

Site Effects Assessment Using Ambient Excitations

SESAME

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

Report on

Optimum deployment strategy for array measurements

University of Potsdam, Germany

WP05

Instrumental layout for array measurements

Deliverable D07.05

December 2002

List of Contents

Summary	3
1 Introduction	4
2 Resolution test for noise-free multiple plane waves and arbitrary array configurations ...	4
2.1 Single plane wave arrival	7
2.2 Three plane wave arrivals (varying azimuth)	10
2.3 Three plane wave arrival (varying slowness)	12
3 Resolution test with synthetic waveforms	13
3.1 Preparation of synthetic waveform data	13
3.2 Array analysis of synthetics I (FM data set)	16
3.3 Array analysis of synthetics II (AM data set)	22
4 Timing accuracy issues	26
4.1 Timing accuracy of data acquisition systems	26
4.2 Influence of seismometer calibration uncertainties	27
5 Comparison to examples for real data sets	35
5.1 Baviere (Liege, Belgium)	35
5.2 TST (Euro-Seistest, Volvi graben, Greece)	36
6 Conclusions	38
Acknowledgements	39
References:	40
Appendix 1 to Appendix 3 (in separate pages)	

List of Contributors

(in alphabetical order after the last name)

Fortunat Kind	ETHZ, Zurich, Switzerland
Andreas Köhler	IGUP, Potsdam, Germany
Matthias Ohrnberger	IGUP, Potsdam, Germany
Gudrun Richter	IGUP, Potsdam, Germany
Alexandros Savvaidis	ITSAK, Thessaloniki, Greece
Frank Scherbaum	IGUP, Potsdam, Germany
Daniel Vollmer	IGUP, Potsdam, Germany
Marc Wathelet	LGIT, Grenoble, France

Project coordinator:	Pierre-Yves Bard
Task B Leader, WP05 Leader:	Frank Scherbaum

Summary

In the following we report on the preliminary 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 05 (WP05 – Instrumental layout for array measurements). Within the context of this work package the dependence of the array performance (for phase velocity determination) on the experimental conditions (array geometry, aperture, number of sensors, sensor types, and timing accuracy) has been assessed.

We have evaluated the influence of the array geometry on the determination of dispersion curves by numerical simulation of the noise-free array response for several array geometries and apertures to situations of multiple plane wave arrivals. Using a synthetic waveform data set we have evaluated the capability to derive dispersion curves with standard array analysis techniques from the synthetic wavefield consisting of 100% Rayleigh type surface waves. Several source configurations have been compared for a realistic geological subsurface structure (generic deep basin situation).

The results of these numerical simulations show, that in general (especially for deeper sediment sites) it is not possible to setup a single best array configuration which allows to determine unbiased phase velocity curves for the complete frequency band of interest. The separation of simultaneously arriving wave fronts from distinct directions, or in presence of higher mode surface waves with distinct phase velocities, can be only achieved with a high number of sensors and large array apertures (~ 1 km). Comparing the successful analysis results of real measurements from various sites with the previously simulated results suggests that in nature we often encounter an ambient noise wavefield which shows more favourable characteristics than the general case investigated in the numerical simulations. In most real data examples we find (at least for a confined frequency band) a dominant direction of wave propagation and little influence of higher mode surface waves, which allows to derive a qualitative acceptable dispersion curve from the data.

In order to extend the valid frequency band for the dispersion curve analysis with a limited number of stations, we suggest the use of a spatially adaptive deployment scheme as optimal deployment strategy for ambient vibration array measurements. Starting with small inter-station distances and small apertures to analyse the shallowest part of the site will provide a good approximation of the shallow shear velocity (v_{S30}). This information can then be used for re-deploying the array configuration with subsequently larger apertures and inter-station distances. The analysis is then iteratively carried out for an adequate narrower frequency band of interest which is adjusted to the given array size and next larger target wavelength.

1 Introduction

The objectives of WP05 of the SESAME project are to develop and assess the dependence of the array performance for phase velocity determination on the experimental conditions. Of main interest within this context are the array geometry, array aperture, the number of sensors deployed, the type of sensors which is used and the timing accuracy of the measurements. If we can give an answer to the above question it will be possible to derive a strategy for optimum array deployments for measuring the dispersion characteristics of the ambient noise surface wavefield for site effect estimation.

Without taking into account neither the nature of ambient noise sources nor the specific wave propagation effects, we have started to address the given problem by just considering the resolution capabilities of a specific array configuration in terms of the array response function. The numerical simulation of multiple plane wave arrivals is presented in section 2 of this report. Taking further into account the distribution of ambient noise sources and wave propagation effects for an interesting geologic situation (generic deep basin structure) we have worked on the array analysis of synthetic waveforms to determine the array performance for a more realistic situation in section 3.

The influence of timing accuracies due to instrumental effects is addressed in section 4 whereas a comparison with the results of dispersion curve analysis for real data examples is given in section 5. In section 6 we finally have concluded the results and give a suggestion for a deployment strategy for the measurement of dispersion characteristics with ambient vibration array recordings.

2 Resolution test for noise-free multiple plane waves and arbitrary array configurations

One of the currently implemented analysis schemes for the determination of dispersion curves from ambient vibration array measurements at the Institute of Geosciences, University of Potsdam, consists in a standard semblance-based beamforming algorithm in the frequency domain (Kvaerna and Ringdahl, 1986). For a given array data set the beamforming algorithm is evaluated in a sliding window analysis for a set of narrow frequency bands. Subsequent frequency bands are spaced logarithmically starting from the lowest frequency band of interest. The bandwidth of the frequency band is given as a fraction F of the respective centre frequency f_c characterising the analysed frequency band. The window length which is chosen for the sliding window analysis is selected according to the analysed frequency band and is calculated by L times the centre period thus allowing a constant time-bandwidth product processing scheme. Both the values for F and L are user selectable and are typically set to 0.1 and 10, respectively (compare deliverable D05.06 of WP06).

In order to determine the resolution capability for a specific array configuration and the above introduced analysis scheme, we compute the theoretical array response function for situations of multiple plane wave arrivals without any additional noise contribution. Neither source nor path or instrument effects are considered, i.e. the spectral shape is flat. By doing so, the

results will provide conclusions with respect to the impulse response of the array geometry acting as a filter in the frequency wavenumber space.

Given a list of I coordinates \mathbf{r}_i for an array configuration and a list of N plane wave arrivals, parameterized via the absolute slowness values s_n , propagation azimuth α_n , and signal strength (amplitudes) A_n , we can calculate the normalized array response function for a narrow-banded frequency interval consisting of K discrete frequencies \mathbf{w}_k and target slowness \mathbf{s} as:

$$P(f_c, \bar{s}) = \frac{1}{KI \sum_n A_n} \sum_{n=1}^N \sum_{k=1}^K \sum_{i=1}^I A_n \exp(w_k (\bar{s} - \bar{s}_n) \bar{r}_i)$$

The n -th plane wave front arrival is expressed as:

$$\bar{s}_n = s_n \cos(\alpha_n) \bar{e}_x + s_n \sin(\alpha_n) \bar{e}_y$$

and the array response is evaluated for a discrete cartesian grid of target slownesses:

$$\bar{s} = (-s_{\max} + l\Delta s) \bar{e}_x + (-s_{\max} + m\Delta s) \bar{e}_y$$

with indices l, m 0 to 100, $s_{\max} = 5$ s/km and Δs set to 0.05 s/km. The narrow frequency bands are chosen equivalent to the previously described analysis scheme for real array data. The lowest centre frequency considered is 0.2 Hz and 25 logarithmically spaced frequency bands are calculated. Examples of the resulting normalized array response functions for a single incident plane wave are depicted for different array geometries and frequency bands in Figure 2-1 and Figure 2-2, respectively.

In Figure 2-1 we show slowness maps for two frequency bands (middle: $f_c = 0.57$ Hz, bottom: $f_c = 1.22$ Hz, $F=0.1 \Rightarrow$ bandwidth 10% of f_c) and an almost ideal array configuration consisting of 100 randomly distributed stations with maximal apertures between 0.2 km and 1.0 km in steps of 0.2 km (array configurations are shown on top from left to right). The maximum of $P(f_c, \mathbf{s})$ is depicted by a star symbol and the slowness vector of the incident single plane wave is shown as vector. The absolute slowness values of the incident plane waves are chosen according to a realistic dispersion curve calculated for a generic deep basin model as is found e.g. in the Lower Rhine Embayment in NW-Germany (Brüstle & Stange, 1999). The azimuth of the simulated plane wave front has been selected randomly.

