Site Effects Assessment Using Ambient Excitations

SESAME

European Commission – Research General Directorate
Project No. EVG1-CT-2000-00026 SESAME

Report on

FK/SPAC Capabilities and Limitations

University of Potsdam, Germany

WP06

Derivation of dispersion curves

Deliverable D19.06

January 2005

List of Contents

Summary 3
Introduction 4
Tools for ambient vibration array analysis 5
Frequency wave-number methods 5
Spatial autocorrelation method (SPAC) and modification 6
Determination and visualization of dispersion curves 7
Simulating ambient vibration data sets 8
Velocity models of virtual test sites 9
Source distributions and virtual array configurations 10
Evaluating phase velocity uncertainties 12
Instrumentation effects on dispersion curve estimates 12
Random arrival time delays and related uncertainties 13
Curved wavefronts and related uncertainties 15
Array geometry, resolution and aliasing 17
Evaluation of the reliability of phase velocity determination using simulated datasets 21
Virtual test site Pulheim – comparison of methods 21
Test site Pulheim – comparison between simulated and real datasets 26
CVFK and CAPON method performance 26
Modified spatial autocorrelation method (MSPAC) 28
Virtual test site Liege – array size 31
Virtual test site Liege – attenuation structure 33
Discussion and conclusions 36
Measurement strategy for ambient vibration array analysis 39
Acknowledgements 41
References: 41

List of Contributors (in alphabetical order after the last name)

Sylvette Bonnefoy-Claudet[1]	UJF, Grenoble, France
Cecile Cornou[2]	ETHZ, Zurich, Switzerland
Giuseppe Di Giuglio	INGV, Rome, Italy
Bertrand Guillier	IRD, Grenoble, France
Denis Jongmans	LIRIGM, Grenoble, France
Andreas Köhler	IGUP, Potsdam, Germany
Matthias Ohrnberger	IGUP, Potsdam, Germany
Alexandros Savvaidis	ITSAK, Thessaloniki, Greece
Daniel Roten	ETHZ, Zurich, Switzerland
Frank Scherbaum	IGUP, Potsdam, Germany
Estelle Schissele[3]	IGUP, Potsdam, Germany
Daniel Vollmer	IGUP, Potsdam, Germany
Marc Wathelet	University of Liege
Project coordinator:	Pierre-Yves Bard, UJF, Grenoble, France
Task B, WP05/06/07 Leader:	Frank Scherbaum, IGUP, Potsdam, Germany

[1] now at xxx, Bratislava, Slovak Republic.
[2] now at UJF, Grenoble, France.
[3] now at Laboratoire de Detection et de Geophysique (LDG), Bruyeres-le-Chatel, France.

Summary

In the following we report on the results obtained within the framework of the SESAME Project (Site Effects Assessment Using Ambient Excitations, EC-RGD, Project No. EVG1-CT-2000-00026 SESAME), Task B (array measurement technique), Work Package 06 (WP06 – Derivation of dispersion curves).

Within the context of this work package we tested the capabilities and limitations of a variety of array analysis methods for the determination of surface wave dispersion characteristics from ambient vibration array measurements. For this task, a new integrated software package (**CAP** – **C**ontinuous **A**rray **P**rocessing) was developed for the offline processing of ambient vibration array data (see also D18.06). In particular, the following analysing methods have been integrated into CAP: **a)** semblance based conventional f-k technique CVFK (Kvaerna and Ringdahl, 1986); **b)** conventional f-k algorithm based on cross spectral matrix approach (standard beamformer); **c)** high resolution f-k algorithm after Capon (1969); **d)** multiple signal classification approach (MUSIC) after Schmidt (1981, 1986); **e)** modified spatial autocorrelation method after Bettig et al. (2001) based on original work of Aki (1957).

In order to assess the reliability of dispersion curve estimates, two main procedures have been followed. First, for frequency wavenumber methods (f-k), the array response has been evaluated theoretically for simple, idealized wavefield situations. The obtained results confirm the known limitations of f-k techniques with respect to wavenumber resolution for longer wavelengths and spatial aliasing for shorter wavelengths (e.g. Tokimatsu, 1997). Secondly, we used ambient vibration simulations for 1D-velocity structures to investigate in detail the reliability of dispersion curve estimates in dependence of the array configuration and wavefield situation presented (Bonnefoy-Claudet et al., 2004, Cornou et al., 2004a,b, Ohrnberger et al., 2004b,c).

From the simulation experiments, we find that for a restricted number of sensors: i) it is in general situations not possible, due to the limitations of f-k methods as indicated above, to find an optimal array layout, which allows the estimation of a high quality dispersion curve for the complete frequency band of interest; ii) for random wavefields containing signal arrivals from all directions, f-k techniques overestimate phase velocities for longer wavelengths due to their insufficient resolution capabilities; iii) the autocorrelation technique is advantageous with respect to the inherent resolution limits for longer wavelengths; iv) the autocorrelation technique results may introduce bias for wavefield situations with a dominant source contribution from a single direction.

From these findings we conclude the following:

- The quality of dispersion curves derived from ambient vibration array analysis depend strongly on the characteristics of the observed wavefield.
- f-k and autocorrelation methods complement each other and should therefore be employed in combination to assure high-quality dispersion curve determination.
- A single optimal array layout for ambient vibration array analysis can not be established with a simple rule.
- In order to derive reliable phase velocity curves, it seems therefore most appropriate to construct wavenumber filters by choosing spatial array dimensions for narrow wavelength ranges. An iterative measurement strategy has been developed to accomplish these ideas and has been presented in WP13.

Introduction

One of the main objectives of WP06 was to critically evaluate the capabilities and limitations of common array analysis techniques for the determination of dispersion curves from ambient vibration measurements. Since the passive seismological observation of microtremor wavefields with small-scale mobile arrays presents a cost-effective alternative to competing geophysical or geotechnical measurements (active seismics or borehole drilling and logging), the use of these methods have become wide-spread in recent times (Horike, 1985, 1996, Ishida et al., 1998 , Miyakoshi et al., 1998, Yamanaka, 1998, Yamanaka et al., 1994, Scherbaum et al., 2003). Despite this common utilisation of microtremor array measurements for site effect estimation via the analysis of dispersion properties of surface waves, there is still an ongoing debate about which analysis method performs best, which experimental setup or instruments should be chosen, or what is the reliability of the results.

In order to assess the reliability of array analysis methods, we have used two approaches. First, we constructed idealized wavefronts propagating horizontally across a virtual array configuration and submitted these data our analysis approaches. By doing so, we neglect on purpose specific wave propagation effects caused by the media response of layered velocity structures. This enabled us to investigate the influence of the selected array geometries and specific source situations on the phase velocity estimates. A preliminary report on this type of experiments has been given in D07.05. Secondly, we continued to study simulated data sets, which were computed for simple 1D velocity structures resembling typical site conditions for specific geological environments. These experiments allow the investigation of more realistic scenarios as wave propagation effects related to the layered velocity structures are taken fully into account (compare Figure 1).

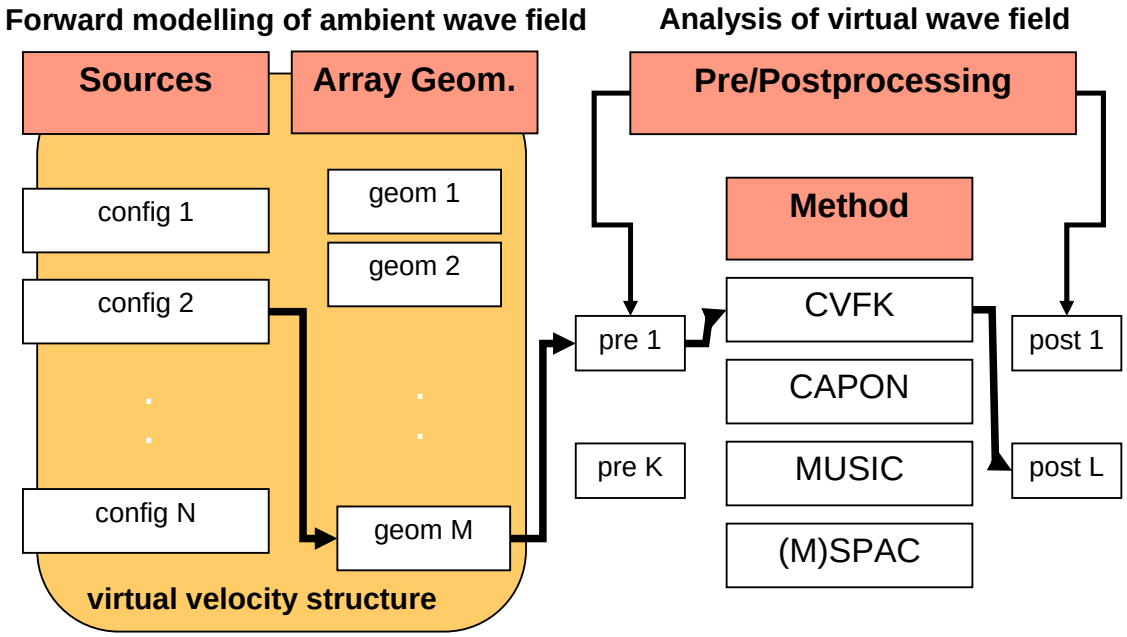


Figure 1: *Sketch of employed strategy for the comparison of array analysis methods within controllable virtual environments using simulated ambient vibration data sets.*

After shortly describing the employed analysis methods and the forward modelling approach of the synthetic data sets, we will report on some important aspects derived from the experiments regarding the performance of individual analysis methods. Finally, we discuss the observations and conclude on the reliability of dispersion curve estimates from ambient vibration array analysis methods.

Tools for ambient vibration array analysis

Frequency wave-number methods

We implemented several frequency wavenumber techniques in order to assess their capabilities and limitations for the use of estimating dispersion curve characteristics from ambient vibration recordings. We distinguish the following methods:

CVFK: a conventional semblance based frequency-wavenumber method after Kvaerna and Ringdahl (1986) evaluated in sliding time window manner and narrow frequency bands around some center frequency. The coherence estimate is given by:

$$RP(\omega,\vec{k}) = \frac{\sum_{l=1}^{L}\left|\sum_{i=1}^{N} X_i(\omega_l)e^{i\omega_l\vec{k}\vec{r}_i}\right|^2}{\sum_{l=1}^{L}\sum_{i=1}^{N}\left|X_i(\omega_l)e^{i\omega_l\vec{k}\vec{r}_i}\right|^2} \qquad \text{Eq. 1}$$

where $X_i(\omega_l)$ are the complex Fourier coefficients of the observed signals at stations i, (i = 1,…,N), at discrete frequencies ω_l, (l = 1,…,L). The phase shifts $e^{i\omega\vec{k}\vec{r}_i}$ account for the delay times related to the horizontal wavenumber $\vec{k}$ from which the direction θ and the horizontal slowness s can be derived as:

$$\theta = \arctan\left(\frac{k_y}{k_x}\right) \text{ and } s = \frac{|\vec{k}|}{\omega}. \qquad \text{Eq. 2}$$

In order to obtain the propagation characteristics of the most coherent plane wave arrival, a grid search over the wavenumber plane is performed. For all presented f-k approaches we used here a wavenumber grid layout sampled equidistantly in slowness and azimuth (azimuthal resolution set to 5 degrees, slowness resolution set to 0.025 s/km).

CVFK2: the conventional frequency wavenumber estimate is based on the evaluation of the cross spectral matrix $R(\omega) = E < X(\omega)X^H(\omega) >$ (CSM), where $E <>$ denotes the expectation value. The CSM is estimated by a block-averaging technique of the observed signal contributions at the array stations $X(\omega) = [X_1(\omega), \ \ldots \ , X_N(\omega)]$ for a given target frequency. The conventional estimator can be written in matrix notation as:

$$P(\omega,\vec{k}) = A(\omega,\vec{k})RA(\omega,\vec{k})^H \qquad \text{Eq. 3}$$

where $A(\omega,\vec{k}) = \left[e^{j\vec{k}\vec{r}_1} \ \ldots \ e^{j\vec{k}\vec{r}_N}\right]^T$ are the steering vectors for wavenumber $\vec{k}$ at frequency ω.

CAPON: the high-resolution frequency wavenumber approach after Capon (1969) is based on the CSM estimate $R(\omega)$. The estimator is constructed to minimize the spectral leakage in the wavenumber domain and is then given by:

$$P(\omega,\vec{k}) = \frac{1}{A(\omega,\vec{k})^H R(\omega)^{-1} A(\omega,\vec{k})} \qquad \text{Eq. 4}$$

This f-k technique is the most widely used within the context of microtremor analysis (Tokimatsu, 1997).

MUSIC: this high-resolution frequency wavenumber method introduced by Schmidt (1986) is based on the decomposition of the cross-spectral matrix into signal and noise subspace. MUSIC is a noise subspace approach and determines estimates of the signal characteristics from:

$$P(\omega,\vec{k}) = \frac{1}{A(\omega,\vec{k})^H E_N(\omega)^H E_N(\omega) A(\omega,\vec{k})} = \frac{1}{\sum_{i=q+1}^{N} \left| A_i(\omega,\vec{k}) \vec{E}_{N_i}(\omega) \right|^2} \qquad \text{Eq. 5}$$

$\vec{E}_{N_i}(\omega)$ are the sorted eigenvectors of the CSM $R(\omega)$ and $q(1 \le q \le N-1)$ is the number of eigenvectors spanning the signal subspace. The parameters of propagation of the q multiple signal arrivals are determined by the minimization of the projection on the noise subspace of the q corresponding wavenumber vectors (Eq. 5). To take full advantage of the high-resolution capabilities of the MUSIC algorithm an optimal separation of the signal and noise subspaces is required. Here, we used a statistical approach, based on an information theory criterion (Akaike, 1973, Wax and Kailath, 1985). Analog to the sliding window analysis scheme of the CVFK algorithm, we evaluated slowness estimates on a statistical basis (MUSIC), or estimated the CSM by block-averaging (MUSIC2). For a detailed overview of these and other f-k estimators, see Zywicki (1999).

Spatial autocorrelation method (SPAC) and modification

Whereas the frequency wavenumber techniques are based on the assumption of the validity of the plane wave signal model, the spatial autocorrelation method (SPAC, Aki, 1957) bases its theoretical foundation on the precondition of a stochastic wavefield which is stationary in both time and space. Aki (1957) showed, that, given this assumption, the existing relation between the spectrum densities in space and time can be used to derive the following expression:

$$\bar{\rho}(r,\omega) = J_0\left(\frac{\omega r}{c(\omega)}\right) \qquad \text{Eq. 6}$$

$\bar{\rho}(r,\omega) = \int_0^{\pi} \rho(r,\omega,\theta) d\theta$ represents the azimuthally averaged spatial autocorrelation $\rho(r,\omega,\theta)$ for station pairs separated by distance r and the inter-station direction θ, and J_0 denotes the Bessel function of the first kind and zero-th order. The above relation allows deriving the single valued phase velocity $c(\omega)$ at a given frequency ω by inversion from observed averaged spatial autocorrelation coefficients. Aki (1957) suggested the use of dense semicircular array deployments to readily obtain these autocorrelation coefficients for various

radii and target frequencies and applied this technique for the analysis of surface wave dispersion characteristics from microtremor recordings.

Bettig et al. (2001) suggested a modification of Aki's original SPAC formula which allows applying the spatial autocorrelation method for less ideal experimental array configurations. The modification concerns the evaluation of the averaged spatial autocorrelation coefficients from station pairs taken from rings of finite thickness $r_1 \le r \le r_2$ instead of using a fixed radius r. The modified formula for the averaged autocorrelation coefficient is:

$$\overline{\rho}(r_1, r_2, \omega) = \frac{2}{r_2^2 - r_1^2} \int_{r_1}^{r_2} r J_0\left(\frac{\omega r}{c(\omega)}\right) dr = \frac{2}{r_2^2 - r_1^2} \frac{c(\omega)}{\omega r} \left[r J_1\left(\frac{\omega r}{c(\omega)}\right)\right]_{r_1}^{r_2} \qquad \text{Eq. 7}$$

Practically, the values for minimal and maximal radii r_1 and r_2 are determined from displaying the co-array configuration for arbitrary array geometries and selecting stations pairs with similar inter-station distances and good azimuthal coverage for the computation of the averaged autocorrelation coefficients.

Determination and visualization of dispersion curves

For the determination of dispersion curves we employ the following schemes for the different f-k approaches. Using a sliding window analysis for the semblance based CVFK in narrow frequency bands, we determine the maximum in each slowness map and record the propagation parameters for the corresponding horizontal slowness vector. This allows for longer time series (several minutes) to evaluate the frequency dependent distribution of propagation parameters. The determination of dispersion curves from these distributions is more appropriate than computing first and second order statistical moments like mean, standard deviation or median and median deviation.

We visualize the histograms obtained from the CVFK analysis as density plots, overlaid by the median and median deviation estimates. For the methods based on the block-averaged CSM (CVFK2, CAPON, MUSIC2), we obtain a single slowness map for the whole analyzed time series. In order to obtain some uncertainty estimate for those methods we record additionally the propagation parameters of the 2% highest power estimates in the slowness maps. The distribution of these values are displayed as density maps overlaid by the maximal f-k estimates. For the determination of dispersion curves, these distributions, viewed as confidence regions, help considerably to judge the reliability of the slowness estimates.

