAC FROM SYMMETRIC ARRAYS — FIELD EXAMPLE

The krSPAC method is tested on data from the Saratoga test site, San Jose, California, where borehole V$_S$ logs to 300 m depth are available and a range of active and passive seismic methods have been applied (see Figure 17 of Boore and Asten, 2008). Figure 5 shows how a useful curve fitting is possible in this example over the frequency band of 1–4 Hz, but the fit above 4 Hz is meaningless (the gray-shaded areas of the plots in Figure 3). However, when the observed and model data are plotted versus $kr$ (where the coherency curve for each pair of stations is transformed to the $kr$-space before azimuthal averaging), the fit of the observed and model SPAC curves extends to 8 or 10 Hz (Figure 3c and 3d). Thus, the krSPAC approach delivers an additional octave of useful frequencies for the chosen array, which approximates to a factor of five in usable (shorter) wavelengths.

The very close correspondence between the krSPAC representation of synthetic data and an effective-mode layered-earth fitted model provides validation of the krSPAC method developed here, where assumptions inherent in the formulation are valid. We note that although the krSPAC formulation presented and demonstrated here uses averaging over only three pairs of stations, there is no reason why the algorithm cannot be applied to an arbitrary number of pairs of variable spacing. For example, the three radial pairs and the three side-length pairs of the triangle could be combined, but we have chosen to keep the two groups separate to be able to compare fitted krSPAC spectra on two complementary plots. We now consider examples of the method applied to asymmetric arrays of real field data. A separate paper (Stephenson et al., 2019) demonstrates the krSPAC methodology applied to the twin challenges of a very deep Quaternary basin (to depths of order 2 km), and where cultural and access limitations caused arrays to become grossly asymmetric, with side lengths differing by a factor of two or more.

krSPAC COMPARED WITH MMSPAC FROM SYMMETRIC ARRAYS — FIELD EXAMPLE

The krSPAC method is tested on data from the Saratoga test site, San Jose, California, where borehole V$_S$ logs to 300 m depth are available and a range of active and passive seismic methods have been applied (see Figure 17 of Boore and Asten, 2008). Figure 5 shows how a useful curve fitting is possible in this example over the frequency band of 1–4 Hz, but the fit above 4 Hz is meaningless (the gray-shaded areas of the plots in Figure 3). However, when the observed and model data are plotted versus $kr$ (where the coherency curve for each pair of stations is transformed to the $kr$-space before azimuthal averaging), the fit of the observed and model SPAC curves extends to 8 or 10 Hz (Figure 3c and 3d). Thus, the krSPAC approach delivers an additional octave of useful frequencies for the chosen array, which approximates to a factor of five in usable (shorter) wavelengths.

We provide a quantitative example of the reduction of uncertainty when the krSPAC method is used instead of the MMSPAC method on the asymmetric array of Figure 2c. Figure 4 shows how a perturbation method may be applied to a chosen parameter of the layered earth model, with the resulting change in standard deviation $\sigma(R_s)$ of the fit between model and (synthetic) field data providing a measure of the resolution of that parameter. We make a subjective choice of an acceptable change in standard deviation then apply perturbations to the upper layer thickness $h_1$ of Figure 2a to determine the best-fit thickness and error bounds allowed by the specified maximum change in standard deviation.

The same perturbation process is applied to the observed and model data of the asymmetric array of Figure 3a (SPAC) and 3c (krSPAC), with the results listed in Table 1. For this example, the uncertainty bounds are closely similar for the conventional MMSPAC interpretation using the large symmetric triangle of Figure 2c compared with the krSPAC interpretation using the asymmetric triangle. However, the uncertainty more than doubles when using MMSPAC with the asymmetric triangle.

The very close correspondence between the krSPAC representation of synthetic data and an effective-mode layered-earth fitted model provides validation of the krSPAC method developed here, where assumptions inherent in the formulation are valid. We note that although the krSPAC formulation presented and demonstrated here uses averaging over only three pairs of stations, there is no reason why the algorithm cannot be applied to an arbitrary number of pairs of variable spacing. For example, the three radial pairs and the three side-length pairs of the triangle could be combined, but we have chosen to keep the two groups separate to be able to compare fitted krSPAC spectra on two complementary plots. We now consider examples of the method applied to asymmetric arrays of real field data. A separate paper (Stephenson et al., 2019) demonstrates the krSPAC methodology applied to the twin challenges of a very deep Quaternary basin (to depths of order 2 km), and where cultural and access limitations caused arrays to become grossly asymmetric, with side lengths differing by a factor of two or more.
shows the layout of a pair of triangular arrays and an asymmetric triangle used to test the krSPAC method. Figure 6 shows the best fit layered-earth model and the $V_S$ log; note that the model differs to a small degree from that shown in Boore and Asten (2008) due to the recent use of the effective mode $R_e$ in SPAC processing at this site.

The plots in Figure 7a and 7b show an example of MMSPAC interpretation using the large equilateral triangle of Figure 5. Figure 7c and 7d shows the fit achieved if using MMSPAC processing on the asymmetrical triangle of Figure 5. The loss of high-frequency information (shown as the shaded areas of the plots) compared with that of the symmetric triangle is obvious. Figure 7e and 7f shows the equivalent plots for the asymmetrical triangle when using krSPAC. It is evident that the krSPAC averaging preserves information in the higher order maxima and minima of the $J_0$ function, even though the interstation distances differ by up to 30% from the mean.