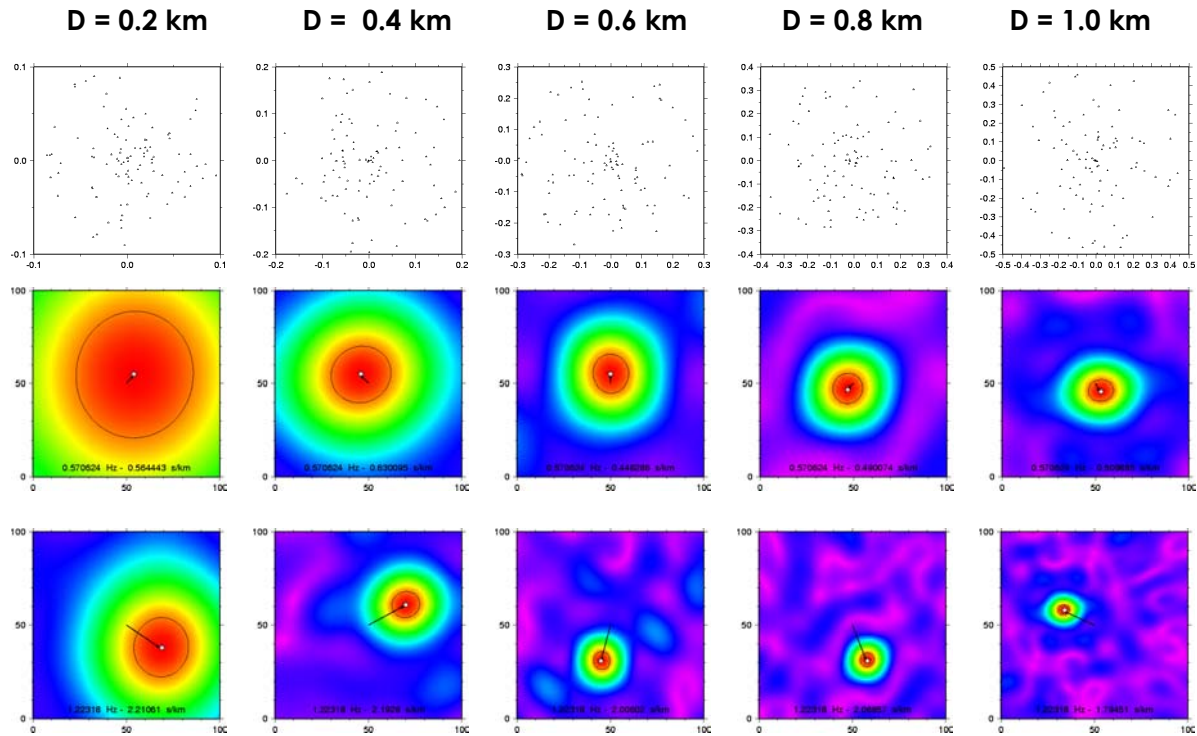


Figure 2-1 Array transfer functions for two different frequency bands (middle and bottom rows) and array apertures (0.2 – 1.0 km left to right on top). Array configurations (shown in top row) consist of 100 stations which are randomly distributed (in polar coordinates). The good spatial sampling which is achieved by this huge number of stations results in an array transfer function with no significant side lobes. The larger the aperture, the narrower the main lobe, the better the separation of multiple plane wave arrivals in the wavenumber space. The centre frequencies of the evaluated frequency bands and the estimated phase velocities of the maximum in $P(f,s)$ are given for each slowness map.

Figure 2-2 shows the same scenario for a (more realistic) array configuration consisting of 13 stations arranged in two hexagons around a central station. The ratio of the radii of the inner and outer hexagons is one to three. The apertures are again ranging from 0.2 km to 1.0 km. Comparing Figures 2-1 and 2-2 one can recognize the effect of the superior spatial sampling of the wavefield for the case of the random station distribution with 100 sensors (Figure 2-1). No significant side lobes can be observed in the array response, whereas for the less dense and furthermore symmetric station configuration (Figure 2-2) we find a symmetric pattern of considerably high side lobes in the slowness maps.

The effect of the array aperture on the wavenumber resolution capabilities can be observed equally well from this simple example. The larger the array aperture, the narrower gets the observed main lobe allowing for a better resolution of multiple plane wave arrivals with similar slowness vectors. A theoretical expression between slowness resolution and array aperture for two incident plane wave arrivals has been given e.g. by Woods and Lintz (1973) for both correlated and uncorrelated signals. Focussing on the objective of determining dispersion curves from ambient noise measurements, we have investigated the resolution capabilities of array configurations for multiple plane wave arrivals which are shown in the following sections.

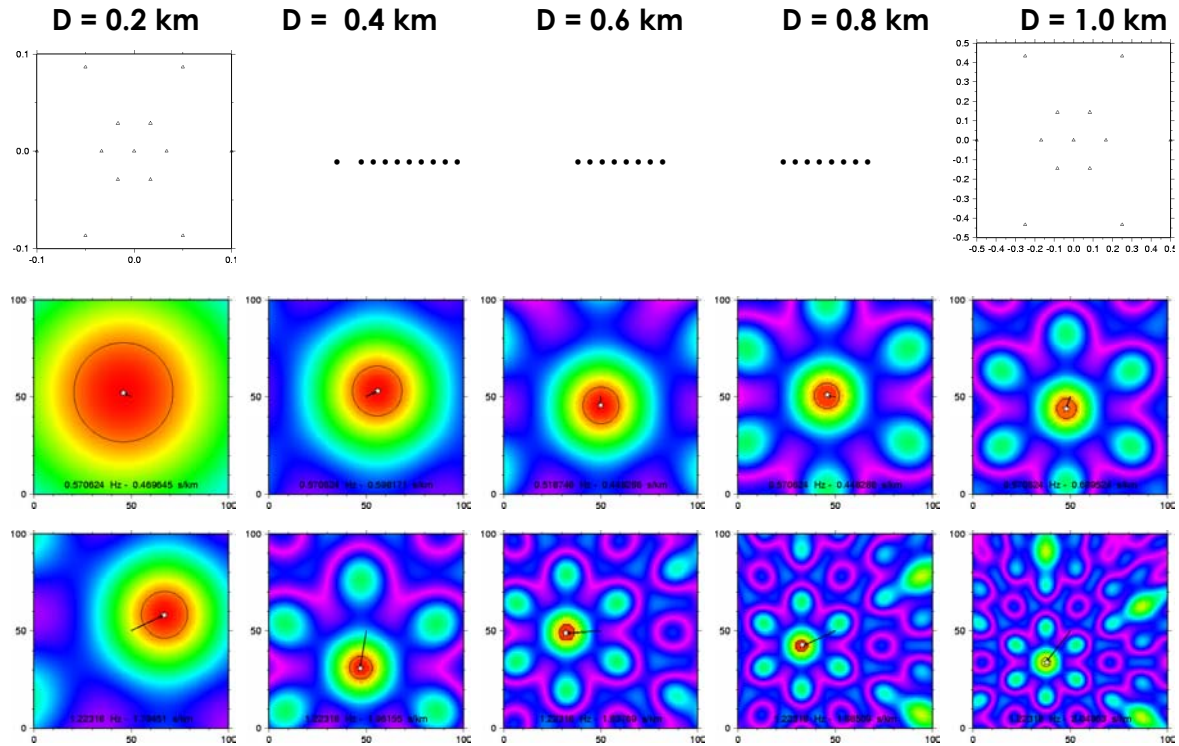


Figure 2-2 As Figure 2-1 for a 13 station configuration arranged as two hexagons surrounding a central station. The reduced spatial sampling of this configuration compared to the 100 station setup and the highly symmetric geometry lead to significant side lobes symmetrically distributed in the resulting array response function.

2.1 Single plane wave arrival

We introduce the procedure that has been followed for investigating the resolution of a given array geometry starting with the simplest situation: a single plane wave arrival. In a first step we define a number of logarithmically spaced frequency bands with 50% overlap ($F=0.1$, $f_c(n+1) = 1.1 \cdot f_c(n)$). Using the theoretical phase velocity curve for fundamental mode Rayleigh wave for the generic deep basin model (Brüstle & Stange, 1999), we determine the range of possible phase velocities $[c_{\min}, c_{\max}]$ which lie within the selected frequency band $[0.9f_c(n), 1.1f_c(n)]$. For each frequency band we simulate a single plane wave front with its respective phase velocity taken randomly from the range $[c_{\min}, c_{\max}]$ and arbitrary azimuth (randomly selected $[0, 2\pi]$). After computing the array response we determine the maximum within the discrete slowness map as the best phase velocity estimate.

In Figure 2-3 we have displayed the result of this numerical simulation for four different array configurations: 100 station randomly distributed station geometry and a 13 station hexagonal station configuration with apertures of 0.2 km and 1.0 km (for the geometries compare Figure 2-1 and Figure 2-2). For comparison we have plotted the results of this simulation both proportional to apparent phase velocity as well as proportional to slowness. The red curve depicts the theoretical fundamental mode Rayleigh wave dispersion curve. Boxes are drawn to indicate the frequency bands chosen and the compatible phase velocity range for the frequency band under consideration. The blue arrows give the direction of the simulated plane wave and the ordinate of the vector's origin is given by the absolute value of phase velocity taken for the simulation. The length of the vectors is proportional to the amplitude of the

respective plane wave arrival. The estimates of the phase velocities which have been determined for each frequency band as the maximum of array response function within the discrete slowness grid are displayed as red circles.