In Figure 2 we show two examples for the visualization of analysis results. For a repeatedly active source at fixed position with respect to the array setting, we show both the slowness-frequency curves evaluated from the CVFK and CAPON analysis (lower panels) as well as the directional estimates (upper panels). We choose to view the dispersion curves proportional to slowness instead of phase velocities, as this allows a linear relation to the measurement errors (time delays). Additional arguments for this way of displaying are given in Boore and Brown (1998) and Brown et al. (2002). The theoretical Rayleigh wave dispersion curves for fundamental and first higher modes are plotted for comparison (black curves). In addition, aliasing curves are plotted for the minimal (red dashed), mean (black dashed) and maximal (green dashed) interstation distance within the array configuration. In this example the wavefield consists of fundamental Rayleigh waves only and deviations from the theoretical

dispersion curves (f<0.6Hz, f~1.1Hz, f>2Hz) can be related to sharp drops in the vertical component spectra.

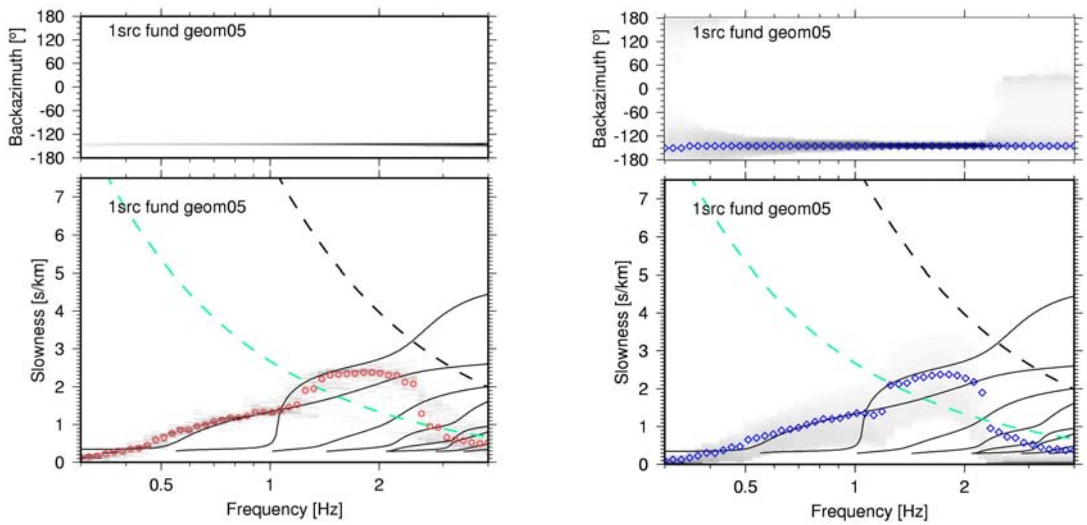


***Figure 2**: Example of visualization of f-k analysis results. Left panels: semblance based sliding window CVFK analysis. Right panels: Capon estimator. Lower panels: slowness estimates (red, blue symbols), theoretical Rayleigh wave dispersion curves (black solid) and aliasing conditions (dashed lines). Underlying density plots show the distribution of estimates. Upper panels: directional estimates. The direction of the fixed source location is correctly estimated within the resolution of the employed wavenumber grid sampling (5 degree azimuthal resolution).*

The derivation of dispersion curves from the analysis result of the spatial autocorrelation method involves the non-linear inversion of Eq. 6 (Eq. 7 for the modified approach, respectively). In order to avoid any difficulties related to the inversion process, we use for the interpretation of analysis results directly the autocorrelation curves. Therefore, for comparison with the results obtained from f-k methods or with the theoretical dispersion curves, we convert the phase velocity curves (or slowness estimates from f-k methods) into the autocorrelation domain which allows a better judgement of the reliability of autocorrelation measurements.

Simulating ambient vibration data sets

In order to allow a consistent comparison of the different array techniques for estimating Rayleigh wave dispersion curves from microtremor recordings, we created several synthetic datasets simulating ambient seismic noise. The composition of ambient seismic noise wavefields is still debated in the scientific community. Different estimates for the energy ratio between body and surfaces waves, the ratio between Love and Rayleigh waves or the ratio between fundamental mode and higher mode surface waves contained in the ambient noise wavefield have been given. Considering the complexity of source excitation of anthropogenic activity and site specific propagation effects for distinct geologic subsurface structures, it is unlikely that a general rule can be found. Thus, as no optimal procedure for realistic ambient noise simulations can be given, we restricted ourselves to simplified assumptions. In particular we used point-source excitations at the earth's surface with impulsive source time functions throughout this study.

In order to account for different wavefield situations, we varied the spatio-temporal source densities and force orientations of the surface sources. We used the modal summation technique (Hermann, 1996) to compute several types of datasets:

- fundamental mode of Rayleigh waves without higher modes
- all modes of Rayleigh waves.
- fundamental modes of Love and Rayleigh waves.
- all modes of Love and Rayleigh waves.

Velocity models of virtual test sites

We mainly distinguished two distinct 1D velocity structures which resembled closely the characteristics of two sites, where field experiments had been undertaken. The first model is related to the geological situation encountered at site Pulheim (PLH), located in the NW of the city area of Cologne in the Lower Rhine Embayment (LRE) in NW-Germany. The situation can be described as a deep sedimentary basin, where the sedimentary cover of tertiary and quaternary ages reach thicknesses of 200 to 300 m overlying the bedrock of Devonian age. The subsurface structure of this region is relatively well known and preceding work has been accomplished for this site (Budny, 1984, Ibs von Seht and Wohlenberg, 1999, Parolai et al., 2002, Scherbaum et al., 2003, Hinzen et al., 2004, Ohrnberger et al., 2004a). For P- and S- wave velocity models, we chose the generic deep basin model (Brüstle and Stange, 1999) of the new German earthquake code and adjusted this model slightly according to the study of Scherbaum et al. (2003) to better resemble the real situation at site PLH. The model comprises two major impedance contrasts, one at 35 m depth and one at 210 m depth. The attenuation structure used for modeling has been taken from the work of Budny (1984), who derived empirical models from downhole measurements in the LRE (compare Figure 3).

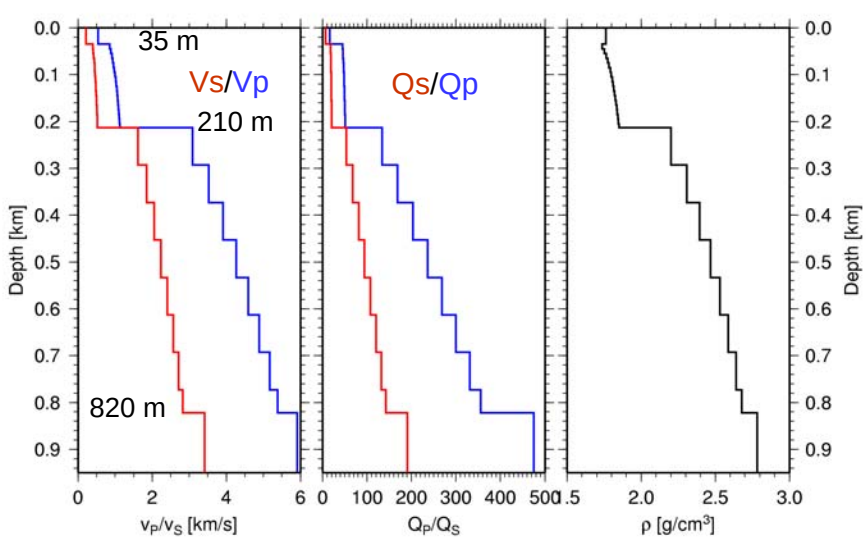


Figure 3 : *1D velocity model used for forward modeling of the virtual test site Pulheim.*

The second site used in our numerical evaluation of dispersion curve determination from synthetic waveform data is related to the geological situation found in Liege. The site is characterised by two thin layers of constant material properties representing the

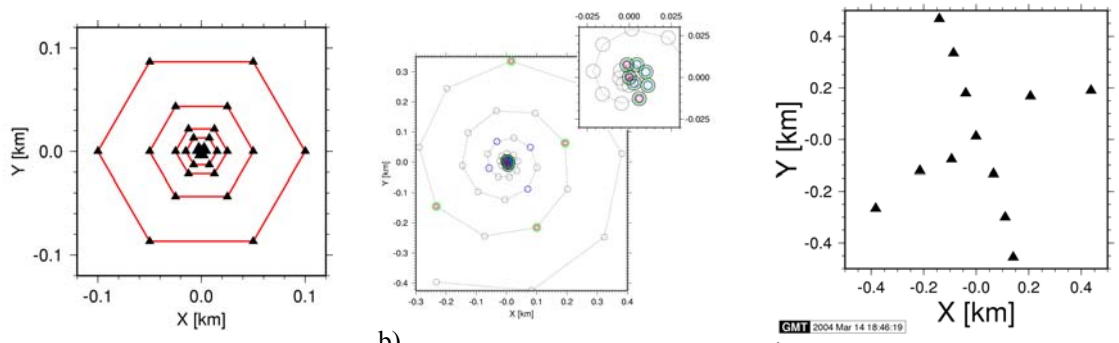
unconsolidated sedimentary cover overlying the bedrock. The site exhibits a very high impedance contrast. We chose this virtual test environment to study specifically the influence of the attenuation structure on the results of dispersion curve analysis, therefore three models with varying attenuation structure have been used to compute simulated ambient vibration time series. The parameters of this 1D-structure are given in Table 1.

Layer	Thickness [m]	V_P [m/s]	V_S [m/s]	ρ [g/cm^3]	Q_P low / mod / high	Q_S low / mod / high
1	7.8	310	193	2.0	**20 / 100 / 250**	**5 / 25 / 100**
2	20.0	1112	694	2.0	**50 / 100 / 250**	**15 / 25 / 100**
HS	--	2961	2086	2.0	**100 / 100 / 500**	**25 / 25 / 200**

Table 1 : *1D velocity model for site Liege. Three models have been comuted using different Q-factors.*

Source distributions and virtual array configurations

The computation of simulated data sets has been performed for a set of receiver configurations and for source distributions ranging from simple and idealized situations to complete random (and more realistic) settings.
The virtual array configurations used for sites PLH and Liege are depicted in Figure 4. For Liege, simple hexagonal array geometries with varying aperture have been used (4m to 200m aperture). At site Pulheim, two configurations were used: one spiral shaped configuration (50 sensors) which allowed taking subsets for the analysis (see Table 2 below) and one large aperture cross shaped array layout, which has been used for comparison with the real field measurement layout.



a) b) c)

Figure 4 : *Virtual array layouts for test sites. a) site Liege: hexagonal array shapes with varying aperture; b) spiral shaped array layout for site Pulheim (colored circles depict different subsets of arrays chosen for analysis); c) 12-element cross shaped array simulating the situation of field experiment.*

The characteristics of the array subsets taken from the spiral configuration for site Pulheim are given in Table 2. The selections have followed the idea to use array configurations which span a large variation of the involved inter-station distances to obtain good resolution and aliasing characteristics for a larger frequency band.

Name of the geometry	Number of sensors	Dmin (m)	dmean (m)	dmax (m)
geom01	7	4.4	11.4	21.5
geom04	11	4.4	184	557
geom05	11	4.4	63	187
geom06	7	7.6	272	557

Table 2 : *Summary of characteristics of sub-arrays selected from the spiral shaped configuration in Figure 4. These arrays have been used for the dispersion curve analysis in section "Virtual test site Pulheim – comparison of methods".*

The source distributions for the wavefield simulations are graphically shown for the site Pulheim in Figure 5. The number of source locations varied from 1 source in 5km distance to the array center ('1src' data set) excited 50 times within a 5 minutes time span, to 450 excitations of 150 source locations in distances from 1 km to 10 km for configurations '2village' and 'randcart' and finally to 6000 excitations of 2000 source locations in distances from 0 to 5 km from the array for the 'randclose' source configuration.

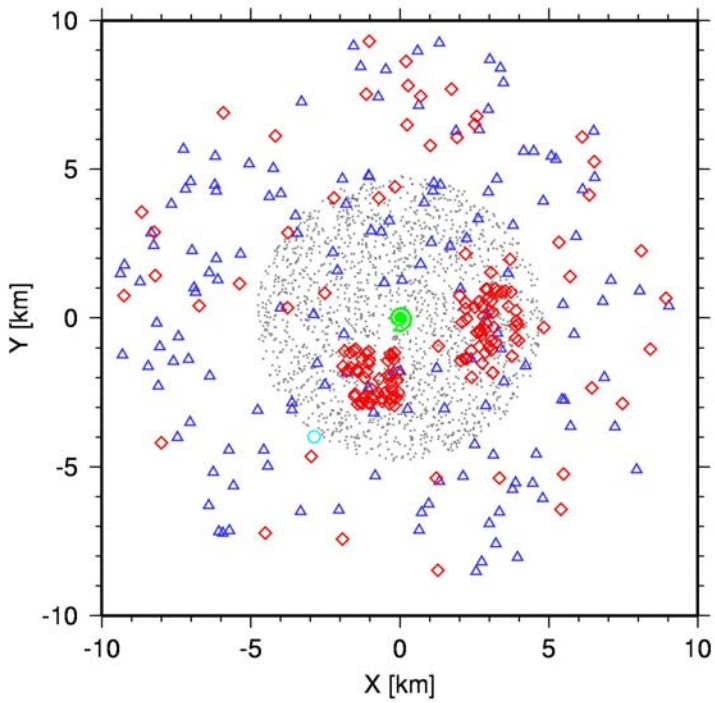


Figure 5: *Spatial distributions of the sources used to calculate the simulated data sets. Site Pulheim: data sets '1src' (turquoise circle), '2village' (red diamonds), 'randcart' (blue triangles) and 'randclose' (black dots) datasets. The green line represents the spiral shaped array setting of Figure 4b. For site Liege only single sources at variable distances and a dense random excitation of sources have been considered (similar to 'randclose').*

The parameters are summarized in Table 3. For site Liege we considered only single source data sets with variable distance to the array center and a full random source distribution. The parameters are summarized in Table 4.

Source configuration	Number of source locations	Number of excitations	Sources-array distances(km)	Modal summation seismograms
1src	1	50	5	16384 pts @ 50 Hz
2village	150	450	1-10	16384 pts @ 50 Hz
randcart	150	450	1-10	16384 pts @ 50 Hz
randclose	2000	6000	0-5	32768 pts @ 50 Hz

Table 3 : *Parameters for source configuration used for wavefield modelling at site Pulheim.*

Source configuration	Number of source locations	Number of excitations	Sources-array distances(km)	Modal summation seismograms
1src-0050 – 1src-4000 (distance in m)	1	1000	0.05, 0.10, 0.250, 0.50, 1.00, 2.00, 4.00	32768 pts @ 50 Hz
random	5000	7500	0-5	32768 pts @ 50 Hz

Table 4 : *Parameters for source configuration used for wavefield modelling at site Liege..*

Evaluating phase velocity uncertainties

Potential sources of error can be attributed to one or a combination of the following three issues: recording, processing and nature of ambient vibration wavefield. Considering the use of common array analysis techniques (f-k or SPAC) for the determination of dispersion curves, the main source of uncertainties and/or bias in the estimation procedure can be associated with relative arrival time errors between any pair of sensors within the array setting. Thus, the occurrence of timing errors may be related to different sources, such as: instrumentation, phase determination, violation of the plane wave signal model. Furthermore, we can distinguish between systematic and random time shifts at individual stations or between stations.

Instrumentation effects on dispersion curve estimates

In a previous report (D07.05), we have evaluated the amount of systematic time shifts related to instrumental effects (sensor phase delays) and time synchronization difficulties. It was shown, that the difference of phase delays between sensors due to small relative calibration inaccuracies can be unexpectedly large, even for a homogeneous set of instruments. Furthermore it was proven, that timing delays of DCF time receivers have always to be taken into account in order to allow for a reliable estimate of phase velocities in field experiments. In D07.05, we have focussed on the quantification of the amount of time delays introduced by these two error sources. For the sensor delay differences this was achieved by a numerical experiment by evaluating the frequency dependence of phase delay differences for a pair of sensors for small perturbations of the individual seismometer constants f_0 and h. The results of these simulations are shown in Figure 6. For details of the numerical experiment, we refer the reader to D07.05.