Plots of individual coherency spectra for the radii of the asymmetrical triangle DBC (Figure 5) are shown in Figure 8. These plots further illustrate the utility of transformation from frequency to $kr$ space when performing the spatial average on data from an asymmetrical triangle.

**krSPAC APPLIED TO THE PLEASANTON ARRAY**

Ambient noise has been recorded on an asymmetric seismic array with a diameter of approximately 1 km, as installed in suburban Pleasanton, California (Figure 9a). These data are used in a direct comparison of interpretation by MSPAC implemented in geopsy software and by krSPAC (Figures 9b and 10, respectively). For this example, the modeling is limited to the $R_0$ mode only, for two reasons. First, the MSPAC models provided used only the $R_0$ mode, and second, the direct-fitting krSPAC shows anomalous behavior in...
the frequency band 0.7–1.4 Hz, which (unlike other sites discussed here) cannot readily be fit using the effective-mode algorithm; possible reasons for such anomalous behavior are beyond the scope of this paper.

Figure 9b shows that with MSPAC, the maximum useful frequency was approximately 1 Hz, and the interpretation yielded a three-layer model with the top layer being 150 m thick and undifferentiated (Figure 11c and 11d). The useful upper frequency using direct fitting of frequency SPAC shows an increase of about one octave (Figure 10a and 10b). However, when using krSPAC, we see that useful frequencies extend upward by an additional octave (Figure 10c and 10d). Thus, the useful frequency range for krSPAC
is 4x larger than for MSPAC, and the corresponding useful Rayleigh wavelengths are shorter by a factor of approximately 10, giving resolution of $V_S$ for the upper 10 m of soil.

Figure 11 shows a comparison of phase-velocity DCs and $V_S$ versus depth profiles, for the two models derived at the Pleasanton site from MSPAC and krSPAC analysis under the assumption of $R_0$ mode propagation. The resolved depth of the velocity contrasts at depths 150 and 1500 m is similar for the two methods, but the additional resolution of layering in the upper 200 m of the model is obvious.

Averaged $V_S$ for the upper 30 and 100 m is given in Table 2; it is evident that the information accessible from higher frequencies by krSPAC has a major influence on estimation of soil conditions, and hence on the assessment of earthquake hazard, for the upper 100 m of the Pleasanton site. Here, the difference in the 30 m average of $V_S$ is more than 40%, and the difference in the 100 m average is more than 20%. In terms of standard site classifications based on $V_S$ 30 (BSSC, 2001), the site class is D with krSPAC versus C with MSPAC, which illustrates the importance of gaining the best possible accuracy in these interpretations.

### Table 2. Average velocity $V_S$ for upper layers for Pleasanton, using the best-fit layered-earth model from each of the two processing methods.

<table>
<thead>
<tr>
<th>$V_S$</th>
<th>MSPAC</th>
<th>krSPAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 m</td>
<td>422 m/s</td>
<td>267 m/s</td>
</tr>
<tr>
<td>100 m</td>
<td>422 m/s</td>
<td>342 m/s</td>
</tr>
</tbody>
</table>

Figure 11. (a and b) Computed phase velocity DCs $V_S$ versus depth for the krSPAC model and the geopsy model, respectively. Colors as for Figure 2c. (c and d) Interpretations of Pleasanton data for (left) deep layers and (right) shallow layers. Thick red line: model obtained using krSPAC on the central triangle of Figure 9a. Thin red line: model obtained using MSPAC on a seven-station array. Arrowed interfaces are resolved by frequency bands approximately centered on the frequencies marked.
CONCLUSION

The krSPAC method, using direct fitting of coherency spectra spatially averaged in the wavenumber $k_r$ domain, allows inversion of microtremor Rayleigh-wave data from asymmetric arrays over a wider frequency band than alternative methods, with $k_r$ (unitless) values in the range 25–60 achieved in examples described here, where station separations differing by up to 30% are included in the spatial average. In the examples presented here, we have demonstrated that with large station spacings ranging from 400 to 900 m, this method permits resolution of $V_S$ within the upper 10 m of near-surface sediments, and down to a depth on the order of 2 km. The method proves robust when compared with alternative SPAC methodologies using symmetric and asymmetric arrays on a sample of synthetic data and field data from the Santa Clara Valley, California, and Pleasanton, California, where station spacings are irregular on triangular arrays, which vary from 70 to 800 m side lengths.

The krSPAC method will have applications in which logistical or noise constraints force the use of modified asymmetrical array geometries at the observation sites. Use of the Rayleigh effective-mode concept allows it to be applied where multiple modes exist with a power partition approximating the theoretical partition expected for vertical point sources. However, we expect the averaging process produced by addition of multiple coherency spectra will, as with all SPAC methods, rely on some level of spatial (azimuthal) distribution of sources for its success.

ACKNOWLEDGMENTS

M. Asten was assisted by a grant from the U.S. Geological Survey Earthquake Hazards Program (USGS EHP) 05HQGR0020 and a USGS collaborative research travel grant. R. Sell, C. Dietel, M. Meremonte, and A. Leeds assisted in acquisition of the Saratoga and Pleasanton data through USGS EHP funding. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. government. The paper benefited from thoughtful insights provided by M. Moschetti and three anonymous Geophysics reviewers.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

REFERENCES