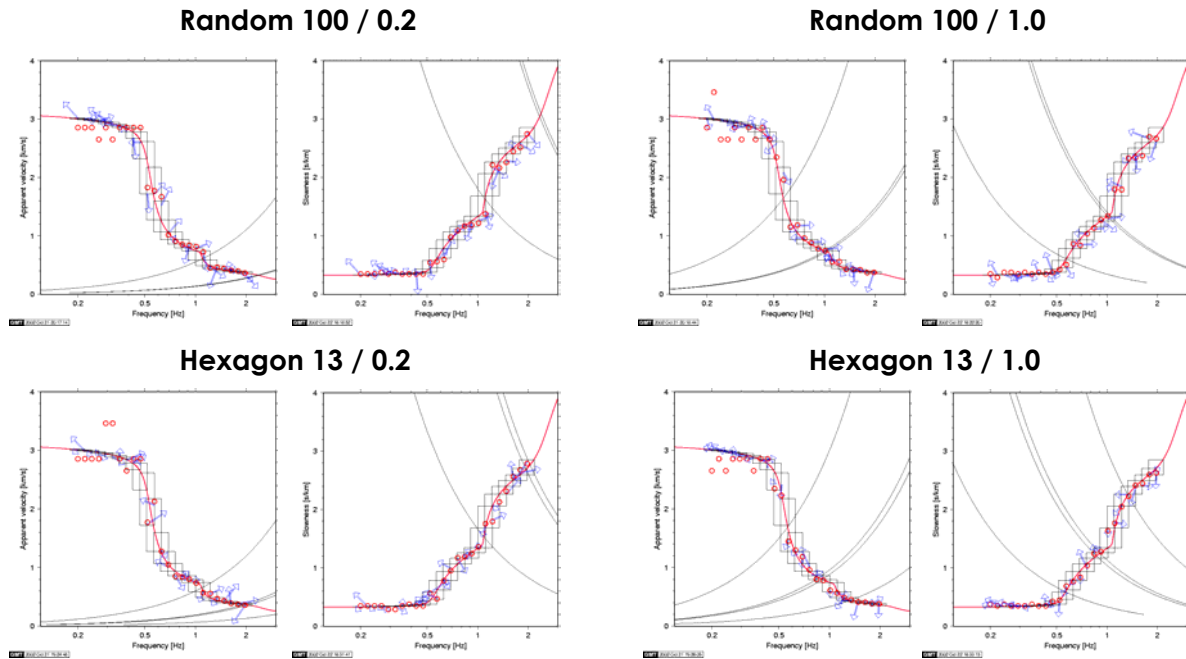


Figure 2-3 Result of numerical simulation for four different array configurations: 100 station randomly distributed station geometry (top) and a 13 station hexagonal station configuration (bottom) with apertures of 0.2 km (left) and 1.0 km (right). Results are plotted both linear in apparent phase velocity as well as proportional to slowness. Red curves depict the theoretical Rayleigh wave dispersion curve for the fundamental mode for the LRE model. Boxes indicate the selected frequency bands and compatible phase velocity range within the frequency band under consideration. Blue arrows give the direction of phase velocity. The “observed” phase velocity which has been determined from for each frequency band as the maximum of array response function in the discrete slowness grid are displayed as red circles. Thin black lines indicate approximated aliasing curves for mean, median and smallest inter station distances and an approximated curve relating the penetration depth and the array aperture ($\lambda/3$ criteria). For more details and discussion of results, please see the text.

The thin black lines in the figures indicate an approximation of the aliasing condition. The spatial aliasing condition can be given exactly only for a given direction of wave propagation $\vec{\xi}$ and equally spaced sensors $d_{\min}(\vec{\xi})$ as:

$$k_{\max}(\vec{\xi}) = \frac{2\pi}{\lambda_{\min}(\vec{\xi})} = \frac{\pi}{d_{\min}(\vec{\xi})}.$$

The aliasing condition is direction dependent and for general geometries not easily given. Therefore we evaluate the above equation for the minimum of distances between any station pair in the array geometry as well as for the median and mean of all distances. We expect to observe spatial aliasing effects somewhere in the region between those curves.

Additionally we have plotted the so-called $\lambda/3$ criteria (e.g. Tokimatsu, 1997), which correlates the maximum aperture d_{\max} with the estimated penetration depth of the largest effective wavelength λ_{\max} of a surface wave like:

$$d_{\max} = \frac{\lambda_{\max}}{3} = \frac{2\pi}{3k_{\min}} \Leftrightarrow k_{\min} = \frac{2\pi}{3d_{\max}}.$$

Although most of the phase velocity (slowness) estimates match the simulated input value, it can be seen from Figure 2-3 that the evaluation of the array response function on a discrete slowness grid can lead to erroneous estimates of the true slowness input. The maximum error due to the cartesian discretization of the slowness space is $\sqrt{2} \cdot \Delta s$ and scales linearly with slowness, but hyperbolic in terms of phase velocity ($c=1/s$). Therefore we can observe in the plots proportional to phase velocity larger errors for lower frequency bands (high phase velocities \Leftrightarrow small slowness values). The discretization error can have either sign, deviating to higher or lower phase velocities (slowness values), depending on the direction of incident plane wave front and the discretization step Δs .

It is possible to reduce the discretization error of the observed phase velocity estimates due to the sampling in the slowness space by applying a densely sampled grid layout which is equidistantly sampled in phase velocity. However, we do not recommend this procedure for the following reason. For common array analysis procedures, the physical quantities to be measured are the time lags of arriving wavefronts between the individual sensors. Measurement errors can be caused by effects of the available instrumentation (i.e. seismometer phase delays and limitation of clock accuracies, see section 4) and by limitations of the applied algorithms according to digital signal processing principles (dependent on the processing scheme and experimental parameters). Again the measurement errors are proportional to time. Hence the achievable timing accuracies of the experiment are linearly related to the uncertainty of the slowness estimates, but inversely related to the uncertainties of the phase velocity estimates. Thus, applying a densely sampled grid layout proportional to apparent velocity for the computation of array responses implicitly assumes high measurement accuracies for closely spaced higher phase velocities. For the same reason we neither recommend to present the results of array analysis proportional to phase velocity, as for the reader it is not easily realized that the uncertainty of the estimates scales hyperbolic with the obtained phase velocity estimate.

From practical considerations, we suggest to use grid layouts proportional to slowness. The discretization step Δs should be chosen depending on a realistic assumption of timing inaccuracies Δt both stemming from effects of the available instrumentation (i.e. seismometer phase delays and clock accuracies, see section 4) and the capability to resolve time differences of plane wave arrivals of a certain wavelength between individual sensors considering underlying digital signal processing aspects (processing scheme). Given an average inter-station distance $\langle d_{ij} \rangle$ within the array configuration and measurement uncertainty Δt , the discretization step for the slowness grid should be chosen as $\Delta s \geq \Delta t / \langle d_{ij} \rangle$.

2.2 Three plane wave arrivals (varying azimuth)

For the analysis of ambient vibration data we assume hereafter that the wavefield can be described as a superposition of multiple plane wave arrivals, propagating with phase velocities according to a site specific (structural dependent) surface wave dispersion relation. The propagation directions of incident plane waves are assumed to be equally distributed in azimuth. We neither take into account source or path effects, the existence of curved wave fronts nor the distortion of the recordings by measuring devices (seismometers). I.e. the amplitude spectra are flat in the complete frequency band of interest.

Figure 2-4 displays the results from the numerical simulation of array response functions determined for three incident plane waves crossing the respective array configurations with fundamental mode Rayleigh wave phase velocities and varying azimuth. In the upper part of Figure 2-4, the results are presented for the 100 station random array configurations with increasing aperture whereas in the lower part the results for the 13 station hexagonal array layout are shown. It is clearly observed, that the underlying phase velocity curve can only be partly recovered in this simulation. Comparing the estimated dispersion curves for different apertures it can be noted, that the frequency band for which correct slowness estimates are obtained is extended to lower frequencies for larger apertures.

This observation can be explained by the limited resolution capabilities of the investigated array configurations as is demonstrated in Figure 2-5. The superposition of the array responses for the individual plane wave arrivals can lead to a biased position of the observed maximum in the slowness grid. For the 100 station randomly distributed array geometries we observe the superposition of the main lobes of the array response functions for the individual plane waves. The slowness value estimates therefore are biased to lower values (see Figure 2-4, top). However, for less complete spatial sampling, i.e. for the symmetric 13 station hexagonal distribution (as shown in Figure 2-5), it is also possible to observe the superposition of side lobes of the array responses in case of an unfavourable azimuthal distribution of the incident individual plane waves. Thus the biased slowness estimate may also be larger than the input slowness value.