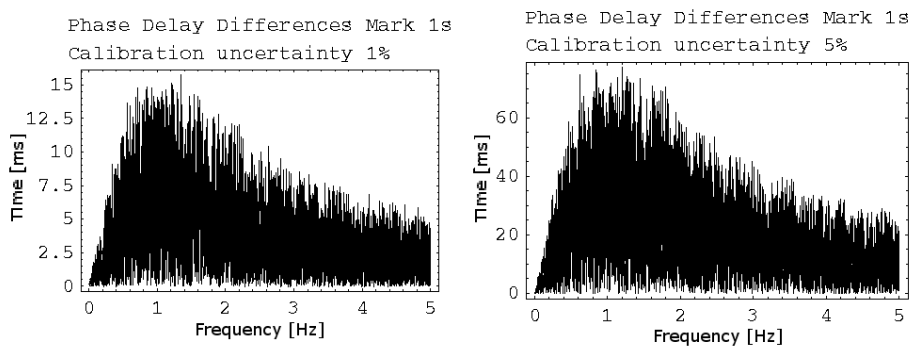


Figure 6 : *Phase delay differences between sensor pairs caused by slight variation of seismometer constants. Left panel: 1% standard deviation in f_0 and h, right panel: 5% standard deviation in f_0 and h.*

The impact of these potential relative (and constant) time delays between sensors on the final phase velocity estimates are difficult to predict, but certainly they are not negligible. The amount of bias introduced depends on the array geometry (inter-station distances involved), the phase delay deviations of each individual sensor and its location within the geometry setup, the number of sensors involved and the direction of the incident wavefield.

In order to keep the bias introduced by instrumental effects to a minimum, we recommend using only well-calibrated instruments and a sufficient number of sensors (at minimum 5) for ambient vibration array measurements. Additionally, it is preferable to use sensors with corner frequencies well below the frequency band of interest at a given site, as the amount of phase delay differences will diminish for frequencies well above the corner frequency of the instruments (compare Figure 6).

Random arrival time delays and related uncertainties

Besides systematic timing errors as described above, we also expect the occurrence of random relative time shifts. Those timing errors will be inevitably caused during the process of estimating the relative travel time differences (usually performed in frequency domain). Indeed these uncertainties occur independent of the applied analysis method (f-k or spatial autocorrelation methods), whereas their magnitudes are related to the signal to noise ratio. We define this in the context of ambient vibration measurements as energy ratio of coherent vs. non-coherent wavefield components. In order to assess the impact of random phase errors at individual stations of the array onto the phase velocity estimates, we used a bootstrapping approach.

For a fixed virtual array setup, we first constructed an ideal plane wave signal by superposing phase shifted sine waves for a target slowness vector at each individual sensor. The forward modelling has been realized in the frequency domain within a narrow frequency band and a small number of discrete Fourier coefficients. This ideal wave front has then been disturbed by adding random phase shifts to the signals at each station. The random time delays were drawn from a normal distribution with zero mean and varying standard deviations. Finally, we analysed the perturbed signal with the conventional frequency-wavenumber algorithm and picked the maximum of the resulting slowness map as estimate of the true slowness vector.

As an example we show in the following the results obtained when considering a circular array geometry consisting of five stations with varying aperture. The bootstrap experiments were performed for different target slownesses and variable source azimuths; however, we fixed the frequency band throughout the analysis to 0.9 Hz to 1.1 Hz. By doing so, the array response function is kept constant for each virtual array setting, and has not to be taken into account for the interpretation of results. The parameters of the random distributions were the same for all individual stations. For all runs we chose zero mean normal distributions with standard deviations ranging from 5 ms to 50 ms. In Figure 7, we depict the results of an individual bootstrap run. As expected, the estimated slowness vectors scatter around the true value. The expectation value of the distribution corresponds to the true propagation characteristics of the undisturbed plane wave front, whereas the scatter can be seen as an uncertainty measure of the estimate.

The overall distribution of the bootstrap samples resembles the shape of the array response pattern. Using a highly symmetric configuration, the distribution of estimates has a nearly circular shape and can be approximately described by a two-dimensional Gaussian.

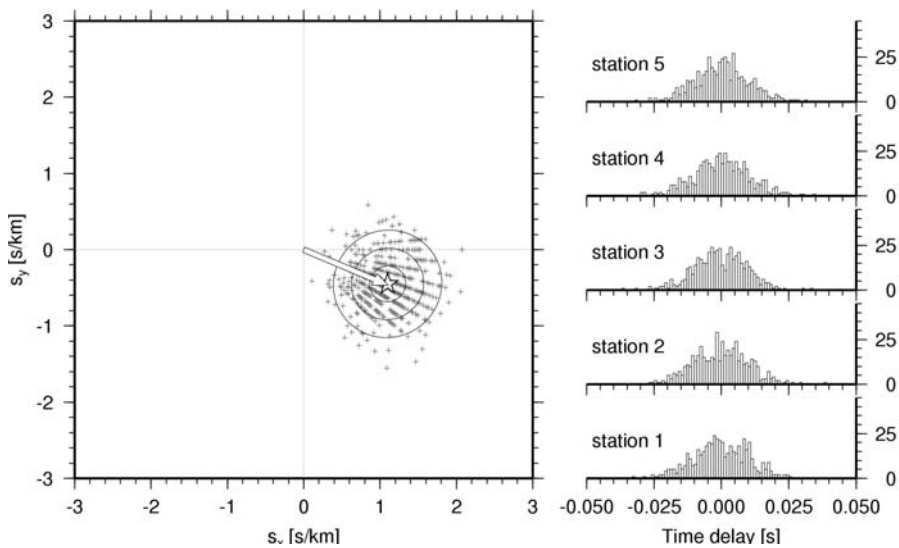
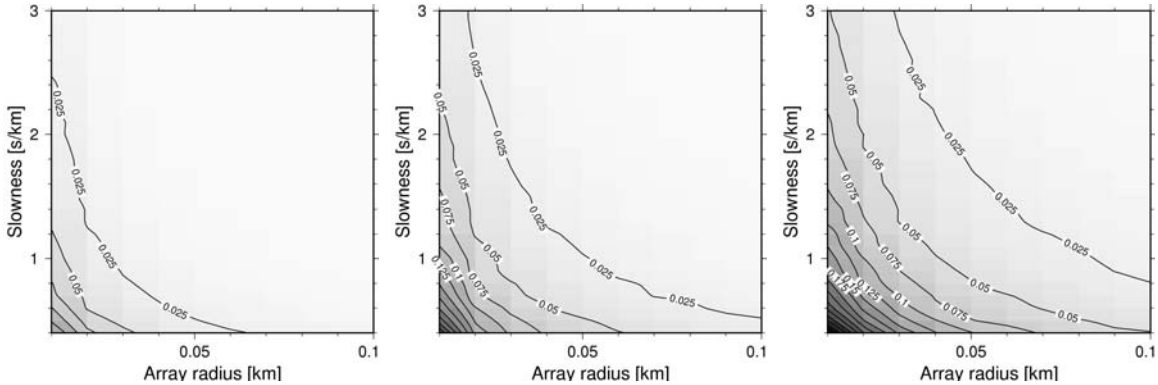


Figure 7: *Scatter plot of estimated slowness vectors (cross symbols) obtained by a semblance based conventional frequency-wavenumber algorithm from 500 perturbed plane wave fronts. The time delays imposed at each of the five stations were drawn from a normal distribution with zero mean and 10 ms standard deviation. The distribution of the samples is shown for each station to the right. Due to the circular array geometry (5 stations, radius 30m) the distribution of slowness estimates resembles closely a two-dimensional Gaussian with nearly equal variances along the principal axis. Error ellipses for 1, 2 and 3 standard deviations are shown as black contour lines. The average slowness vector estimate is given by the white star. The input slowness vector (indicated by the white arrow) has an absolute slowness value of 1.2 s/km.*

In order to summarize the tests, for each bootstrap run with fixed aperture and target slowness we fitted a Gaussian distribution and recorded the average of the standard deviations along the principal axes from these. In this way, we quantify the uncertainty of slowness estimates due to the defined standard deviation of the imposed random time delays and its dependence on array size and slowness of the arriving wave front. In Figure 8, we display the obtained results as shaded contour plots.



a) σ = 15 ms b) σ = 30 ms c) σ = 50 ms

Figure 8 : *Summary of bootstrap tests linking the random time delay error at each station to the resulting uncertainty in slowness estimates. Standard deviations of random time delays were chosen as a) 15ms, b) 30ms, and c) 50ms. Shaded maps depict the value of uncertainties obtained from a bootstrap experiment for a specific array radius and target slowness combination (contour lines at 0.025 s/km intervals).*

From the simulations we can draw the following conclusions: i) the mean of the slowness estimate distribution is an unbiased estimate of the true slowness; ii) the uncertainty of estimates scales directly with the array aperture and the target slowness, in particular the error increases with decreasing array size and faster propagating wave arrivals (smaller slownesses).

It must be noted, that in the simulation we have only considered a single wave arrival. Thus, the evaluated effect of random time delays on uncertainties in the slowness determination seems to be an optimistic outcome compared to real situations (compare also D07.05). Nevertheless, we consider the results as a lower bound of uncertainties which have to be taken into account in any field experiment. From the conclusions drawn above, we can derive some recommendations in order to improve the phase velocity estimates: i) for stationary time series, the amount of phase delay errors introduced from estimating the signal phase in the frequency domain can be reduced by considering the usage of long analysis windows (order 25 to 50 times the center period); ii) for a reliable estimate of the slowness, a large number of analysis windows should be considered allowing for a consistent average phase velocity measure and additionally to provide rough uncertainty estimates.

Curved wavefronts and related uncertainties

In practical terms, the situation investigated above can be considered as the observation of a wavefront arrival showing small undulations (e.g. due to phase estimation errors or small lateral heterogeneities of the site). By assessing the impact on slowness estimates, we were actually investigating the stability of the algorithm to provide correct results for those randomly perturbed wavefronts. In addition, if the sources of ambient vibrations are close to the array, a more deterministic distortion of the wavefront is expected from its curvature. In this case, additional predictable time delays are encountered at individual stations. As all standard frequency-wavenumber analysis algorithms are based on the plane wave model, in this case the estimated slowness will be biased in this case (compare also Almendros et al., 1999).

Again, we tried to quantify the corresponding effects on phase velocity estimates by a numerical experiment. Compared to the previous experiment, only a slight modification of the forward modelling part is needed. Now, the exact curved wave fronts are calculated using the arrival times from an individual source located at a given distance and the backazimuth measured relative to the array centre. No extra random time delays are added, but for each run all source azimuths are considered (5° resolution). The narrowband Fourier spectra of the curved wavefronts are then analysed either using the semblance based conventional frequency wavenumber algorithm or by computing the average autocorrelation coefficient according to Aki (1957). To allow for a better comparison between the results obtained from the CVFK and the SPAC methods, the autocorrelation values are further converted to slowness estimates (Eq. 6) using a simple grid search technique. For each simulation, we then can compare the true slowness (inverse of local propagation velocity of the wavefront) with the estimated one.

In the numerical experiment we used the same pentagon shape array configuration as before for the CVFK analysis, whereas for the SPAC analysis method an 18-station virtual semi-circular array setting was utilised (Figure 9).

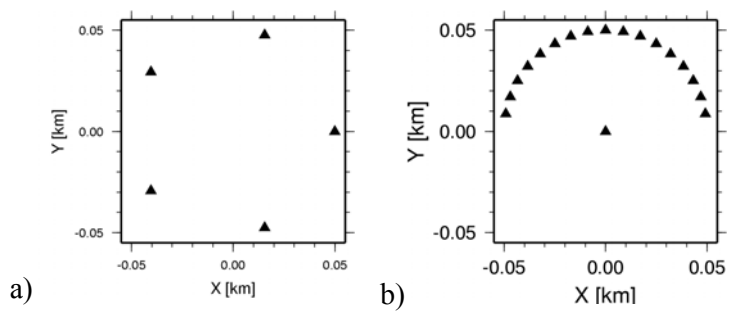


Figure 9 : *Array geometries used in the numerical experiment for assessing the bias introduced by the occurrence of curved wavefronts. a) pentagon shape for CVFK analysis. b) 18-station semicircular shape for SPAC analysis. During the experiment, the array sizes and source distances have been varied.*

In Figure 10 we depict the analysis results for incident curved wavefronts. Each box represents the summary of one test run for a fixed array size and a given source distance but variable azimuths. The true slowness is plotted vs. the estimated slowness value for each individual sample of the experiment (source azimuth). The colour shading of cross symbols has been selected to be proportional to the relative slowness deviation. The ratio between source distance and array radius (d/r) is given in the lower right corner of each box.

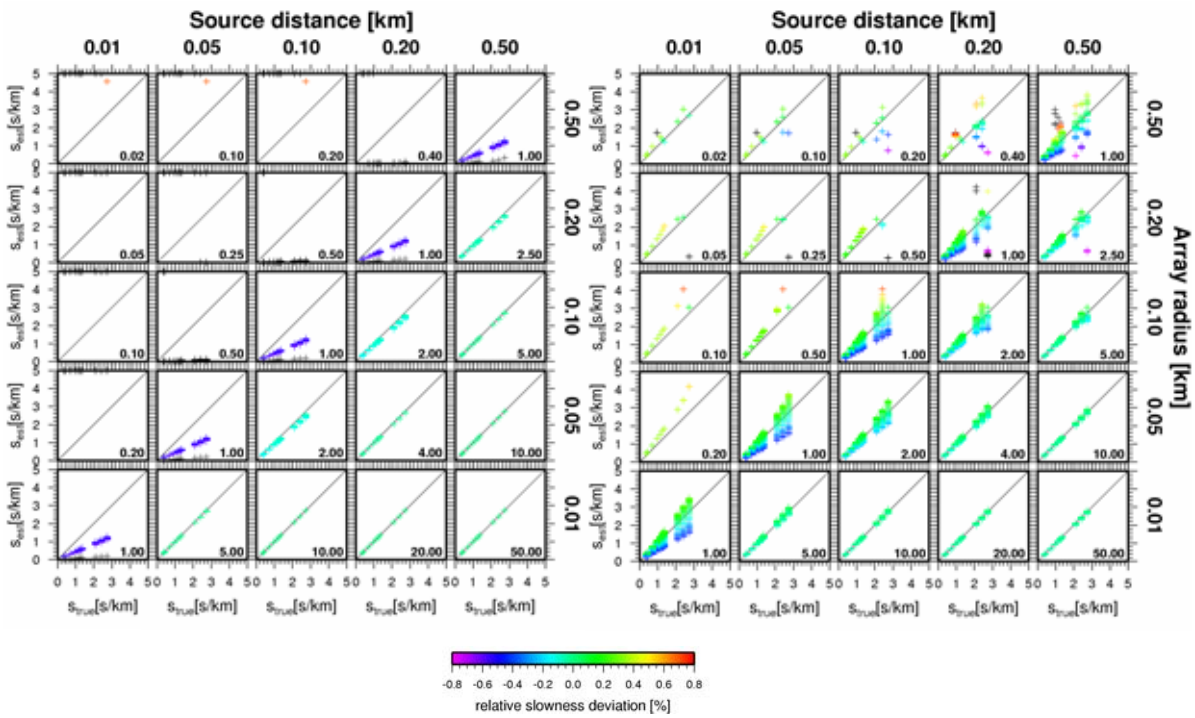


Figure 10 : *Bias of slowness estimates due to the existence of curved wavefronts in the wavefield. In the individual boxes, the true slowness is plotted vs. the estimated slowness value for different array sizes and source distances. Left panel: results of CVFK analysis. Depending on the ratio of source distance to array radius (value given in the lower right corner of each box), a consistent bias to lower slowness values can be observed. Right panel: results from SPAC analysis. Until a source distance to array radius ratio of 1, the SPAC estimates of the slowness show a symmetric scatter around the true slowness value. For details see text.*

From Figure 10 it is clearly recognized that the CVFK results show a significant bias which is introduced by mapping the original curved wavefronts to the underlying plane wave model. The estimated slowness is consistently lower than the true slowness and its magnitude is related to the ratio d/r. For ratios d/r up to 2.5, the deviations are as large as 20%. Ratios below or equal to one (sources inside the array configuration) don't allow any reasonable estimate from the f-k analysis. In comparison, the results of the SPAC analysis show a spread of estimated slowness values in dependence of the azimuth of the incident wavefronts. This spread increases for decreasing values of d/r. However, down to ratios of one, the average of computed slowness values in this experiment still allows a good estimate of the true propagation characteristics. Finally, sources inside the array configuration (d/r < 1) apparently give too high estimates of the true slowness value.

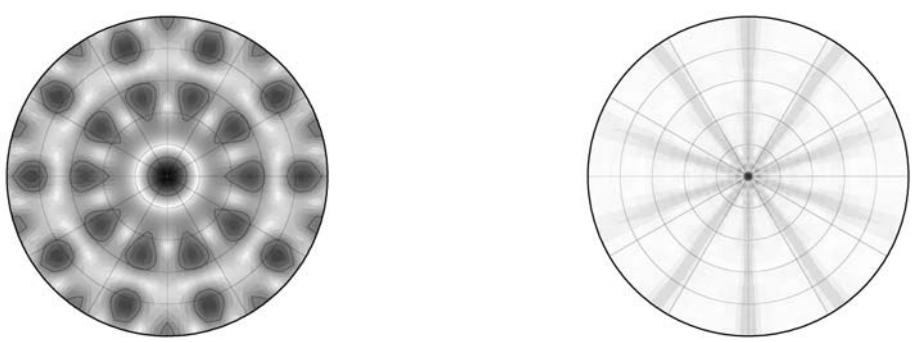
Array geometry, resolution and aliasing

An important issue in any array measurement is the question of optimal geometry (size, shape, number of sensors, etc.) regarding the resolution limits and spatial aliasing effects. This topic has been treated both for large aperture arrays in the context of earthquake/explosion monitoring (Haubrich, 1968, Harjes, 1990) as well as in the field of ambient vibration small aperture array processing (Woods and Lintz, 1973, Asten and Henstridge, 1984, Kind, 2002). The main differences between these different fields of application can be attributed to the signal components of interest which are to be analysed.