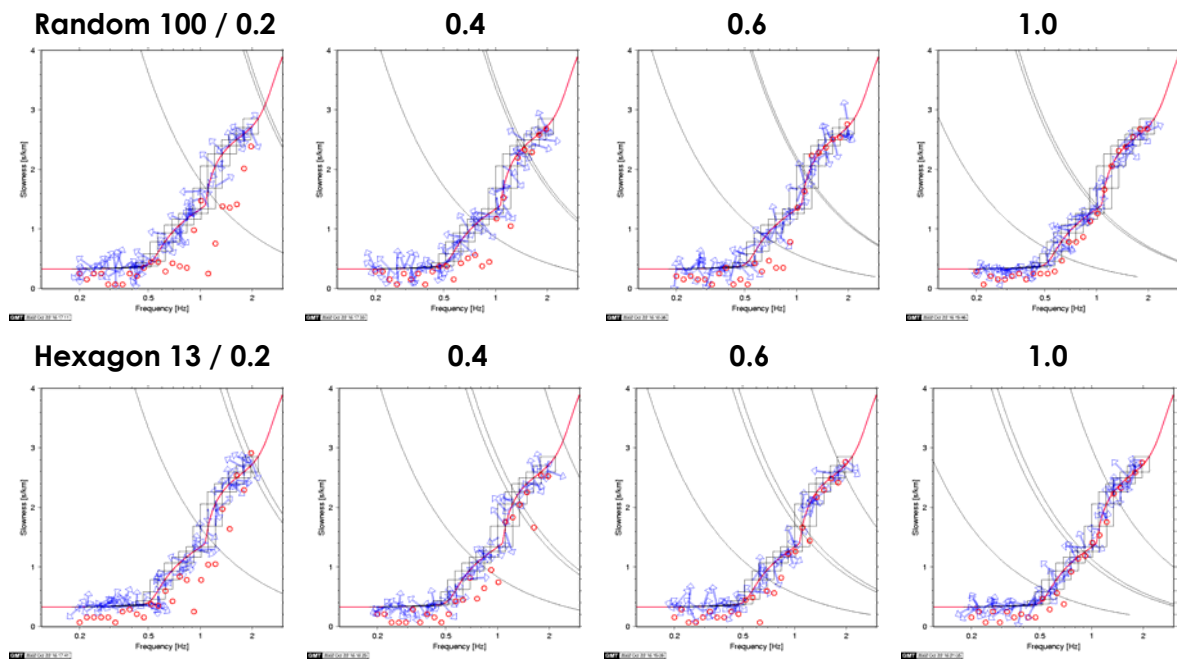


Figure 2-4 Recovery of dispersion curve from simulation of three plane wave arrivals crossing various array configurations. Top: Results for 100 station random station distribution are shown for apertures of 0.2, 0.4, 0.6 and 1.0 km. Bottom: Results for 13 station hexagonal station geometry.

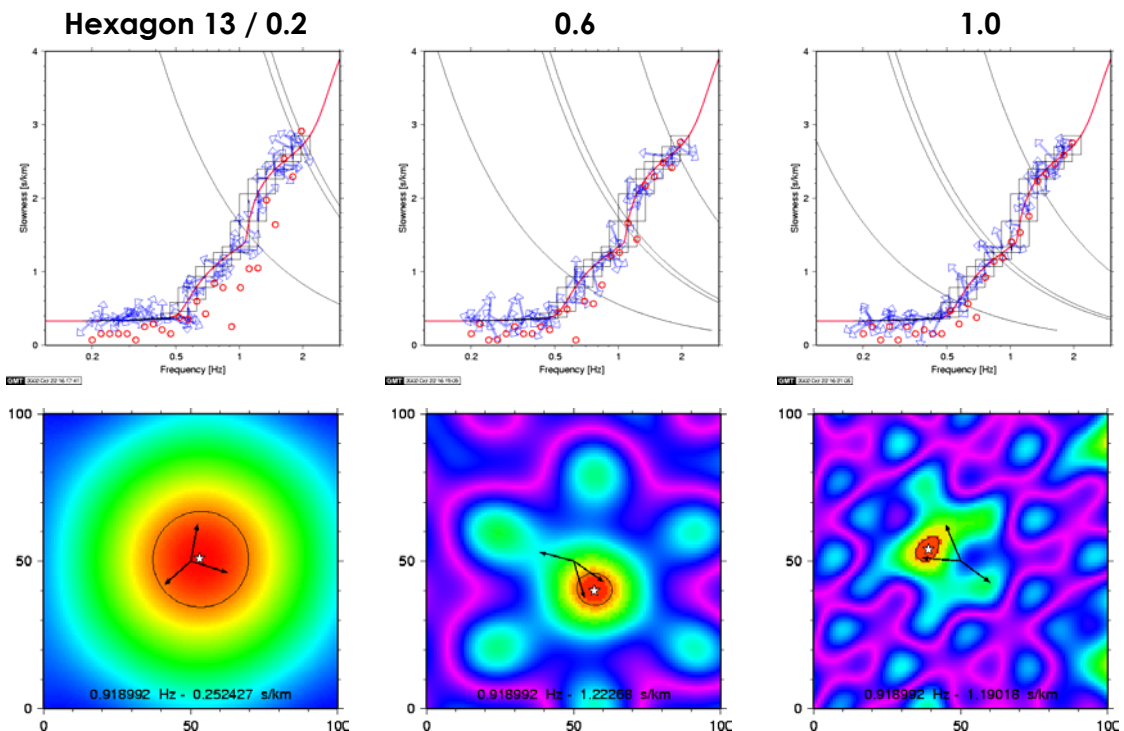


Figure 2-5 Bias of phase velocity estimates by superposition of array response functions for multiple plane wave arrivals. In this example three plane wave arrivals with similar apparent velocity, but differing azimuths are superimposed. Top: extension of valid frequency band which allows to determine correct slowness values for increasing aperture of array configuration. Bottom: examples of superimposed array responses for one frequency band and increasing array aperture.

2.3 Three plane wave arrival (varying slowness)

As additional test we have selected a three plane wave arrival situation, where the azimuth of incident plane waves is kept constant, but the absolute slowness of the incident plane waves is varied according to the phase velocity ranges of all higher modes from the theoretical dispersion curves. As before in section 2.2, it is possible to explain the bias of the slowness estimates by the superposition of array responses. However, even in the case of large aperture and good spatial sampling of the wavefield, an additional complication is given from the fact, that the maximum, which is estimated from the superposed array response function, gives at best the phase velocity of either the fundamental mode or one of the higher modes, respectively. As we lack the knowledge about the energy contribution of the individual mode branches, we are not able to connect the phase velocity estimate automatically to a certain mode branch.

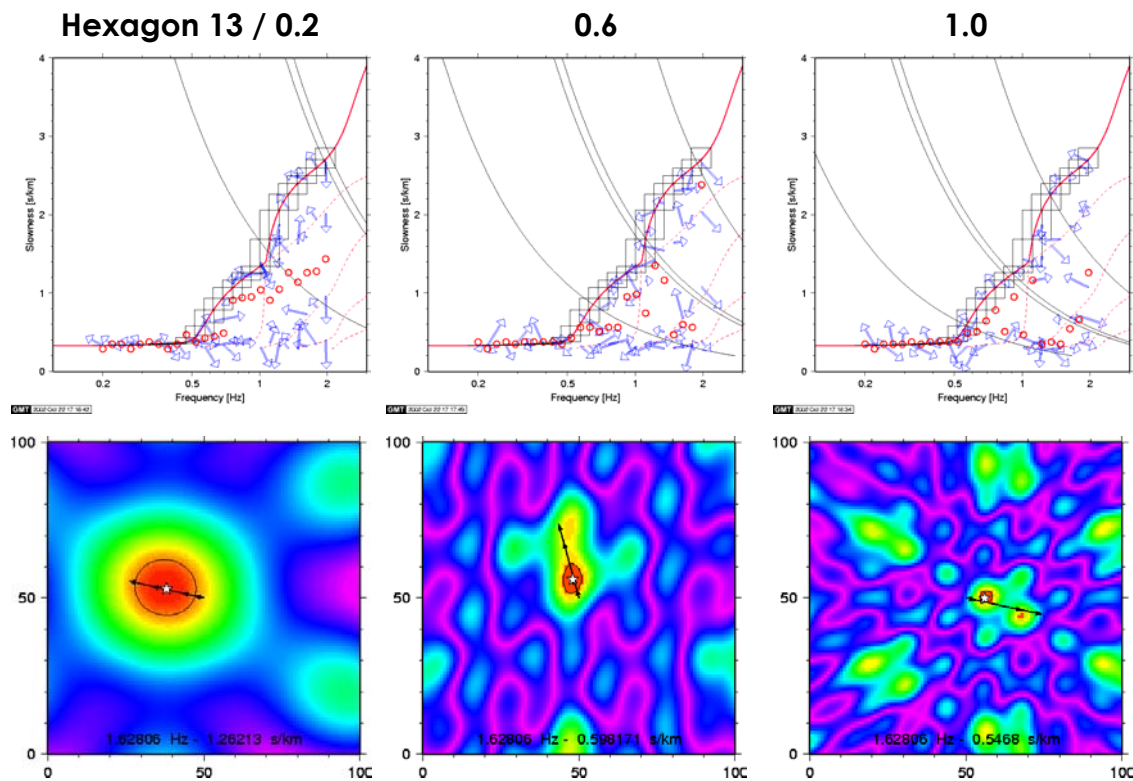


Figure 2-6 Top: Example of the simulation result for three incident plane waves from the same direction but distinct phase velocities (according to fundamental and higher modes for theoretical dispersion curve). Example is given for the 13 station hexagonal station setup and three apertures. **Bottom:** Examples of array response functions for one frequency band. Resolution capabilities increase with increasing aperture. However, even in the case of sufficient resolution, the problem of interpretation which mode is observed remains.