Whereas in the field of earthquake detection, arrays are usually optimized for undispersed broadband transient signal arrivals (body wave detection) from relatively distant sources (plane wavefront), for ambient vibration surveys we have to consider a mostly random wavefield caused by nearby surficial sources which has to be analysed within narrow frequency bands. Therefore, within the SESAME project, the matter of optimal array layouts has been re-evaluated focusing on the specific task of determining the dispersion curve characteristics from microtremor recordings.

Traditionally, the performance of a given array geometry is evaluated by the array response pattern (also known as beam pattern) commonly depicted for a plane incident wave from below (see Figure 6). Several (weak) criteria can be given to judge the quality of an array layout from its response. First, for a good resolution of waves arriving at close wavenumbers, the central peak should be "as sharp as possible". Secondly, the aliasing peaks should be "as separated as possible" from the central peak when evaluating the array response for the frequency band for which the analysis will be performed. Finally, the ideal response pattern for an array should show similar suppression capabilities for all directions of arrivals (DOAs).

On basis of these considerations, Capon (1969) derived his well-known high-resolution f-k estimator (HRFK) by optimizing the windowing function such, that the wavenumber response approaches a delta-function in the f-k domain. Capon's estimator has proven to be not only theoretically appealing, but also practically well-suited for array analysis tasks. Compared to conventional f-k algorithms, the gain in slowness resolution is traded against the lack of time resolution caused by the estimation procedure (termed block averaging by Capon, 1969) of the cross spectral matrix (CSM) and the computational complexity involving the inversion of the CSM. Stabilization of the matrix inversion is generally required to avoid numerical difficulties. Unfortunately, the construction of the wavenumber window in Capon's HRFK involves the signal itself. Thus, the high resolution response can not be studied without the knowledge of the signal, and one is left to study the beam pattern instead to derive a lower limit of slowness resolution (Kind, 2002).



a) narrowband: 0.9-1.1 Hz, N = 5, r = 400m **b)** broadband: 0.1-10.0 Hz, N = 5, r = 400m

Figure 11 : *Array response pattern for narrowband and broadband estimates for a pentagon shaped array with radius r = 400m. Maximum slowness has been set to 5 s/km, the radial resolution is 0.025 s/km and the azimuthal resolution 5°. The frequency bands for the evaluation of the beam patterns are given below the plots. The semblance values are depicted as gray scale map with a linear scale from 0 (white) to 1 (black). Black lines depict the -3dB contour relative to the main lobe maximum.*

Due to the aim of investigation, i.e. the determination of frequency dependent phase velocity curves, array methods have to be employed for a narrowband analysis of the ambient

vibration wavefield. Compared to broadband processing, the resolution of narrowband array responses is significantly reduced (Figure 11) and aliasing peaks are fully developed. The improved response behaviour of broadband estimators stems from the superposition of beam patterns for different frequencies. The stacking of those response patterns, which differ with respect to the positions of the corresponding side lobes, leads to a smearing of the power contributions for all slownesses except for the main lobe.

Unfortunately, the use of broader frequency bands is prohibitive for the determination of dispersion characteristics. Thus, an enhancement of the resolution/aliasing capabilities for narrowband analysis can only be achieved by improving the spatial sampling of the wavefield. In turn, this requires the use of a large number of stations in field experiments. Clearly, viewed from the economical perspective, this option is prohibitive in most cases considering the initial equipment cost, the increased logistical effort and additional man power required for field experiments. Given the economical and logistical constraints, there is little one can do with respect to the resolution/aliasing issue for dispersion curves, ***except*** choosing an appropriate array size suitable for the analysis of a narrow wavelength range. *The use of a single array deployment for the determination of well-resolved, non-aliased phase velocity estimates over a large frequency band seems unfeasible under common circumstances.*

The question remains, how to optimize then the array geometry for a narrow target wavelength range. The resolution capability of some given station geometry is determined by the largest inter-station distance effectively seen by the incident wave arrivals from a certain direction as $\lambda^{max} \sim d_{i,j}^{max}$. This condition is a conservative approximation according to the Rayleigh criterion and for conventional beamforming techniques. For Capon's estimator, the resolution limit has been determined empirically as $\lambda^{max} \sim 3d_{i,j}^{max}$ (e.g. Tokimatsu, 1997). Just as well, the aliasing condition for 1D-station configurations with equidistant spatial station intervals can be given as $\lambda^{min}/2 \geq d_{i,j}^{min}$. Whenever no a priori information about the direction of the incoming wavefield is available, the variation of the maximal (minimal) inter-station distance when considering different DOAs should be minimized to obtain an azimuthally constant resolution/aliasing criterion. Geometries showing a good azimuthal sampling are therefore the preferable choice for ambient vibration experiments. Particularly helpful for judging the azimuthal coverage of any array configuration is the co-array concept (Haubrich, 1968), which graphically depicts the distribution of all station pair difference vectors.

The criteria given above for the aliasing limits are approximate values, which strictly just apply for equidistant 1D station deployments. For a 2D-array setting the aliasing condition varies strongly with the direction of the incoming wave arrival, as the effective lattice seen by the wavefield changes significantly. In addition, the number of the shortest inter-station distances $d_{i,j}^{min}$ may be small compared to the total number of larger $d_{i,j}$ involved in a particular observation direction. In order to visualize and quantify the effective aliasing behaviour for a general 2D array deployment, we have taken the following approach. For discrete azimuth values, we computed the projections of inter-station distances $d_{i,j}$ onto the observation direction. The effective lattice was derived by counting the number of $d_{i,j}$ in 10m intervals. The density plot in Figure 12 shows the corresponding azimuthal variation of $d_{i,j}$ distributions. Considering the existence of a random wavefield (uniform distribution of DOAs), we averaged these lattices to obtain the mean effective inter-station density which can be interpreted as the average spatial filter for the wavefield components (gray curve in right panel, Figure 12). Finally, we can derive the percentage of inter-station distances which fails to fulfil the aliasing criterion for a given wavelength by computing the cumulative distribution and normalizing to the total number of inter-station distances involved (black curve in right panel, Figure 12). We called this the cumulative aliasing criteria curve (CACC), which allows

the estimation of a relative amount of observed aliasing features for a particular wavelength. In the example shown in Figure 12 (pentagon shaped array with 100m radius), we determine a critical effective half wavelength $\lambda^{min}/2$ of approx. 90m, for which 50% of the involved inter-station distances show an insufficient spatial sampling of the wavefield.

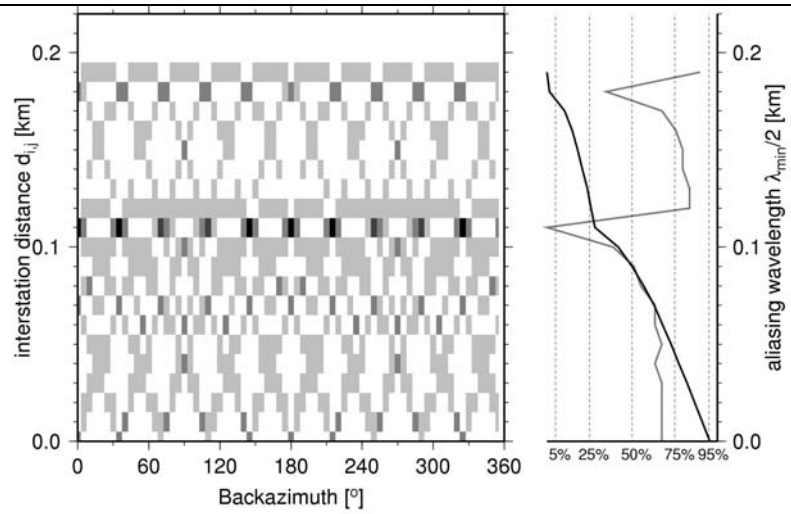


Figure 12 : *Density plot of inter-station distances $d_{i,j}$ for a pentagon shaped array with 100m radius. For different viewing directions (resolution 5°), the length of the projected inter-station difference vectors is binned in 10m intervals. The effective lattice for waves propagating from a single direction will determine the aliasing and resolution capabilities along the viewing direction. The average effective lattice for a large number of wave arrivals from different directions (as expected for ambient vibration wavefields) is given in the right panel as gray curve. The cumulative aliasing criteria curve (black curve, right panel) estimates the percentage of inter-station distances $d_{i,j}$ which fail on average to fulfil the aliasing criterion.*

It is clear, that the requirements on the array geometry setup in terms of resolution and aliasing lead to different conclusions. For high resolution capabilities, particularly at long wavelengths, the array size should be as large as possible. However, in order not to violate the aliasing criterion for even the shortest wavelengths of interest, the inter-station distances should be as small as possible. Some compromise is therefore needed to optimize the array geometry for a given narrow wavelength range.

We suggest the usage of circular deployments with an odd number of stations. These particular array shapes show the best possible azimuthal suppression capabilities for a given number of sensors. Further, the number of smallest inter-station distances is always large and equals the number of deployed sensors. Thus, the aliasing criterion is indeed well approximated by $\lambda^{min}/2 \geq d_{i,j}^{min}$ and the overall aperture D of the array is close to the largest involved inter-station distances d_{max}. As a rule of thumb, one can then derive the wavelength limits of the array configuration from the chosen aperture alone, by using $d_{max} \sim D$ and $d_{min} \sim \pi D/N$, with N being the number of sensor, we get: $\lambda_{max} \sim D > \pi D/N > \lambda_{min}/2$. As an example, we give the approximate wavelength ranges resulting from applying the classic resolution/aliasing limits as specified above to this kind of array geometries with 5, 7 and 9 stations and radius $r \sim D/2$ in Table 5.

N	exact $d_{i,j}^{min} – d_{i,j}^{max}$	λ^{min} - λ^{max} – 3*λ^{max}	approximate $d_{i,j}^{min} – d_{i,j}^{max}$	λ^{min} - λ^{max} – 3*λ^{max}
5	1.17*r – 1.90*r	2.34*r – 1.90*r - 5.70*r	1.26*r – 2*r	2.52*r – 2.*r – 6.*r
7	0.87*r – 1.95*r	1.74*r – 1.95*r - 5.85*r	0.90*r – 2*r	1.80*r – 2.*r – 6.*r
9	0.68*r – 1.97*r	1.36*r – 1.97*r - 5.91*r	0.70*r – 2*r	1.40*r – 2.*r – 6.*r

Table 5 : *Resolution and aliasing limits for circular array layouts with N stations.*

The usage of circular array layouts yields an additional advantage: the spatial autocorrelation curves can be computed in their original form (Aki, 1957, Eq. 6), which is easily recognized by considering the derived co-array configurations. Deviations from the ideal layout due to logistical constraints (feasibility of occupying station positions in urban environments) can be captured by the modified computation scheme suggested by Bettig et al. (2001, Eq. 7).

Theoretically, the SPAC method shows better resolution capabilities than the f-k methods. The resolution limit can be derived from considering the rate of change of the spatial autocorrelation curves. For small values of the argument of the Bessel function ($2\pi fr/c(f) < 0.4$), the sensitivity of the autocorrelation curve to changes of the phase velocity (c(f)) is low compared to the variation of the observed uncertainties. Nevertheless, using 0.4 as lower limit of the argument for the Bessel function, the longest wavelengths which should be resolved with the SPAC method are in the order of $\lambda^{max} \sim$ 15-16*r. The aliasing limitations of the SPAC method are expressed by the oscillating nature of the Bessel function, showing multiple solutions for arguments larger than ~ 3.6. Thus, given a dense azimuthal sampling of the wavefield, the aliasing wavelength is given by approximately $\lambda^{min} \sim$ 1.8*r.

Evaluation of the reliability of phase velocity determination using simulated datasets

Virtual test site Pulheim – comparison of methods

In a first step, we compared the capability of the different frequency wavenumber decompositions (CVFK, Capon, CVFK2 and MUSIC2) to retrieve the dispersion curves for all available datasets. Exemplarily, we present the ‘1src’ and ‘randcart’ source configurations for both the pure fundamental mode simulation and the dataset containing all higher mode surface waves. These datasets have been calculated for the deep sediment velocity model and the results for one fixed station configuration (geom04, see Figure 4 and Table 2) are shown in Figure 13 for the source configuration ‘1src’ and in Figure 14 for simulation dataset ‘randcart’.

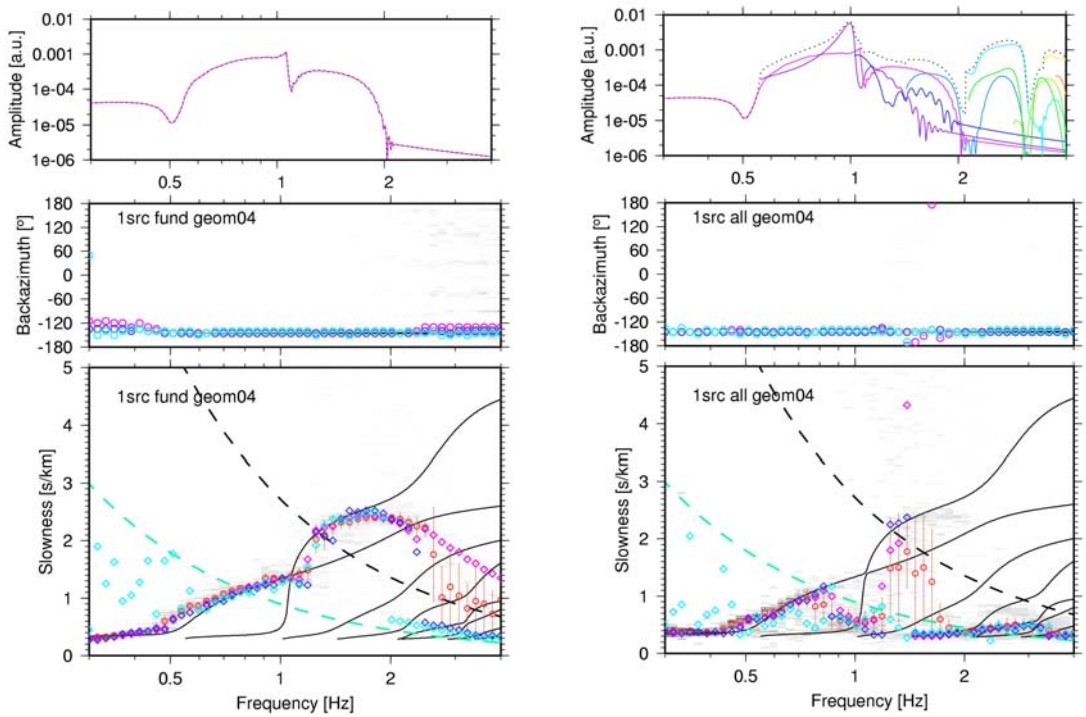


Figure 13: *Results from the frequency wavenumber decomposition using all different methods for the '1src' source configuration including only fundamental mode (left panels) and all modes (right panels). The phase velocity results are depicted as: CVFK: histogram and red dots for median values, Capon: blue diamonds, CVFK2: pink diamonds and MUSIC2: turquoise diamonds. The upper left panel represents the spectra of the fundamental mode (distance of propagation equals 5km) and the upper right panel represents the spectral contribution of the different modes (from pink to orange). The black dotted curve gives the sum of all contributions.*

a) Case '1src' – fundamental mode

All methods perform well in the frequency band [0.6Hz-1.8Hz] and the frequency evolution of the slowness values follows the theoretical dispersion curve of the fundamental mode. A stable estimate of -145° is obtained for the directions of propagation which corresponds to the direction of the true source signal. Nevertheless, close to 1Hz, all methods show some scatter coinciding with the sharp drop in energy of the fundamental mode as observed in the spectra plotted in the upper left panel of Figure 13. Additionally, there is a mismatch between the slowness values and the theoretical curve outside the [0.6Hz-1.8Hz] frequency band, where low spectral energy levels are observed. We think that these small deviations are associated with numerical noise in the forward calculation of the simulated wavefield which deteriorates the phase delay estimates below a certain amplitude level.

b) Case '1src' – all modes

The presence of higher modes increases considerably the complexity of the frequency-slowness distribution. In the frequency range [0.6Hz-0.8Hz], the fundamental mode is dominating and CVFK, Capon and CVFK2 give correct estimates of the slowness whereas MUSIC2 underestimates the theoretical values. In the [0.8Hz-1.8Hz] frequency band, the spectra show that 1) the first higher mode overtakes (~0.8Hz), 2) there is a drop in energy of the fundamental and first higher mode and the second higher mode overtake (~1.1Hz), and 3) the fundamental mode returns to be the most energetic (~1.6Hz). In this case, the bimodal distribution obtained for the CVFK describes well the simultaneous

contributions of both fundamental and higher modes. However, the median values chosen to describe the statistics of the distribution are misleading. Opposed to the CVFK2 and MUSIC2 which show in this case a strong bias, CAPON gives correct estimates of the most energetic contribution (fundamental or higher) within each frequency band. For frequencies higher than 2Hz, all methods show that higher modes dominate the wavefield, and as the frequency increases, the slowness values jump from one higher mode to the next.

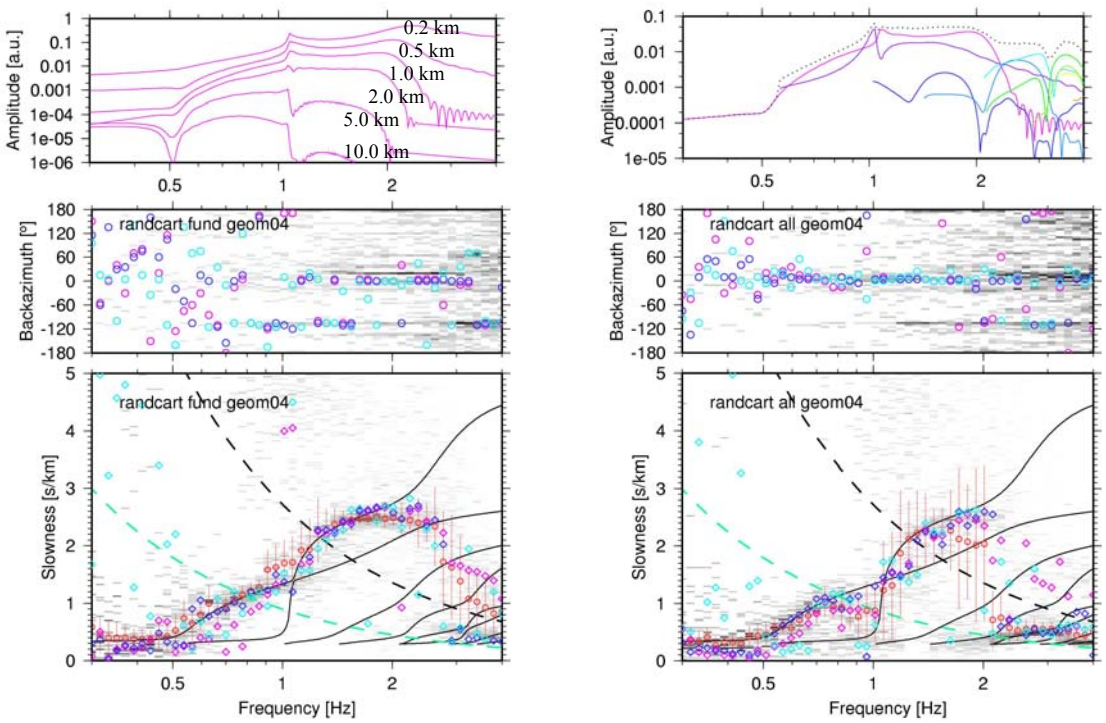


Figure 14: *Lower panels: as Figure 13 but for the 'randcart' source configuration. The upper left panel represents the spectra of the fundamental mode for different distances of propagation (0.2, 0.5, 1, 2, 5 and 10 km) and the upper right panel represents the spectral contribution of the different modes for a 1km distance of propagation. The black dotted curve gives the sum of all contributions.*

c) Case 'randcart' – fundamental mode only

The results are similar to the case '1src' but a larger scattering of the slowness values is observed around the theoretical dispersion curve for frequencies below 1Hz. We attribute this result to the simultaneous arrival of multiple plane waves from distinct directions that the array is not capable to resolve. Interferences between these multiple source contributions lead to biased slowness values, especially for the low resolution capabilities of the CVFK / CVFK2 algorithm. For frequencies above 1 Hz dominant stable source directions (0° and -120°) are estimated. This is consistent with the direction of the closest sources in the 'randcart' source configuration, lying at distances of 1 to 2 km from the array center. The spectral levels of the fundamental mode for closer and more distant sources show larger differences for higher frequencies as can be seen from the spectra plotted on the left top panel. One possible explanation for this behavior is the relatively strong attenuation in the shallow part of the velocity model (Q_S < 20, Figure 3).

d) Case 'randcart' – all modes

As opposed to the '1src' case, the 'randcart' data set seems to be mainly dominated by the fundamental mode in the frequency band [0.6Hz to 2Hz]. However, around 1Hz, all methods show a decrease of the estimated slowness values. We relate this observation to the peculiar feature of our simulated source configuration. The directional estimates for frequencies above 1Hz indicate a dominant energy contribution of the closest sources between 1 and 2 km (see Figure 5). Considering the spectra of the individual Rayleigh wave modes at 1 km distance (upper left panel of Figure 14), the first higher mode exceeds the energy contribution of the fundamental mode around 1Hz. Above 2.2 Hz, only higher modes contribute significantly to the ambient noise wavefield. These facts are consistently reflected in the slowness frequency estimates obtained from the different f-k decomposition techniques.

In a second step, we have evaluated the spatial autocorrelation coefficients for the same datasets and station configurations as used for the f-k analysis by applying the modified SPAC approach by Bettig et al. (2001). A direct comparison between the results obtained from the SPAC method and the previously discussed f-k techniques involves the inversion of observed SPAC curves into a dispersion curve. In order not to introduce additional ambiguities related to the stability of this non-linear inversion problem, we preferred to forward compute the theoretical autocorrelation curves from the theoretical dispersion curves. Therefore we can judge qualitatively the performances of the individual array methods by comparing the results against the same benchmark.

In Figure 15 we show the frequency dependence of the averaged spatial autocorrelation coefficients for the station configuration geom05 (see Figure 4, Table 2). From the co-array configuration we selected seven rings with mean radii ranging from 8 m to 100 m. The number of station pairs in each ring varies from 4 to 10 and the azimuth coverage spanned from 90 to nearly 180 degrees. The computation was performed for the dense random source configuration data sets ('randclose'), both for the fundamental mode Rayleigh wave simulation (left panels) as well as for the case including all higher modes (right panels). We additionally plotted the results of the CVFK and CAPON analysis results for these datasets. We converted the slowness-frequency pairs to frequency-spatial autocorrelation pairs, using Eq. 7 for the minimal and maximal radius for each ring. The distribution of the CVFK results are shown as density plot, the CAPON results are given as individual point estimates.

By comparing the theoretical autocorrelation curves computed from the fundamental (red) and first and second higher modes (green, blue) of the Rayleigh wave dispersion curves to the measured data, we can recognize a very good agreement between the observation and theory for the fundamental mode case. It can especially be noted that for the larger rings a very good match is obtained, which nearly spans the complete frequency band. It is recognized, that both f-k measurements agree very well within the frequency band from 0.7 Hz to 2.5 Hz to the autocorrelation curves. Interestingly, the SPAC outperforms the f-k methods for higher frequencies, following nicely the theoretical dispersion relation, whereas the f-k estimates scatter significantly and don't show any clear relation to the theoretical curve.

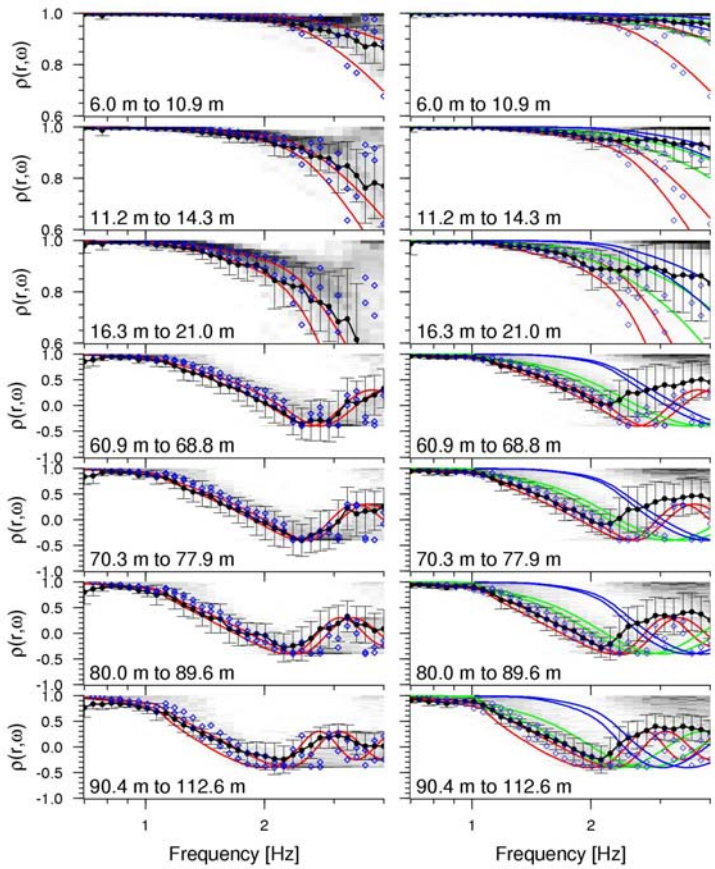


Figure 15: *Averaged spatial autocorrelation coefficients evaluated for 50 frequencies between 0.3 Hz and 4.0 Hz (black dots). From top to bottom, the ring dimension increases from ca. 8 m to 100 m. On the left, we evaluated the simulated fundamental mode Rayleigh wave data set, using the source configuration 'randclose'. Overlaid are the theoretical autocorrelation curves computed from the fundamental mode dispersion curve for the given velocity model and radii* r_1 *and* r_2 *(red curves). On the right we used the same source configuration, but including all higher modes. The theoretical autocorrelation curves for the fundamental and first two higher modes are overlaid (red green blue). For the interpretation, see text.*

A similar good fit between the f-k results and the theoretical and observed autocorrelation curves is obtained for the data set including all higher modes below the frequency limit of 2 Hz. However, above this limit, the observed AC-curves show a clear deviation from the theoretical AC-curves and also the f-k analysis results do not show a clear fit to any of the individual modal autocorrelation curves. We attribute this behavior to the dominating energy contribution of higher modes (compare Figure 14). A simple interpretation can not be given in this case, and advanced inversion strategies including higher mode contributions must be used, which requires additional assumptions about the energy partitioning between the individual modes contributing to the mixed wavefield.

Test site Pulheim – comparison between simulated and real datasets

For the site Pulheim, an ambient vibration array experiment was performed in the year 2000 using an array configuration with 12 elements. The employed sensors were Lennartz LE5D three-component seismometers with an eigenperiod of 5 s. Owing to access constraints, the array was operated as cross array following local dirt roads and/or small trails (compare Figure 4c). Several hours of ambient vibrations were recorded using a sampling frequency of 125 Hz. In order to assess the interpretation of the results that we obtained from the processing of the data set recorded in Pulheim, we simulated ambient vibration wavefields for the same geologic situation that the one in the Lower Rhine Embayment (compare section "Simulating ambient vibration data sets"). We have evaluated both the real and simulated using f-k and spatial autocorrelation methods. We start with the discussion of the f-k analysis results.

CVFK and CAPON method performance

Figure 16 shows both the slowness-frequency curves evaluated from the CVFK and CAPON analysis (lower panels) as well as the directional estimates (middle panels) for the simulated (left panels) and real (right panels) data sets. The theoretical Rayleigh wave dispersion curves for fundamental and first higher modes are plotted for comparison (black curves). In addition, aliasing curves are plotted for the minimal (red dashed), mean (black dashed) and maximal (green dashed) interstation distance within the array configuration. We visualize the histograms obtained from the CVFK analysis as density plots, overlaid by the median and median deviation estimates. For the CAPON method, based on the block-averaged of the cross spectral matrix, we obtain a single slowness map per frequency for the whole analyzed time series. The upper left panel represents the spectral contribution for the fundamental and first higher modes calculated for three different distances of propagation (0.2, 1 and 5 km). The upper right panel represents the spectra of the observed data.

For the simulated data set, we observe that within the [0.6-1.9] Hz frequency band, both CVFK and CAPON follow the fundamental mode of dispersion curve. No stable estimates are observed in the frequency evolution of the direction of propagation, what reflects the random source configuration used for the simulation. However, around 1Hz, we observe a decrease of slowness values for both CVFK and CAPON estimates. We relate this observation to the contribution of higher modes around this frequency. This is confirmed by the spectral contribution of each mode plotted in Figure 16.

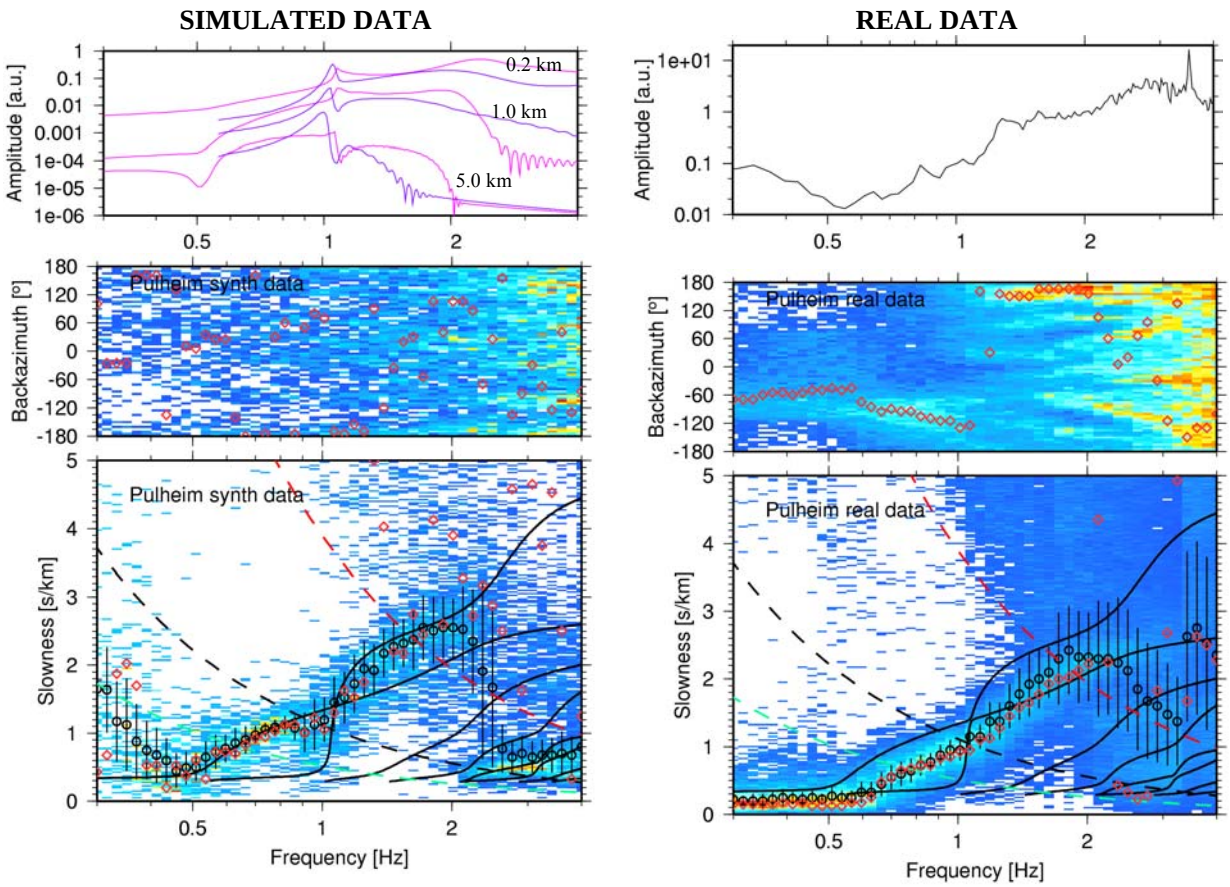


Figure 16: *Results of the CVFK (histograms overlaid by the median and median deviation estimates (black dots)) and the Capon (red dots) analysis for the simulated (lower left panel) and real (lower right panel) datasets. Lower panels: directional and slowness estimates, theoretical Rayleigh wave dispersion curves (black solid) and aliasing conditions (dashed lines). The upper left panel represents the spectral contribution of the fundamental and first higher modes (pink and violet respectively) for three different distances of propagation (0.2, 1 and 5km) and the upper right panel represents the spectra of the observed data.*

As the sources are randomly distributed in space and time, it is difficult to determine whether there exists a dominant wavefield energy contribution from a certain distance range. Therefore, we computed the spectral contributions for the fundamental and first higher modes for three different distances of propagation (0.2, 1 and 5 km). All spectra show that the energy of the first higher mode is overtaking the energy of the fundamental around one 1Hz. Apparently, this creates a mixed multimode wavefield situation that we are able to detect but not to resolve. For this simulated example, we consider that the valid frequency band of reasonable slowness values is restricted to the [0.6-1.9] Hz frequency band. For lower frequencies (< 0.6Hz), we observe large scattering of the slowness values estimates that we attributed to low spectral energy level in the spectra. We think that these small deviations are associated with numerical noise in the forward calculation of the simulated wavefield which deteriorates the phase delay estimates below a certain amplitude level. For higher frequencies (>1.9Hz), the array analysis is limited by aliasing features, due to insufficient spatial sampling. In the same frequency band, the slowness distribution obtained from the CVFK exhibit contributions from higher modes of surface waves.