From the numerical simulation of the array response function for multiple plane wave arrivals we can make the following preliminary conclusions. For the determination of dispersion curves from array measurements of ambient vibrations, it is necessary to sample the wavefield as dense as possible in the spatial domain. Larger array apertures are to be favoured in terms of achievable wavenumber resolution capabilities. An extension of the valid frequency band for obtaining reasonable estimates of a dispersion curve can be obtained by improving the

resolution capabilities of the array configuration, hence increasing the aperture while maintaining inter station distances as small as possible to prevent spatial aliasing for higher wavenumbers. However, this suggests the use of a high number of sensors for larger array apertures, a requirement which can not be fulfilled considering the array analysis of the ambient noise wavefield as an economically reasonable method for site investigations.

Using a manageable number of stations (e.g. 13 stations as in the numerical examples) the enlargement of the array aperture will be still not easily realized within urban areas. For sites where fast lateral changes of the subsurface geology can be expected, the enlargement of array apertures may not even be desired, as 2-D effects of wave propagation will at least introduce additional complications for the interpretation of the obtained results. Furthermore, whether -or under which conditions- it is possible to observe a coherent ambient noise wavefield over larger distances remains an open question.

Recalling that the numerical simulation shown here, evaluates the traditional array response function on basis of an expression derived from the standard beamforming approach, an improvement of the resolution capabilities can be achieved using alternative f-k algorithms, i.e. HRFK (Capon, 1969) or MUSIC (Schmidt, 1981, 1986). The drawback of these approaches lies in the reduced robustness of estimates for ambient noise data (e.g. Zywicki, 1999).

We suggest the use of a spatial sampling strategy which is optimized for a narrower wavelength window, but then allows using a smaller number of sensors for the measurements. In order to cover the complete frequency (wavelength) range of interest for the determination of dispersion curves, it will be necessary to apply an iterative measurement approach with successive enlarging of the array aperture and inter-sensor distances.

3 Resolution test with synthetic waveforms

The resolution test described in section 2 is a very simplified approach. As previously noted it does include neither source or path effects nor curved wave fronts. In order to allow for a more realistic resolution test of array configurations for the determination of dispersion curves from ambient noise measurements, we have performed array analysis for several sets of synthetic waveform data. The objective of this study is to investigate the influence of both wave propagation effects as well as the use of different analysis procedures on the results of the estimated dispersion curve result.

3.1 Preparation of synthetic waveform data

In a first step, we have calculated a set of synthetic waveforms for a station geometry consisting in 181 sensors locations and a generic deep sediment basin velocity model (Brüstle & Stange, 1999) on top of a regional crustal model derived for the Lower Rhine Embayment in NW-Germany (Reamer and Hinzen, 2001). Both the station configuration and the velocity model are shown in Figure 3-1. The synthetic seismograms have been calculated using the modal summation code of Herrmann (1996).

In order to make the computation feasible for an arbitrary number of source configurations (source location distribution), we have first calculated synthetic seismograms for a 10 km station profile with 1 m station spacing. As source we have selected a vertical single force located at the surface with an impulsive source time function. The computed wavefield is therefore radial symmetric and only Rayleigh waves are excited. Two data sets of waveform synthetics have been created. The first data set has been restricted to contain just the fundamental mode Rayleigh wave (FM data set), whereas the second set includes also all higher modes (AM data set).

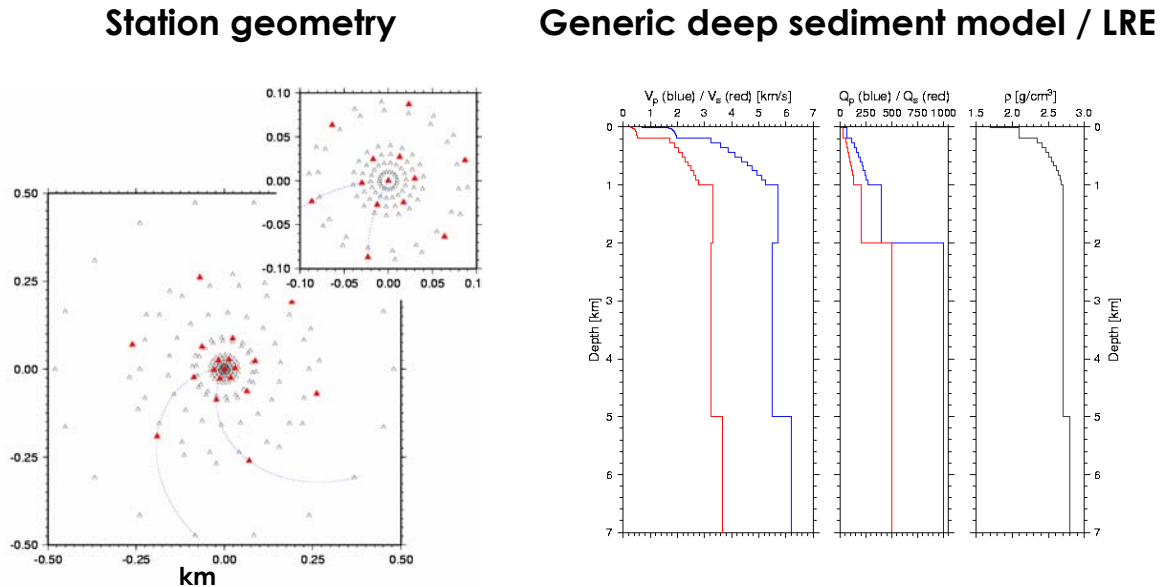


Figure 3-1 Left: Station geometry used for the calculation of synthetic waveforms. 181 stations are distributed in a spiral shaped configuration. 18 stations are placed on each of the 10 rings surrounding a central station. The radii chosen range from 10 m to 500 m. rings are tilted against each other to obtain good azimuthal sampling. The red triangles indicate the 19 station sub-geometry selected for the following resolution tests. Right: Velocity model used for the generation of synthetic waveforms. The upper part ($h < 1\text{ km}$) follows a realization of the generic deep basin model (Brüstle & Stange, 1999) which fits well for a situation in the Lower Rhine Embayment in NW-Germany (Scherbaum et al., 2002). For the deeper part of this model, we have chosen a crustal model from Reamer and Hinzen (2001) derived from travel time curves for the same region.

A complete 3-component waveform data set has been obtained for several source configurations as follows. For each individual source, all source-receiver distances and azimuths have been computed. The previously calculated profile seismograms corresponding to the closest matching distance for the respective source-receiver pair is selected and summed to the output trace for the corresponding receiver and component (north and east components are projected according to the computed azimuth). The individual sources are excited several times with random time shifts and variable amplitudes.

We have defined three source configurations (Figure 3-2) for testing the capability to derive dispersion curves from the synthetic data sets. First we created a data set for a single source location ("Single") which is excited 50 times with random time shifts and amplitudes (min/max ratio equals 1/10). The source-receiver distances lie in the range of 4 to 5 km. The second data set ("City") consists of 150 source locations of which 100 are distributed in a restricted rectangular area around 4 to 6 km from the array location. The remaining 50 sources have been distributed randomly around the array not exceeding 9.5 km from the array centre.

Each of the 150 sources is excited between 1 and 4 times with varying amplitudes. Finally, the last data set (“Random”) consists in 150 randomly distributed source locations not exceeding a distance of 9.5 km from the array centre. The number of source excitations for each individual source has been selected randomly between 1 and 4 with varying amplitudes.

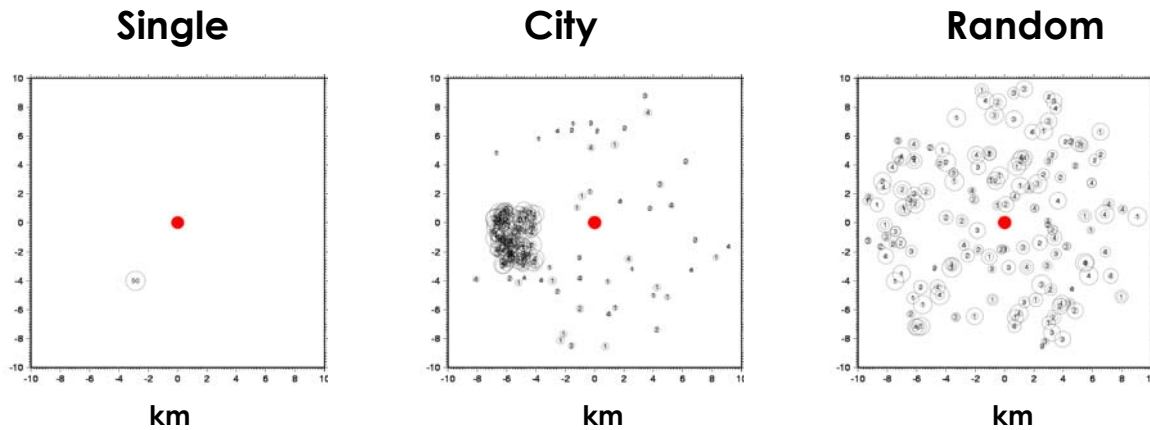


Figure 3-2 Source configurations used for constructing synthetic data sets. The open circles indicate the position of the source, the number gives the number of excitations for the respective source and the source strength is given by the circle sizes (minimal and maximal strengths). The filled red circle in the centre covers the area of the source configuration shown in Figure 3-1. Left: Single source location excited 50 times randomly in time with varying amplitudes. Source↔receiver distance ca. 4 km. Middle: 150 source locations excited between 1 and 4 times randomly in time and varying amplitudes. 100 source locations are distributed in a restricted rectangular region in ca. 4 to 6 km distance to the array centre. The remaining 50 source locations have been selected randomly with maximal offset of 9.5 km to the array centre. Right: 150 source locations randomly distributed with maximal distance of 9.5 km to the array centre. Number of random excitations in time is 1 to 4 with varying amplitude.