As for the simulated data set, we observe stable estimates of the slowness values within the [0.6 - 1.9] Hz frequency band for the real data set. From the source distribution we recongize a different situation than in the simulated case, since particularly stable direction of propagation are observed within this frequency band. Compared to the theoretical dispersion corresponding to the simulated wavefield, both CVFK and CAPON values exhibit slower slowness values. Nevertheless, the global shape of the observed dispersion curves is similar to what is expected. Interestingly, we do not observe any contribution of higher modes around 1Hz, as it was the case for the simulated data. For frequencies f<0.6 Hz, we still observe stable estimates of the slowness values. However, the spectral low at 0.55Hz (left uppermost panel) of the observed vertical component spectrum coincides with the H/V peak location (Scherbaum et al., 2003). Therefore, interpreting the H/V ratio from the Rayleigh wave ellipticity, we should not observe phase velocities of the fundamental mode Rayleigh wave around this frequency. We assume that for this frequency range either the array configuration shows too low resolution to separate wavefield contributions from diferrent azimuths or the wavefield is composed of body waves traveling at higher (apparent) phase velocities. In any case, these values can be discarded from interpretation as the very high shear wave velocity for the bedrock (around 5km/s) are not realistic. Finally, for frequencies higher than 1.9Hz, the aliasing features are clearly visible, as for the simulated data set, but no clear contribution of higher modes are observed.

Modified spatial autocorrelation method (MSPAC)

We have evaluated the spatial autocorrelation coefficients for the two corresponding datasets by applying the MSPAC (modified SPAC) approach by Bettig et al. (2001). As in the previous section, we preferred to forward compute the theoretical autocorrelation curves from the theoretical dispersion curves to avoid the inversion problem for estimating phase velocities from autocorrelation curves.

In Figure 17 we show the frequency dependence of the averaged spatial autocorrelation coefficients for the simulated and real data sets. From the co-array configuration we selected six rings with mean radii ranging from 150 m to 730 m. The number of station pairs in each ring varies from 7 to 16 and the azimuth range spanned ranges from 90 to nearly 180 degrees (except for the second ring which presents an azimuth coverage lower than 90 degrees). We additionally plotted the results of the CVFK and CAPON analysis results for these datasets. We converted the slowness-frequency pairs to frequency-spatial autocorrelation pairs for the minimal and maximal radius for each ring. The distribution of the CVFK results are shown as density plot, the CAPON results are given as individual point estimates.

For the simulated data set, the comparison between the theoretical autocorrelation curves computed from the fundamental mode (red curve) of the Rayleigh wave dispersion curves and the observed correlation curves (black circles) are in good agreement, especially for the smaller rings. In addition, around 1Hz, we can clearly observe the contribution of the first higher mode. The auto-correlation curves show a clear deviation from the fundamental mode to the first higher (green curve) for various rings.

SIMULATED DATA **REAL DATA**

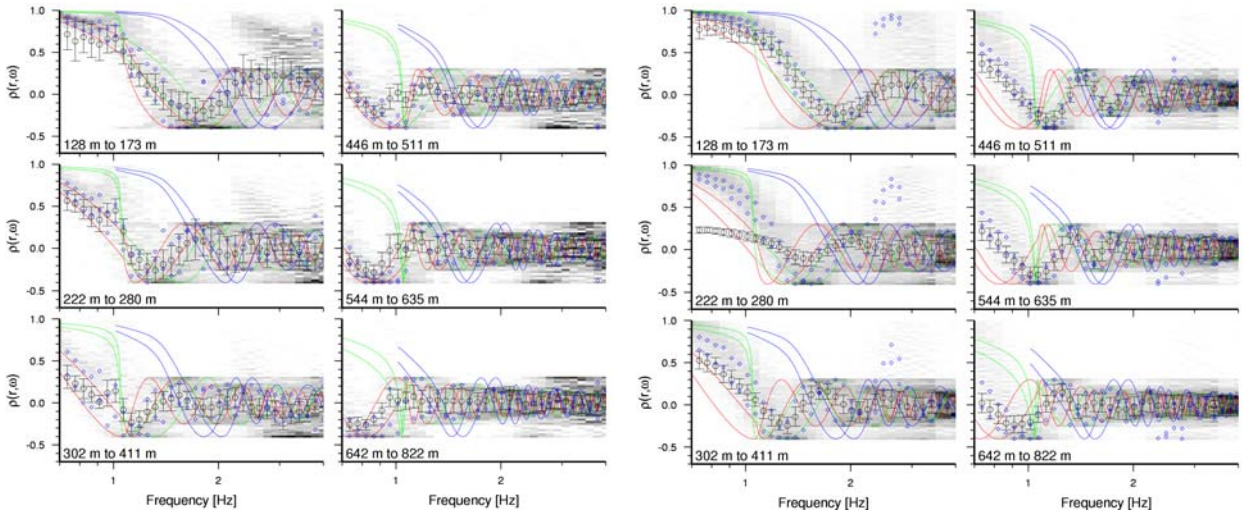


Figure 17: *Averaged spatial autocorrelation coefficients evaluated for 50 frequencies between 0.7 Hz and 4.0 Hz (black dots). From top to bottom, the ring dimension increases from ca. 150 m to 730 m. The real and simulated are evaluated on the left and right panel respectively. Overlaid are the theoretical autocorrelation curves computed from the fundamental (red) and first two higher (green blue) modes for the given velocity model and for the minimum and maximum radii considered in each subplot. For the interpretation, see text.*

For the real data set, the observed autocorrelation curves apparently exhibit similar oscillation frequencies as the simulated ones. However, the curves appear to be shifted in frequency when compared to the theoretical curves. This observation is consistent with the results obtained from the frequency wavenumber decompositions where the observed dispersion show lower slowness values when compared to the synthetic model curves. As already pointed out by the f-k decomposition, we can find no indication of higher modes in the real data set.

For the second ring (222 to 280m), a mismatch of the observed auto-correlation coefficients is observed within the [0.7-1.2] Hz frequency band. The range of azimuth spanned by the co-array station pairs for this particular ring is less than 90° and as observed from the f-k analysis, the wavefield presents a very stable direction of propagation for this frequency range (for details, see Figure 18). As a consequence the approximation of the averaged spatial autocorrelation coefficients by the Bessel function (Eqs. 6 or 7) is no longer valid. This effect is also known as directional aliasing (Henstridge, 1979). It should be noted, that this mismatch is not observed for the simulated data in Figure 17, as the spatial randomness of the source distribution compensates the small azimuthal coverage of station pairs.

Summarizing the observations, we come to the following conclusions: For the simulated data set we are able to detect the presence of higher mode contributions in the synthetic wavefield from an unexpected deviation of dispersion and spatial autocorrelation curves for a distinct frequency band (0.8-1.1 Hz). We can attribute this to the first higher mode, which energetically dominates the wavefield. However, this higher mode contribution can not be clearly separated nor can it be followed for a broader frequency range. Comparing the analysis results from different methods we determine the valid frequency band of dispersion curve interpretation to lie in the range between 0.6 and 1.9 Hz. Limitations are given from the occurrence of aliasing at higher frequencies and lack of energy for lower frequencies.

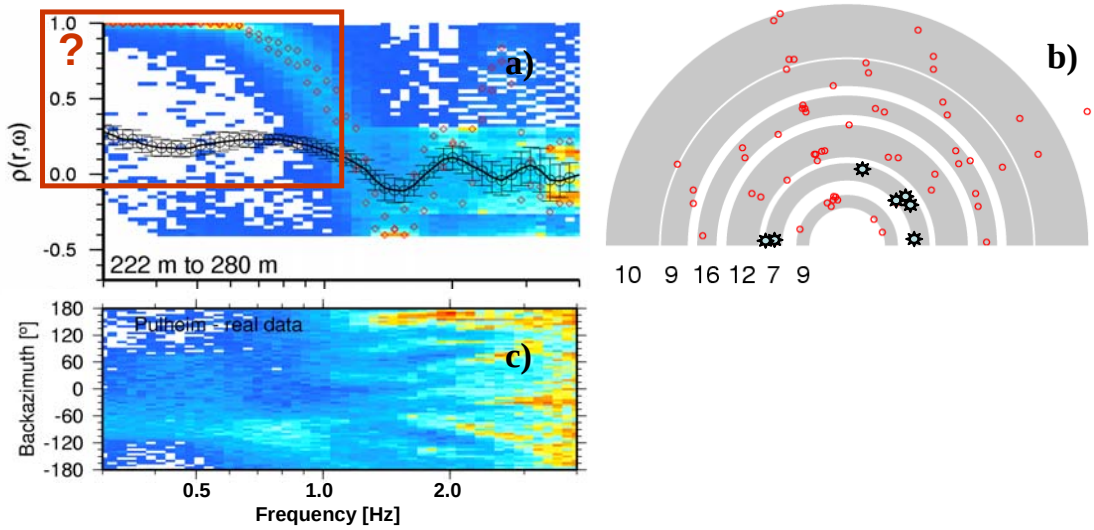


Figure 18: *Possible bias of autocorrelation results caused by insufficient azimuthal sampling. a) Autocorrelation curve (black circles with error bars, MSPAC) computed for the second ring of the co-array distribution shown in panel b) (blue stars). The underlying density map shows for comparison the results of individual f-k estimates (CVFK) converted to autocorrelation values by Eq. 7 for r_1 and r_2 (f-k mean values depicted as red diamonds). An unexpected distortion of the autocorrelation curve is observed for frequencies below 1Hz. The effect is explained, by the existence of a dominant wavefield direction as determined from the f-k analysis in c) in combination with only partially sampled azimuth range by the contributing station pairs as shown in b). The wavefield propagation directions are concentrated along wave paths in W or WSW azimuth and the station pair directions cover only a restricted azimuth range (lacking directions from N to WNW).*

For the real ambient vibration array recordings at PLH we are not able to detect any indication for the presence of higher mode surface waves contributing to the observed wavefield. However, a strong deviation of the stable slowness estimates from the fundamental mode predicted by the general reference model (Brüstle and Stange, 1999) is observed. Considering that the shear wave velocity profile for this model has been confirmed in previous studies (Scherbaum et al., 2003) this observation can not be easily explained. We speculated therefore whether the measured phase velocity curve could be better explained by higher mode contributions (e.g. first higher mode for frequencies above 1.3 Hz, Figure 16). Nevertheless we reject this possible explanation for the following reason: considering the spectral contributions of individual mode branches for the simulated data set (relying on the reference model), the energetically dominance of the first higher mode should lie around 1 Hz and is relatively independent of the source – receiver geometry. In the analysis result we can not detect any remarkable behavior of the smoothly estimated dispersion curves at these frequencies. We conclude therefore, that we observe a wavefield dominantly composed of fundamental mode Rayleigh waves and the shift between the predicted and observed dispersion characteristics are due to discrepancies of the generic velocity model and the real site structure. As for the simulated analysis results, we set the valid frequency band for the interpretation of the dispersion characteristics to the range between 0.6 and 1.9 Hz. Limitations are given from the occurrence of aliasing at higher frequencies and inconsistently high phase velocities for the lower frequencies.

Virtual test site Liege – array size

Regarding the resolution and aliasing limits for f-k analysis methods, we have concluded, that it is not possible to analyse the phase velocity curves for large frequency bands. We tested the proposed iterative deployment strategy (see also D07.05 and WP13) in this virtual testbed. We have used a set of hexagonal array layouts with varying aperture for the wavefield simulations. Thus, we simply analysed the wavefield by CVFK and CAPON methods for those virtual array settings. The analysis results are shown in Figure 19 for array apertures 4m (geom-10), 30m (geom-12), 50m (geom.-13), and 100m (geom.-14). In order to ease the interpretation and to avoid a mixing with other wave propagation or source configuration effects, we used the ideal situation of a single source location and fundamental mode Rayleigh wave propagation in all cases.

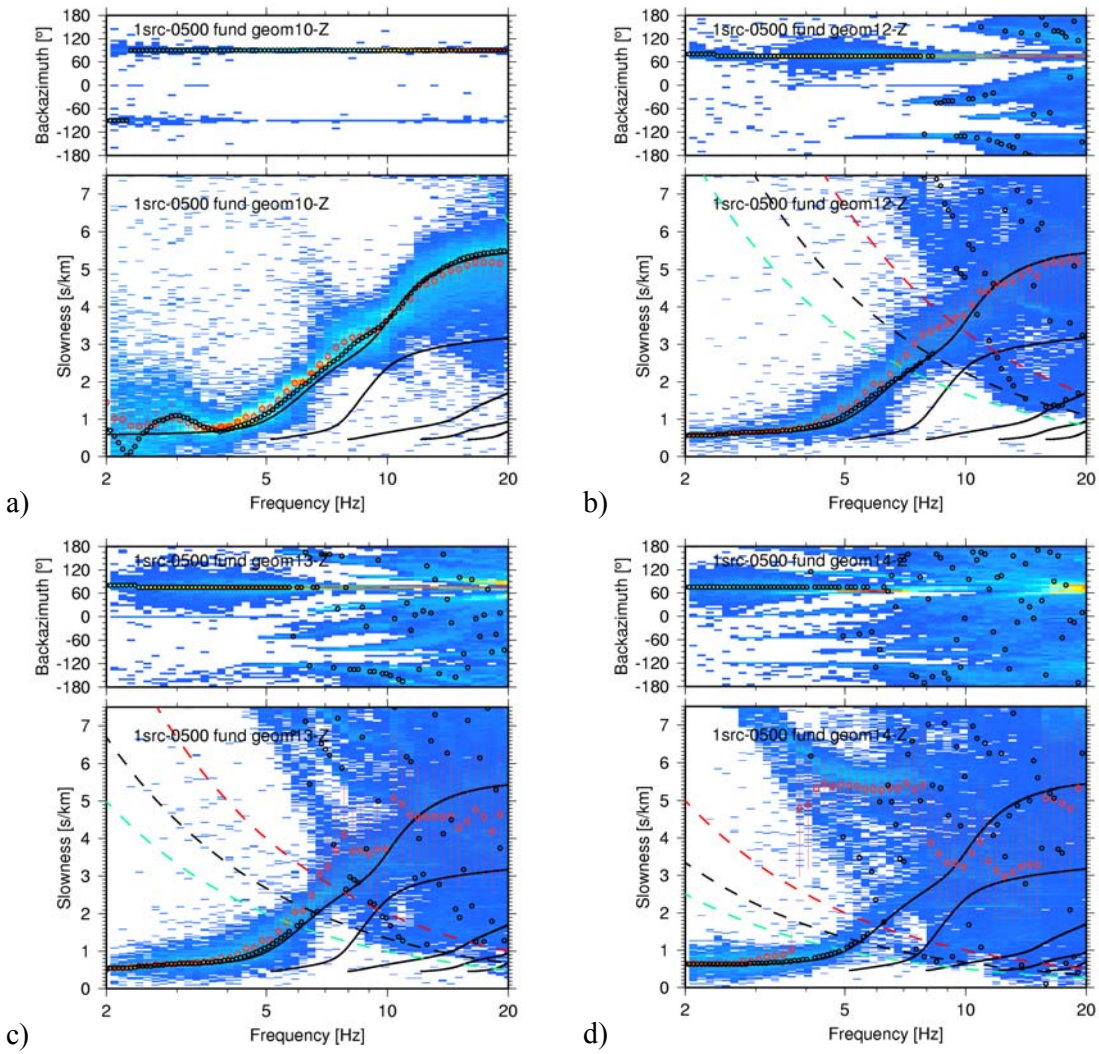


Figure 19 : *Testing the re-iterative deployment strategy, starting at small array apertures from 4m in panel a) to 100m aperture in panel d). Aliasing conditions for minimal, average and maximal inter-station distances are depicted as red, green and black dashed line, respectively. The partially resolved phase velocity branches shift from high to low frequencies.*

It can be well recognized from Figure 19, that, when increasing the aperture of the array, the part of the dispersion curve which is well resolved and provides high quality estimates of the phase velocities is shifted to lower frequency bands. As expected, the aliasing features also start at smaller frequency values and it is relatively easy to recognize at which point, the phase velocity curves are no longer reliable.

Virtual test site Liege – attenuation structure

While analysing the simulated datasets for fundamental mode Rayleigh waves and the simplest source configuration (single source location), one peculiar feature was observed. The phase velocity curves derived from "far-distant" (>1 km) sources showed, especially for higher frequency bands, poor quality when compared to the results obtained from closer located excitation points of the wavefield (compare sequence of f-k results in Figure 20).

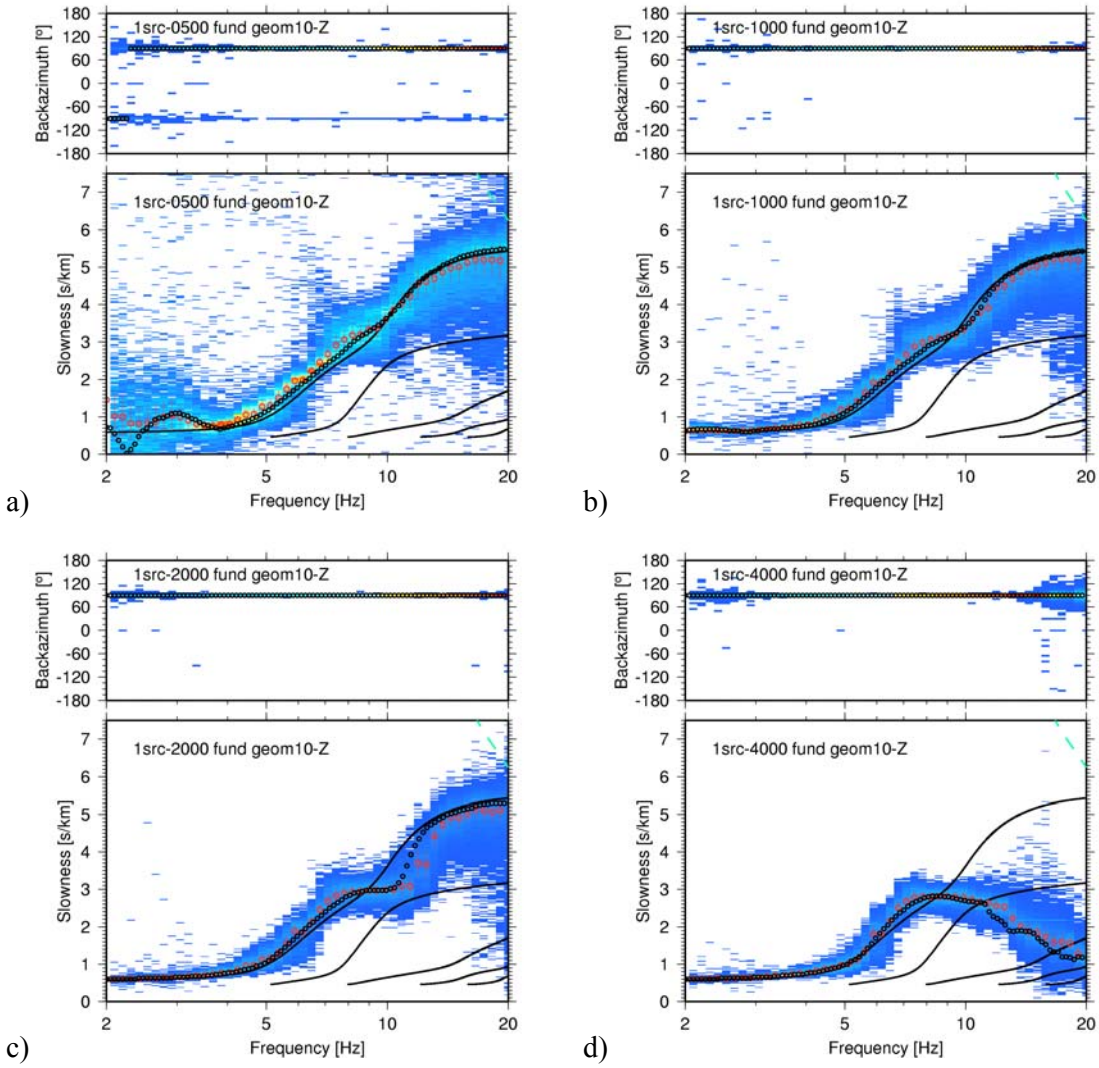


Figure 20 : *Influence of attenuation structure on phase velocity estimates. For the datasets comprising Rayleigh wave fundamental mode only, the single acting source is moved from 500 m distance from the array center in a) to a distance of 1000m in b), 2000m in c) and 4000 m in d). Whereas for short travel distances of the fundamental mode Rayleigh wave, the phase velocities are well estimated for frequencies above 10 Hz, the true propagation velocities can no longer be resolved for far-distant source locations.*

This effect was contradictory to our expectations regarding the influence of curved wavefront arrivals as discussed above (section "Curved wavefronts and related uncertainties"). The spectral analysis of the individual time series revealed, that the cause of this outcome is related to a lack of coherent energy which can be attributed to the intrinsic attenuation properties of the medium. In order to confirm this observation, we tried to derive the influence of the site's attenuation properties on the array analysis results by comparing simulated datasets for two velocity models which just differed in their respective Q-factors in the individual layers (compare entries low and high Q in Table 1). A single source at a distance to the array center of 500 m only was used for the simulation and all Rayleigh wave modes have now been included in this computation.

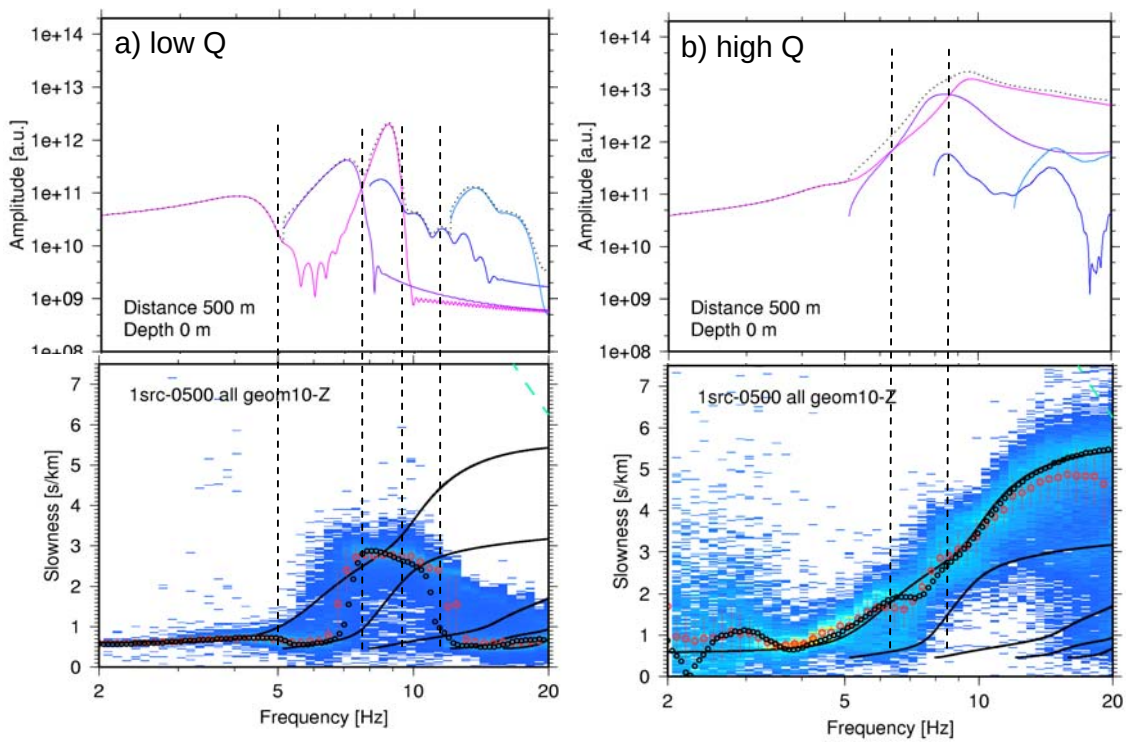


Figure 21 : *Influence of attenuation structure on the phase velocity estimates. a) Dispersion curve analysis by CVFK method for a shallow sedimentary site with low Q-factors (high intrinsic attenuation) and corresponding spectral energy contributions of individual Rayleigh wave modes on top. The simulated wavefield is composed of repeated excitation of a single source located at 500m distance from the array center. The colored density map depicts the histogram of slowness estimates from all analysed time windows within narrow frequency bands. Black curves indicate the theoretical Rayleigh wave dispersion curve relations for the velocity structure. b) as a), but for model with high Q-values (low intrinsic attenuation). Vertical dashed lines show the location of cross-over points, where the dominating energy contribution in the wavefield changes from one mode branch to another.*

The phase velocity estimation results obtained from the conventional and high resolution f-k algorithms (CVFK and Capon) are shown in Figure 21 together with the corresponding spectral energy contributions of the individual Rayleigh wave modes. The comparison of the spectra indicates the appearance of pronounced holes for the fundamental and higher Rayleigh wave modes for the stronger attenuating model. Interestingly, the higher mode energy contributions dominate for certain frequency bands, whereas for the less attenuating structure, the fundamental mode is the strongest wavefield component for large portions of the analysed

frequency band. As a consequence, the estimated phase velocities jump between the theoretical dispersion curves of fundamental and higher mode branches, depending on their modal energy contributions of the wavefield. Intermediate phase velocities are found for comparable energy levels between different modes of propagation.

The existence of intermediate phase velocity estimates has been previously described by Tokimatsu et al. (1992a,b) and Tokimatsu (1997). The authors termed the observed dispersion curve results "apparent phase velocities" and suggested to include the media response of higher mode contributions into the forward computation problem when inverting for the velocity model.

Our simulations suggest that the attenuation structure heavily influences the outcome of the estimated dispersion curves. This implies an additional dependence of the results on the involved source-receiver distances which would rule out any attempt to find a correct treatment in the inversion task. However, from experimental results in real environments, we know that ambient vibration array analysis results are reproducible and remain temporally relatively stable. Without further investigations about the true nature of the ambient vibration wavefield, it seems difficult to bring these conflicting observations into agreement. However, for the moment, we prefer the following interpretation: The stability of real ambient vibrations measurements is a consequence of the randomness of the source distribution and the attenuation properties of the media. For uniform spatial random distribution of source locations, the number of contributing sources increases proportional to the areal increase around the receiver array, i.e approximately linear with the distance. Assuming a more or less homogenous excitation strength A_0 of the individual sources, the superposed amplitude contributed to the wavefield is a linear function of distance (neglecting interference effects). On the other hand, geometrical spreading for surface waves shows a root square dependence of the amplitudes with distance. Finally, due to the intrinsic attenuation, distant source contributions are further diminished with an exponential depending on the wave velocity, frequency and effective Q. Therefore, the total amplitude contribution A from sources located in some distance x from the array center may be expressed as:

$$A \propto A_0 x \frac{1}{\sqrt{x}} e^{\frac{2\pi f x}{c(f) Q_{eff}(f)}} = A_0 \sqrt{x} e^{\frac{2\pi f x}{c(f) Q_{eff}(f)}} \qquad \text{Eq. 8}$$

Further, the average energy density carried by a wave can then be written as:

$$\overline{E} = \frac{1}{2} \rho A^2 \varpi^2 \propto \frac{1}{2} \rho \varpi^2 A_0^2 x e^{\frac{\pi f x}{c(f) Q(f)}} \qquad \text{Eq. 9}$$

In Figure 22 we have plotted Eqs. 8 and 9 for a set of 10 frequencies ranging from 1.0 to 10.0 Hz, taking the phase velocity values of the corresponding fundamental Rayleigh wave mode dispersion curve of site Liege and assuming a frequency independent Q value of 25.

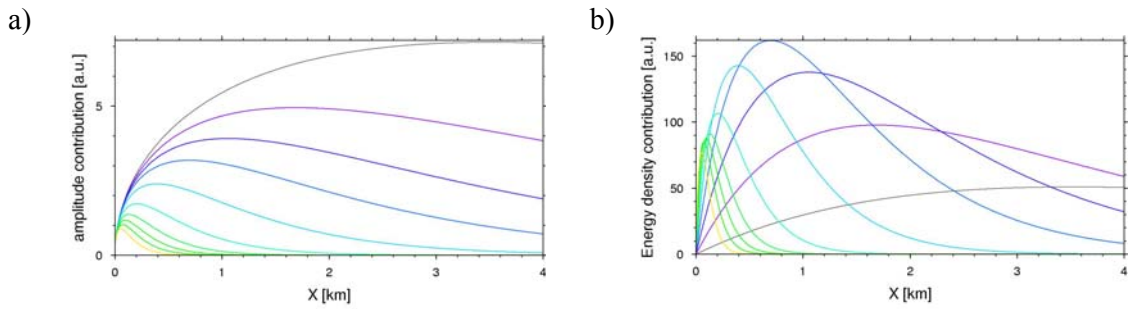


Figure 22 : *a) Relative amplitude and b) relative energy density contributions from a uniform random source distribution, assuming simple cylindrical geometrical amplitude decay for surface waves and frequency independent Q factor. The set of colored curves have been evaluated for frequencies ranging from 1 Hz (orange) to 10 Hz (black) in steps of 1Hz, using the phase velocities of the fundamental Rayleigh wave mode for site Liege. For details, see text.*

Although the frequency independence of effective Q for Rayleigh waves is probably a bad assumption (Q should increase for longer wavelengths, as the penetration depth increases), it seems that the distances for sources which contribute a significant amount of energy to the observed microtremor wavefield energy is limited to a relative narrow range. Clearly, this is a very crude and straightforward calculation, however, it seems interesting, whether field experiments could be conducted to validate or disapprove these thoughts.

Discussion and conclusions

We compared the performance of different frequency wavenumber approaches as well as the spatial autocorrelation method in order to determine under which circumstances high-quality site-specific surface wave dispersion characteristics can be obtained from the analysis of microtremor wavefields. Using simulated ambient vibration wavefields, we studied the influence of both the array layout (number of sensors, geometry and interstation distances) and source characteristics as well as propagation effects of the noise wavefield for each of the employed methods.

For the simplest case, one dominant source and fundamental mode only, we found that all methods perform equivalently well. Considering more realistic wavefield situations, randomly distributed sources and/or higher mode contributions, the direct interpretation of frequency-slowness values becomes more difficult. In particular we observed significant bias of slowness values in case of insufficient resolution capabilities of the array configurations. In addition, in case that higher mode Rayleigh waves contribute significantly to the total energy, mixed dispersion curve characteristics will be obtained. For well resolving configurations it is, for favorable wavefield situations, possible to separate individual mode contributions. However, the interpretation of which mode is observed is not straight forward. For the inversion of shear velocity profiles advanced strategies have to be employed to make use of this information.

For the practical task of dispersion curve determination from ambient vibration array recordings we suggest the use of combinations of various f-k methods. As most stable

algorithm we regard the conventional f-k (CVFK). Although it offers only low resolution capabilities, it allows the determination of robust propagation characteristic distributions when applied in a sliding window analysis. We find that these distributions are especially advantageous for the visualization and determination of reasonable uncertainty measures to recognize the validity of dispersion curve estimates. Capon's high resolution f-k approach (CAPON) is well suited to complement the CVFK analysis, as it gives less biased estimates and allows phase velocity determination for higher frequencies where aliasing is already deteriorating the CVFK results. In addition to CVFK and CAPON frequency-wavenumber decompositions, we suggest to use the spatial autocorrelation method in order to further cross-check the results on the same array data set. Compared to frequency-wavenumber methods, the SPAC gives reliable estimates of the dispersion characteristics within a larger frequency band, and allows additionally an easier recognition of the presence of higher modes from the unexpected occurrence of oscillations in the autocorrelation curves. Less suitable for the goal of dispersion curve estimation are the CVFK2 and MUSIC2 approaches. CVFK2 exhibits a limited resolution capability whereas the main drawback of the MUSIC2 algorithm is the difficulty to reliably assess the number of sources spanning the signal subspace for the typical multi-source situations in the ambient vibration wavefield.

From the tests performed we conclude the following for the interpretation of dispersion curves: the valid frequency band for interpreting the dispersion characteristics obtained from f-k analysis techniques is limited on both sides. For low frequencies the limitation is either caused by insufficient resolving capabilities of the chosen array layout or by the vanishing spectral energy contribution of the vertical Rayleigh wave displacements (for example related to the degenerated horizontal ellipticities around the frequency of the H/V spectral peak location). The upper limit is given by the occurrence of aliasing patterns due to insufficient spatial sampling of the wavefield. In order to improve the determination of dispersion curves, we recommend the use of adaptive array deployment strategies (see below). Array apertures and interstation distances should be adapted for distinct target wavelength ranges. Thus, starting from short wavelengths and going to higher wavelengths well resolved partial dispersion curves can be obtained even in complex wavefield situations. However, ambient noise excitation as well as particular propagation effects may lead to misinterpretation of phase velocities or autocorrelation coefficients obtained from array analysis. The use of various combinations of analysis methods may allow to prevent this eventual misinterpretation by providing complementary information on the ambient vibration wavefield characteristics. Contradictory results obtained from the individual methods may be an indicator to recognize such situations.

For a short overview, we have summarized some of the key-observations made from the numerical experiments and the analysis of simulated wavefields in Table 6. On basis of each observation, it is possible to derive conclusions and/or recommendations with respect to the determination of reliable dispersion curve estimates.

Experiment	***Observation***	***Conclusions/Recommendations***
Plane wavefronts (randomly perturbed)	FK: - mean is unbiased estimate - variance scales with array aperture - variance scales with slowness	- larger time windows needed to improve phase delay estimate. - large number of windows necessary to obtain good statistics - larger array sizes perform better
Curved wavefronts	FK: - biased estimate of true slowness (lower) - bias scales with 'curvedness' - 'curvedness' related to ratio of source distance to array radius d/r SPAC: - directional dependent deviation - average of estimates unbiased until d/r~1 - d/r<1 biased estimates (larger)	- sources inside array geometry give unreasonable results (CVFK) or overestimate the slowness (SPAC). - nearby sources d/r<2.5 must be avoided →semblance based estimate outperforms power-based algorithms - very small arrays preferred (allows certain control of inside sources / d/r increases fast from site of operation).
2D-Aliasing (narrow-band estimate)	FK: - multiple plane wave arrivals lead to biased estimates (always lower) - bias scales with resolving capability of the array (related to array size) - aliasing patterns related to peaks in cumulative criteria failure curve SPAC: - similar aliasing criterion as for FK methods. - better resolution capabilities for longer wavelengths.	- preferable symmetric arrays showing similar array response behaviour for all directions and favourable cumulative criteria failure curve. - circular array setups allow combined usage of f-k and SPAC techniques.
Wavefield simulation: single vs. multiple arrival directions	FK: - dominant source directions are favourable situation, multiple signal directions lead to biased results (s.a. 2D-Aliasing) SPAC: - dominant direction of wavefield propagation may introduce biased results due to insufficient directional sampling using unfavourable array layouts.	- complementary use of both f-k and spatial autocorrelation techniques is required to resolve peculiar wavefield situations.
Wavefield simulation: low attenuating vs. high attenuating site structure	All methods: - high attenuating structures may favour the existence of higher mode contributions in the wavefield. Apparent phase velocity curves are observed and interpretation is complicated.	Kinks in dispersion curve estimates at certain frequencies obtained for different analysing methods indicate the existence of higher mode contributions in the wavefield. An appropriate treatment in the inversion procedure should be considered.