An example of the resulting seismograms at several receivers for the random source configuration is shown in Figure 3-3. On the left side of Figure 3-3, seismograms are trace normalized, whereas on the right hand side, seismograms are scaled together. For this waveform example coherent wave packets can be clearly observed for all stations.

In the following sections, we show the results of the array analysis obtained for the three source configurations “Single”, “City” and “Random” for both the synthetic waveform data sets FM (fundamental mode only) and AM (all modes). The array analysis has been performed by the software package “cap” (compare 2nd progress report of SESAME) using the database structure of GIANT (Rietbrock and Scherbaum, 1998) and interactive waveform control via PITSA (Scherbaum and Johnson, 1992).

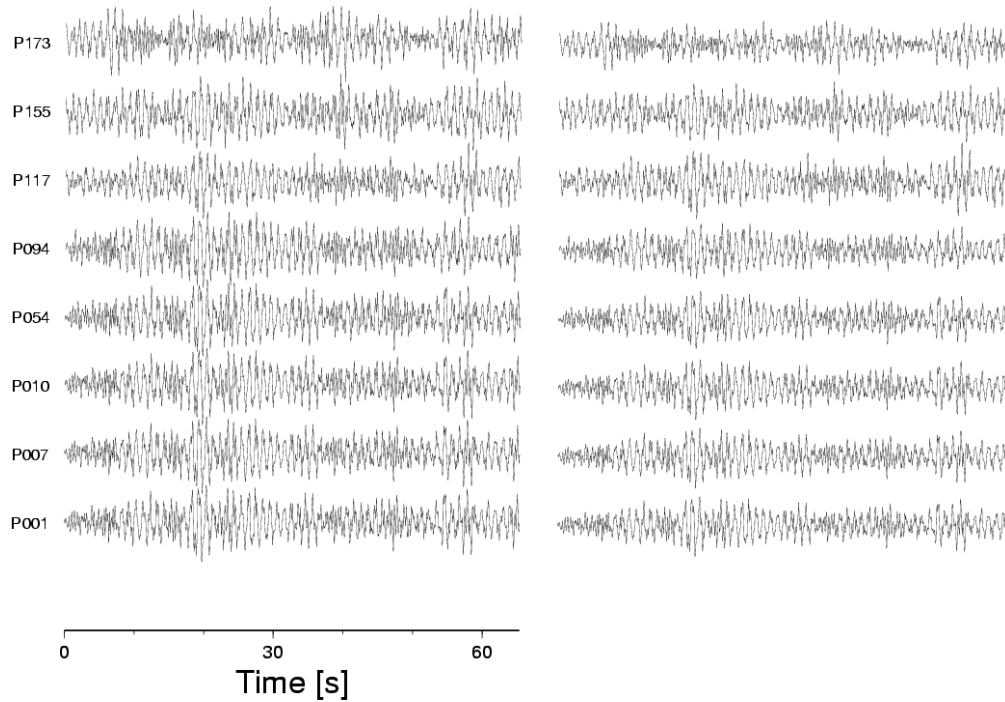


Figure 3-3 Synthetic waveform example for the randomly distributed source configuration (compare Figure 3-2) and 8 receivers. On the left side, seismograms have been plotted trace normalized, whereas on the right side all traces are scaled together. Waveforms are sampled with 50 Hz and total length is 16384 samples (327.68 s).

3.2 Array analysis of synthetics I (FM data set)

For the analysis of the FM data set, we have chosen a 19 station array configuration which consists in three tilted hexagons surrounding a central station (compare red triangles in Figure 3-1). The radii of the hexagons are 30 m, 90 m and 270 m, respectively. Three array methods have been used to analyse the data, a semblance-based beamforming algorithm (further denoted as FK) in frequency domain (e.g. Kvaerna and Ringdahl, 1986), the high-resolution f-k algorithm (HRFK) after Capon (1969) and the modified SPAC (MSPAC, Bettig et al., 2002) method. For all of the methods we use Z-component data only and a window processing scheme as described in section 2. The window lengths and frequency bands have been selected in a way that a constant time-bandwidth product is obtained for all frequency bands in the analysis. Whereas for the FK and MSPAC methods we perform a sliding window analysis with highly overlapping successive windows, for the HRFK analysis the original non-overlapping segment averaging used in Capon's work has been applied for a robust estimate of the spectral covariance matrix.

The results of the FK array analysis is obtained in the following scheme. For the parameter settings used and the time length of the synthetic data sets, we calculate slowness maps for each frequency band and ca. 220 successive time windows. The slowness vector connected to the highest semblance value in each map is automatically selected as best estimate of the slowness vector. We determine for each frequency band the mean, median, standard deviation as well as the 25% and 75% quantil from all slowness estimates obtained for the individual time windows. Furthermore we calculate the mean from only those time windows which exceed 99% of the highest semblance value ever observed in all time windows ("average of

best”). By doing so, we pick the slowness values of the most coherent plane wave arrivals which have been observed in the data. Finally, we have built histograms both for the absolute slowness estimates as well as for the estimated backazimuths from the complete set of slowness vector estimates.

The results of this analysis are presented in Figure 3-4 to Figure 3-6 as follows. For each data set (“Single”, “City”, and “Random”) we display the azimuthal histogram distribution and slowness histogram distribution as a colour map in the background (first two panels from top). On top of the slowness histogram (second panel) we plot first the available a priori information: the theoretical dispersion curve for fundamental and higher Rayleigh wave modes (black lines, solid and dashed) and the approximate aliasing curves and $\lambda/3$ criteria for the selected station configuration as dashed turquoise lines. We have provided an additional scale at the right hand side of this panel indicating the average travel-times of a propagating wavefront between the stations with the slowness value given in the left scale. Overlaid are the mean (black circles) and median (green circles) slowness estimates and the vertical green lines span the region from the 25% to the 75% quantil of the distribution. The red dots show the “average of best” values obtained from the analysis. The three lower panels in the figures show a) the smoothed averaged spectral energy of the array station spectra for Z/N/E component (red/green/blue colours), b) the H/V ratio calculated for each individual station in the array, and c) the averaged H/V ratio for all stations in the array. For b) and c) we have additionally plotted the theoretical ellipticities of the fundamental Rayleigh wave mode.

The HRFK algorithm provides a single slowness map for the complete data window and each frequency band analysed. We pick the highest value within this slowness map automatically as estimate for the absolute slowness and azimuth for the corresponding frequency. The slowness estimates obtained via the HRFK algorithm are displayed as white triangles in the slowness vs. frequency plot (second panel from top) together with the results of the FK-analysis.

For the MSPAC method we have followed the procedure suggested by Bettig et al. (2002). We have computed the correlation values for each station pair, for each time window and frequency band. For station pairs with similar relative distance, we compute then the averaged modified spatial autocorrelation values for each frequency band and “ring radius” (compare Bettig et al., 2002). The radii obtained for the given station configuration are: 30 m, 56 m, 112 m, 156 m, 180 m, 197 m, 244 m, 282 m, 351 m, 468 m, and 540 m. In Figure 3-4 to Figure 3-6 the averaged spatial autocorrelation curves are displayed as black dots in the right subfigure, with increasing radii from bottom to top (30 to 540 m). The subjacent coloured lines indicate the theoretically predicted spatial autocorrelation curves for the dispersion curves of the fundamental and first ten higher Rayleigh wave modes, respectively. On top of the figure we have displayed the average amplitude spectra of the array stations for all components (Z/N/E = red/green/blue) allowing a judgement of energy contribution within the different frequency bands.