Table 6 : *Summary of observations and corresponding conclusions/recommendations.*

Measurement strategy for ambient vibration array analysis

From the above experiments and argumentation it seems clear that, due to the complex interaction between measurement layout, analysing method and wavefield situation, the determination of highly reliable, multi-modal dispersion curve estimates with small error bounds in general remains a difficult task. Nevertheless, we could identify sources of error and the limiting factors of the estimation procedure. Most important herein seems the peculiar observed wavefield situation, which is a result from both excitation, and propagation properties:

- source location and excitation: space-time density (single source, dominant source area, multiple sources) of ambient vibration sources and their spectral energy content (source time function, source orientation and source depth).
- wave propagation: filter effect of layered velocity structure and Q-structure, resulting in changes of the energy partitioning of surface wave modes).

In order to improve the dispersion curve determination from microtremor recordings, we can at best react to the given ambient wavefield situation by improving the measurement and analysis scheme. *Due to the absence of a priori knowledge about the specific site conditions in terms of source excitation and wave propagation characteristics (i.e. the site structure), it is necessary to acquire the necessary information from analysis of the recordings in the field.* Then, we are able to judge the wavefield situation on site and adapt our measurements (array layout / processing) accordingly. In particular, we propose the following general work flow:

Pre-experimental steps:

(1) Select adequate equipment for measurements (f_{seis} << lowest frequency of interest)

(2) Huddle test – verify calibration of instruments (i.e. estimate possible instrument related delay times)

Measurement/Processing/Analysis scheme:

(3) Set up very small array, e.g.: circle with N sensors with radius r = 2…5 [m], the number of sensors N has to be chosen as a compromise between cost and resolution capabilities. A minimum of N=5 sensors should be considered in any situation.

(4) Estimate amplitude spectra and H/V ratios at individual array stations. By doing so, some information of the analyzable frequency bands can be acquired. Furthermore, a detection of a possible 2D- or 3D-site condition can be inferred!

(5) Apply narrowband array analysis methods for the full frequency band which has been determined as appropriate (4). Make use of competing algorithms (CVFK, CAPON, (M)SPAC). Use sufficiently large number of long time windows for the analysis.

(6) Derive frequency band(s) which show(s) consistent phase velocity estimates for all methods. Contradicting results may indicate complex wave field situations which may not be resolvable (mixed mode contributions) and mark those results as critical information which may require extra treatment in the inversion process.

   - derive corresponding wavelength range
   - cross check results with capabilities/limitations of array geometry in terms of aliasing and resolution

- derive (first) partial branch of dispersion curve

(7) Define next narrow target wavelength range which extends the previously analyzed wavelength interval to longer wavelengths and seek the array aperture being optimal for this wavelength range.

(8) Rearrange array geometry according to (7)

(9) Repeat steps (4) to (8) until a well estimated, consistent dispersion relation is obtained for the frequency band of interest (i.e. necessary frequency band for the inversion of velocity models).

The proposed procedure is summarized in the following Figure 23:

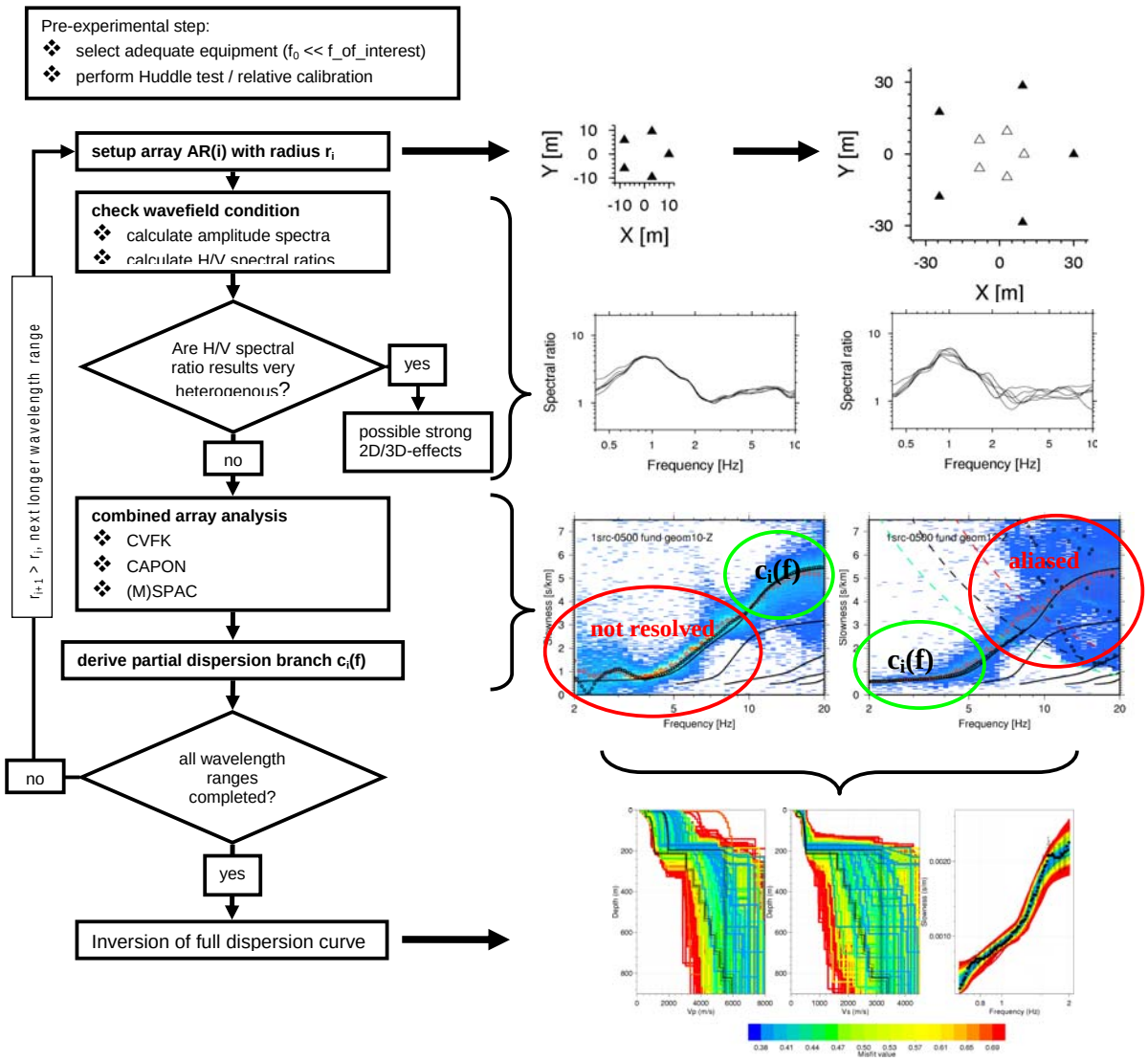


Figure 23 : *Proposed measurement scheme for determination of improved broadband dispersion curve characteristics by iterative deployment of adapted array configurations for narrow wavelength ranges.*

Acknowledgements

This project is supported by the European Commission – Research General Directorate (Project No. EVG1-CT-2000-00026 SESAME). We thank all colleagues and friends from our partner institutions, who have accompanied us within the last three years. Our gratitude goes to all those who have contributed to the work by participating in wonderful workshops, put their helping hands during field experiments, or simply gave critical and constructive comments to the work. All maps and figures were generated with the Generic Mapping Tools GMT Version 3. (Wessel and Smith, 1998)

References:

Akaike, H., 1973: *Information theory and an extension of the maximum likelihood principle*, Proc. 2nd Int. Symp. Inform. Theory, 267-281.

Aki, K., 1957: *Space and time spectra of stationary stochastic waves, with special reference to microtremors*, Bull. Earthquake Res. Inst. Tokyo Univ., 35, 415-456.

Asten, M.W. & Henstridge, J.D., 1984: *Array estimators and the use of microseims for reconnaisaance of sedimentary basins*, Geophysics, 49(11), 1828-1837.

Bettig B., Bard P.-Y., Scherbaum F., Riepl J., and Cotton F., 2001: *Analysis of dense array noise measurements using the modified spatial auto-correlation method (SPAC). Application to the Grenoble area*, Bolletino di Geofisica Teorica ed Applicata, 42(3/4), 281-304.

Boore D.M. and Brown L.T, 1998: *Comparing shear-wave velocity profiles from inversion of surface-wave phase velocities with downhole measurements: systematic differences between CXW method and downhole measurements at six USC strong-motion sites*, Seismol. Res. Lett., 69(3), 222-229.

Bonnefoy-Claudet S., Cornou C., Kristek J., Ohrnberger, M., Wathelet M., Bard P.-Y., Moczo P., Faeh D., Cotton F., 2004: *Simulation of seismic ambient noise: I. Results of H/V and array techniques on canonical models*, Paper No. 1120, XIII World conference on Earthquake Engineering, Vancouver, B.C., Canada, August 1-6, 2004.

Brown L.T., Boore D.M., and Stokoe II K.H., 2002: *Comparison of Shear-Wave Profiles at 10 Strong-Motion Sites from Noninvasive SASW Measurements and Measurements Made in Boreholes*, Bull. Seism. Soc. Am., 92(8), 3116-3133.

Brüstle, W. and Stange, S., 1999: *Geologische Untergrundklassen zum Entwurf von Normspektren für DIN4149 (neu)*, LRGB Baden-Württemberg, AZ: 3480.01/98–4764, 1999.

Budny M., 1984: *Seismische Bestimmung der bodendynamischen Kennwerte von oberflächennahen Schichten in Erdbebengebieten der Niederrheinischen Bucht und ihre ingenieurseismologische Anwendung*. Ph.D. Thesis (in German), Special publications No. 57, Geologisches Institut der Universität zu Köln, pp. 209.

Capon, J., 1969 : *High-resolution frequency-wavenumber spectrum analysis*, Proceedings of the IEEE, 57(8), 1408-1418.

Cornou C., Di Giulio G., Ohrnberger M., Kristek J., and Wathelet M., 2004a: *Simulated vs. observed seismic ambient noise in the Colfiorito basin: site effects estimation and noise wavefield characteristics*, European Seismological Commission XXIX General Assembly, SCF-3, Potsdam, Germany, September 12-17, 2004.

Cornou C., Kristek J., Ohrnberger, M., Di Giulio G., Schissele E., Guillier B., Bonnefoy-Claudet S., Wathelet M., Fäh D., Bard P.-Y., and Moczo P., 2004b: *Simulation of seismic ambient noise: I. Results of H/V and array techniques for real sites*, Paper No. 1130, XIII World conference on Earthquake Engineering, Vancouver, B.C., Canada, August 1-6, 2004.

Harjes, H.-P., 1990: *Design and siting of a new regional array in Central Europe*, Bull. Seism. Soc. America, 80, 1801-1817.

Haubrich, R.A., 1968: *Array design*, Bull. Seis. Soc. America, 58, 997-991.

Henstridge J.D., 1979: *A signal processing method for circular array*, Geophysics 1979: 44, 179-184.

Hermann R.B., 1996: *Computer programs in seismology, Version 3.0*, St. Louis University, http://www.eas.slu.edu/People/RBHerrmann/ComputerPrograms.html.

Hinzen, K.-G., Scherbaum F. and Weber B., 2004: *On the resolution of H/V measurements to determine sediment thickness, a case study across a normal fault in the Lower Rhine Embayment, Germany*, Journal of Earthquake Engineering, 8(6), 909-926.

Horike M., 1985: *Inversion of phase velocity of long-period microtremors to the S-wave-velocity structure down to the basement in urbanized area*, J. Phys. Earth, 33, 59-96.

Horike M., 1996: *Geophysical exploration using microtremor measurements*, Xth WCEE, Acapulco, Paper no. 2033, Elsevier Science Ltd.

Ibs von Seht M. and Wohlenberg R., 1999: *Microtremor measurements used to map thickness of soft soil sediments*, Bull. Seism. Soc. Am., 89, 250-259.

Ishida H., Nozawa T. and Niwa M., 1998: *Estimation of deep surface structure based on phase velocities and spectral ratios of long-period microtremors*, in: Irikura K., Kudo K.,

Kind, F., 2002: *Development of Microzonation Methods: Application to Basle, Switzerland*, Dissertation ETH No. 14548, Swiss Federal Institute of Technology, Zurich, 110 p.

Kvaerna, T., and Ringdahl, F., 1986: *Stability of various f-k estimation techniques*, Semmiannual technical summary, 1 October 1985 – 31 March 1986, NORSAR Scientific Report, 1-86/87, Kjeller, Norway, 29-40.

Miyakoshi K., Kagawa T. and Konioshita S. "Estimation of geological structures under the Kobe area using the array recordings of microtremors". Irikura K., Kudo K., Okada H., and Satasini T., Editors. The Effects of Surface Geology on Seismic Motion, Rotterdam, Balkema, 691-696, 1998.

Ohrnberger, M., F. Scherbaum, F. Krüger, R. Pelzing and Sh.-K. Reamer, 2004a: *How good are shear wave velocity models obtained from inversion of ambient vibrations in the Lower Rhine Embayment (NW-Germany)*, Boll. Geof. Teor. Appl., 45(3), pp. 215-232.

Ohrnberger, M., Schissele E., Cornou C., Wathelet M., Savvaidis A., Scherbaum F., Jongmans D., and Kind F., 2004b: *Microtremor array measurements for site effect investigations: comparison of analysis methods for field data crosschecked by simulated wavefields*, Paper No. 0940, XIII World conference on Earthquake Engineering, Vancouver, B.C., Canada, August 1-6, 2004

Ohrnberger, M., Schissele E., Cornou C., Bonnefoy-Claudet S., Wathelet M., Savvaidis A., Scherbaum F., and Jongmans D., 2004c: *Frequency wavenumber and spatial autocorrelation methods for dispersion curve detmerination from ambient vibration recordings*, Paper No. 0946, XIII World conference on Earthquake Engineering, Vancouver, B.C., Canada, August 1-6, 2004.

Parolai S., Bormann P. and Milkereit C. ”New relationships between vs, thickness of sediments, and resonance frequency calculated by the H/V ratio of seismic noise for Cologne Area (Germany)”. Bull. Seism. Soc. Am. 2002: 92, 2521-2527.

Schmidt, R.O., 1981: *A signal subspace approach to multiple emitter location and spectral estimation*, Ph.D. Dissertation, 201 pp., Stanford University, Stanford, California.

Schmidt, R.O., 1986: *Multiple emitter location and signal parameter estimation*, IEEE Trans. on Antennas and Propagation, 34, pp. 276-280.

Tokimatsu, K., Tamura, S. and Kojima, H., 1992a: Effects of multiple modes on Rayleigh wave dispersion. Journal of Geotechnical Engineering, ASCE, 118(10): 1529-1543.

Tokimatsu, K., K. Shinzawa, and S. Kuwayama, 1992b. “Use of shortperiod microtremors for VS profiling”. J. Geotech. Eng. 1992: 118(10), 1544–1588.

Tokimatsu, K., 1997: *Geotechnical site characterization using surface waves*, in: Earthquake Geotechnical Engineering, Ishihara (ed.), pp. 1333-1368, Balkema, Rotterdam.

Wax M. and Kailath, T. “Detection of signals by information theoretic criteria”. IEEE Transactions on ASSP, 1985: 33(2), 387-392.

Wessel, P., and W. H. F. Smith, 1998: *New, Improved Version of Generic Mapping Tools Released*, EOS Trans., AGU, 79 (47), p. 579.

Woods, J.W. and Lintz, P.L., 1973: *Plane waves at small arrays*, Geophysics, 38, 1023-1041.

Yamanaka H. “Geophysical explorations of sedimentary structures and their characterization”. Irikura K., Kudo K., Okada H., and Satasini T., Editors. The Effects of Surface Geology on Seismic Motion, Rotterdam, Balkema, 15-33, 1998.

Yamanaka H., Takemura M., Ishida H., and Niwa M., 1994: *Characteristics of long-period microtremors and their applicability in exploration of deep sedimentary layers*, Bull. Seism. Soc. Am. , 84(6), 1831-1841.

Zywicki D. J. “Advanced signal processing methods applied to engineering analysis of surface waves”, PhD. Thesis Georgia institute of technology, 227 p., 1999.