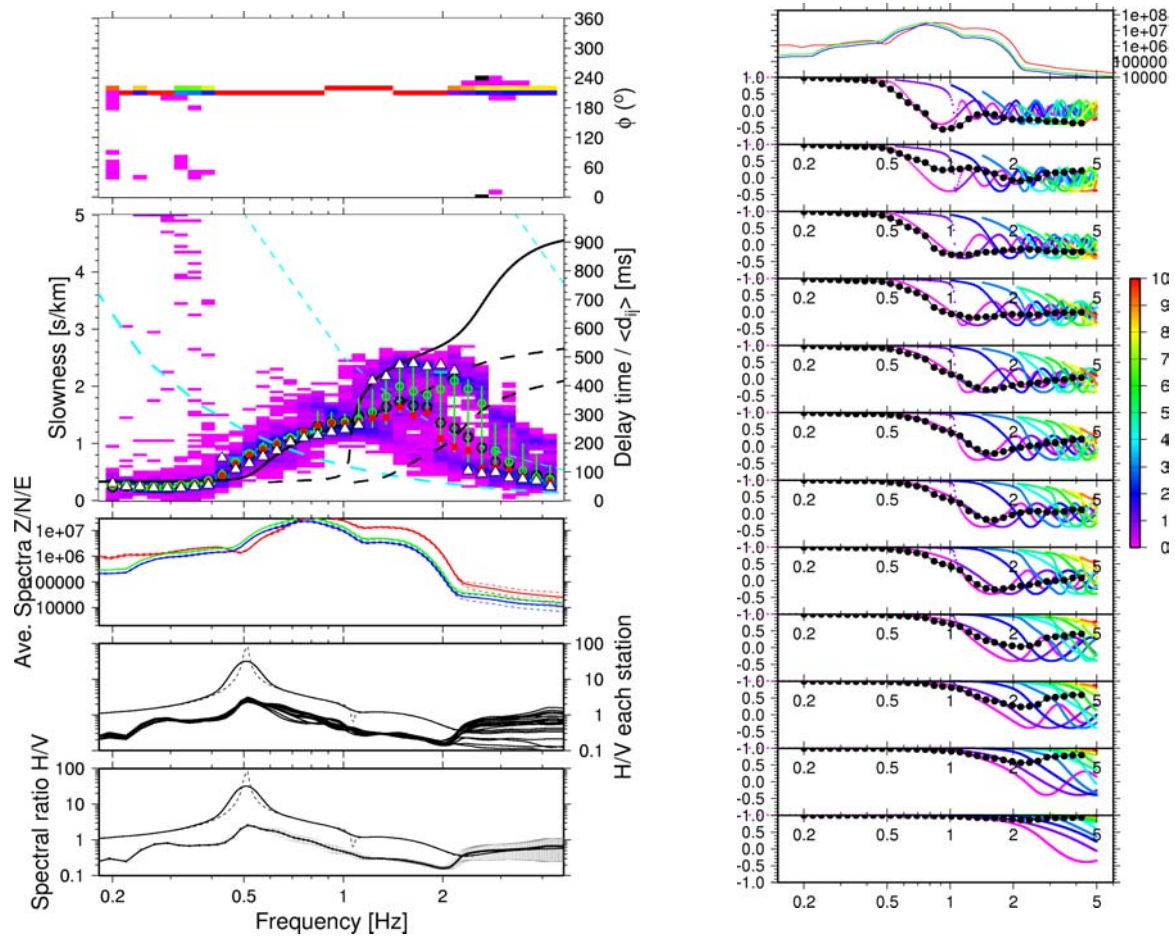


Figure 3-4 Results of array analysis for the « Single » data set, fundamental mode Rayleigh wave only. LEFT: The results of the FK and HRFK are displayed together with the average spectral energy and H/V spectral ratios (both individual stations and average) in the lowest three panels. In the top panel the histogram of the azimuth distribution is shown for all estimates obtained via the FK algorithm. The correct source direction is determined in the whole frequency range in this example. Below the histogram for the absolute slowness value distribution obtained via the FK-algorithm is shown. As in the figures of section 2, the black (solid and dashed) lines indicate the theoretical dispersion curves for the generic deep basin model (fundamental and higher modes, respectively). The approximated aliasing curves and the $\lambda/3$ criteria are shown as dashed turquoise lines. The black and green circles indicate the average and median slowness estimates obtained from the FK analysis, the vertical green lines extend from the 25% to the 75% quantil of the distribution. The red filled dots correspond to the average slowness value obtained for those time windows exceeding 99% of the maximum semblance value observed. Finally, the white triangles indicate the results obtained from the HRFK method. In order to judge the relative error of obtained slowness estimates due to timing inaccuracies, we have provided the scale on the right hand side of this panel. The delay times in [ms] are given for the average inter-station distance $\langle d_{ij} \rangle$. RIGHT: The averaged spatial autocorrelation curves are shown as black dots for selected ring radii (from bottom to top in [m]: 30, 56, 82, 112, 156, 180, 197, 244, 282, 351, 468, 540) together with the theoretically expected autocorrelation curves (coloured dotted lines) for the fundamental and first ten higher modes, respectively.

The results of the FK and HRFK analysis of the “Single” data set are displayed in Figure 3-4. The azimuthal estimate obtained from the FK analysis resembles the true azimuth for the whole frequency band of interest indicating that the selected array configuration shows a good azimuthal resolution capability. The dispersion curve estimate obtained from both FK and HRFK methods show a fairly good match with the theoretical dispersion curves calculated for the input velocity model. However, some deviations are recognized especially for frequencies around 0.5 Hz for both methods, and frequencies above 1 Hz mainly for the FK technique. As this is a single source situation we can not explain the deviations occurring around 0.5 Hz from the superposition of array responses for multiple plane waves. Comparing the spectrum of the vertical component (third panel from top, red curve), we can recognize that for this particular frequency band we encounter a strong drop in spectral energy towards lower frequencies. Therefore we suggest that the bias in slowness estimates is due to deficiencies when estimating the phase spectra for lower energetic parts of the spectrum.

Deviations of the observed dispersion curve for frequencies above 1 Hz for the FK and above 1.6 Hz for the HRFK method can be explained by increasing influence of aliasing effects to higher frequencies. Especially above 2 Hz typical aliasing features are observed in the coloured slowness histogram obtained from the FK analysis following parallel trajectories to the approximated aliasing curves (turquoise dashed line) for the mean station distance. We can clearly observe that the mean and median values of the FK and the HRFK slowness estimates are aligned along these features above a certain frequency threshold.

In case of the FK-algorithm we want to point out an additional problem. The sharp knee in the theoretical fundamental mode dispersion curve (ca. 1 Hz) can not be imaged by the FK technique. This is a consequence of the relatively broad frequency band (10% of f_c) which is used in the processing. In the frequency band from ca. 1.1 to 1.3 Hz the theoretical slowness changes rapidly from ca. 1.3 s/km to 2.2 s/km. It is therefore no surprise that the estimates obtained for the complete frequency range gives at best some intermediate value. In case of the HRFK estimate we calculate the complex Fourier coefficients only for the target frequency (no smoothing over a frequency band is applied here) and therefore the HRFK estimates match closely the theoretical dispersion curve values.

The results of the MSPAC analysis are not as expected. It is obvious, that the estimates of the spatial autocorrelation values match only partly the predicted spatial autocorrelation curves from the fundamental mode phase velocity curve (magenta curve). Currently we suggest that similar to the FK technique, one limiting factor lies in the “broadness” of the frequency band used for estimating the autocorrelation values, which leads to intermediate autocorrelation values for the sharp knee of the fundamental mode Rayleigh wave dispersion curve around 1 Hz. Furthermore, for frequencies above 2 Hz, the spectral amplitudes of the vertical component are about 3 orders of magnitude lower than for the main frequency content of the signal. We suspect that for higher frequencies numerical effects in the forward calculation of the synthetic waveforms can occur and might influence the results of the analysis.

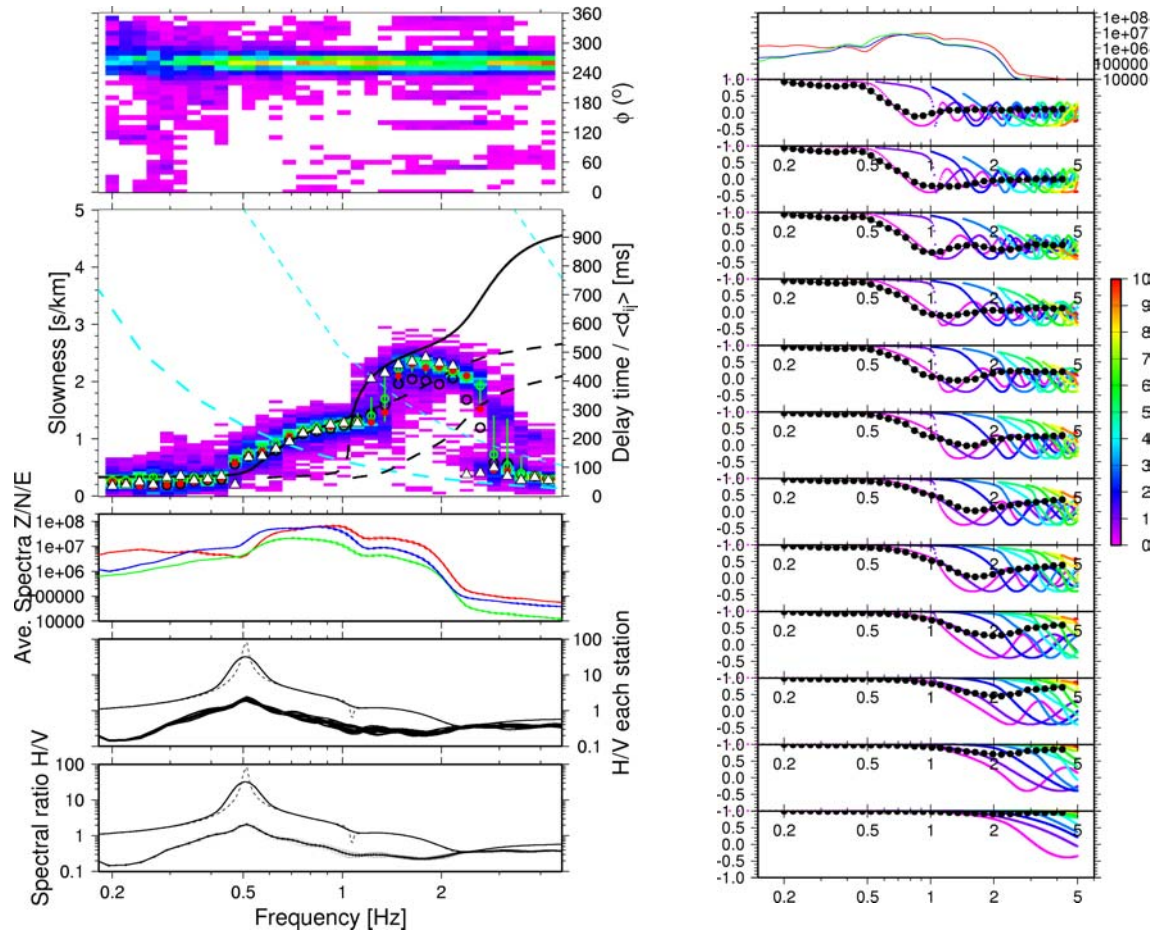
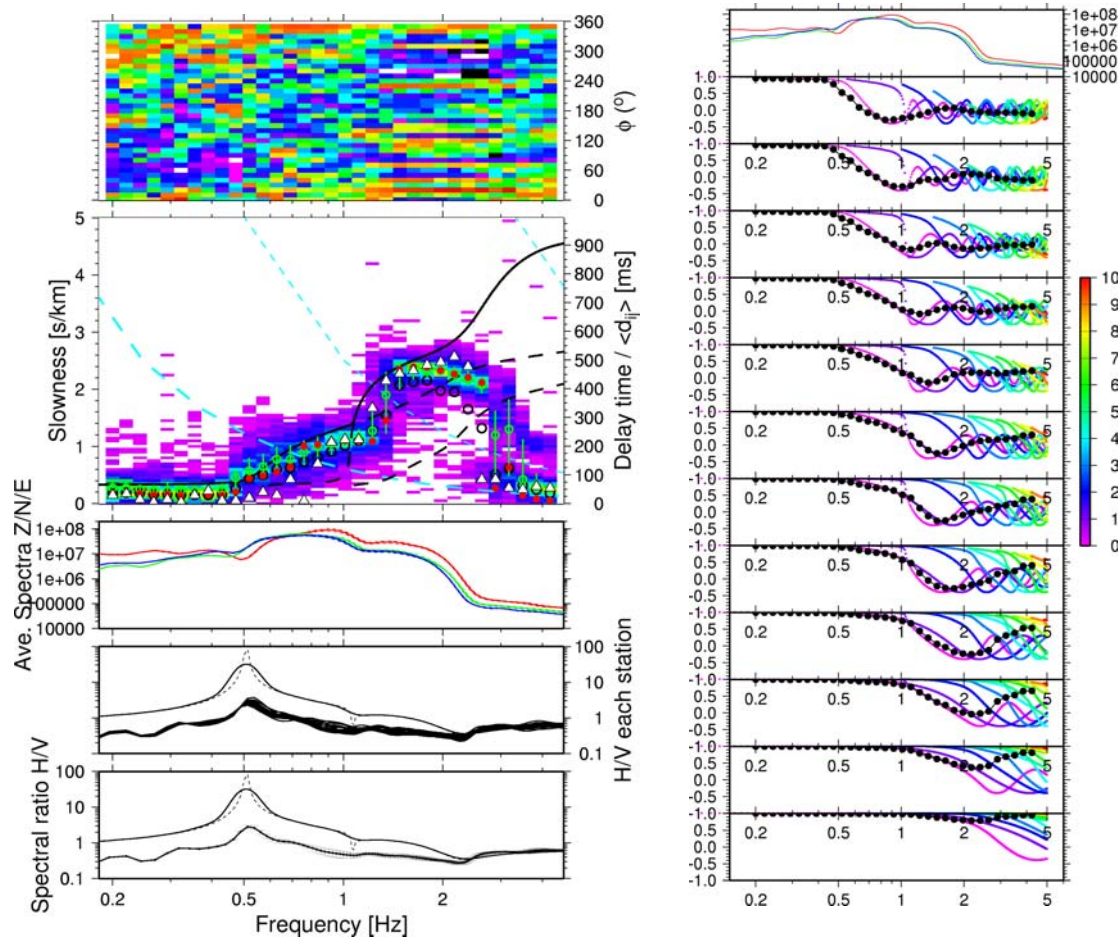


Figure 3-5 As Figure 3-4 for the “City” data set. The results obtained are very similar to the results for the “Single” source situation. It can be concluded that the case of a dominant source region with waves arriving from a confined azimuthal region is a favourable situation for dispersion curve analysis of ambient vibrations.

The results for the “City” data set are displayed in Figure 3-5. We find very similar results to the ones obtained for the “Single” data set. From the azimuthal distribution obtained via the FK technique we can recognize the dominant contribution of signals propagating from the region of concentrated source locations towards the array. As the total number of source excitations is significantly higher than for the “Single” data set, we observe a better developed distribution of slowness estimates from the FK analysis. The deviations found from the theoretical dispersion curve can be explained in the same way as before for the “Single” data set. As a preliminary conclusion we can state, that in case of encountering a clear directional preference in the observed ambient noise wavefield is a favourable situation for determining dispersion curves from ambient noise data.

In contrary, the results for the “Random” data set show a different picture. Both the slowness estimates obtained via the FK and the HRFK techniques show strong deviations from the true slowness within the analysed frequency band. This can be explained from the conclusions made in section 2.2. The superposition of array responses for multiple plane wave arrivals leads to a biased estimate of the true slowness of arriving wave packets.

The mismatch between observed and predicted spatial autocorrelation curves, however, seems to be reduced in comparison to the previous examples. In the pioneer work of Aki (1957) who introduced the spatial autocorrelation method (SPAC), the relation between phase velocity curves and the spatial autocorrelation coefficients have been derived by assuming a stationary random wavefield. As the “Random” data set approximates the underlying assumption in Aki’s work much better than the “Single” or “City” data sets, the improved fit of the observed spatial autocorrelation curves can be understood.



□

Figure 3-6 As Figure 3-4 and Figure 3-5 for the “Random” data set. For interpretation of the results please refer to text.

3.3 Array analysis of synthetics II (AM data set)

In order to investigate the influence of higher mode contributions on the array analysis results, we have analysed the AM data set for the three distinct source configurations “Single”, “City”, and “Random” in the same way as before the FM data sets. The results of this analysis are resumed in Figure 3-7 to Figure 3-9.

For frequencies lower than the cut-off frequency of the first higher mode (~ 0.5 Hz) we observe similar results to the ones we have obtained before for the synthetic wavefield that is composed solely of fundamental mode Rayleigh waves. For higher frequencies, however, the slowness estimates obtained from the FK and HRFK techniques can no longer be interpreted as phase velocities connected to one of the theoretical mode branches. The results have to be interpreted as “apparent dispersion curves”. This expression has been used by Tokimatsu (1997) who demonstrated the significant influence of higher modes by theoretical computation of dispersion curves for a point source acting on the surface of a layered medium. In his work, Tokimatsu (1997) showed for different soil layer models that the apparent dispersion curves can get very complex depending on the response factors (energy contribution) of higher mode Rayleigh waves.

Taking into account the results of Tokimatsu’s study, we would need to compare the obtained dispersion curves to a theoretically derived “apparent dispersion curve” in order to interpret the resolution capabilities of the array geometry in terms of dispersion curve determination from the AM data sets. Until now, we have not yet performed such a comparison. As this complicated appearance of dispersion curves depends mainly on the response factors of the higher modes for a specific velocity model, it is of great importance especially for the interpretation of measured dispersion curves. Therefore we will try to include the concept of “apparent dispersion curves” into the inversion problem of dispersion curves. This will allow us to interpret the quality of dispersion curve estimates for this synthetic example in terms of the quality of the obtained velocity model fit.

A similar observation is made for the MSPAC results for all data sets. A strong influence of the higher mode contribution is clearly observed. Tokimatsu et al. (1992) gave an expression for the effective apparent phase velocity consisting of multiple modes for stationary stochastic waves. We would therefore expect to get this effective phase velocity curves from the inversion of the spatial autocorrelation curves. An incorporation of higher mode contributions into the inversion process for obtaining the phase velocity curves for each mode branch is not possible, as the response factors depend on the velocity structure of the site.

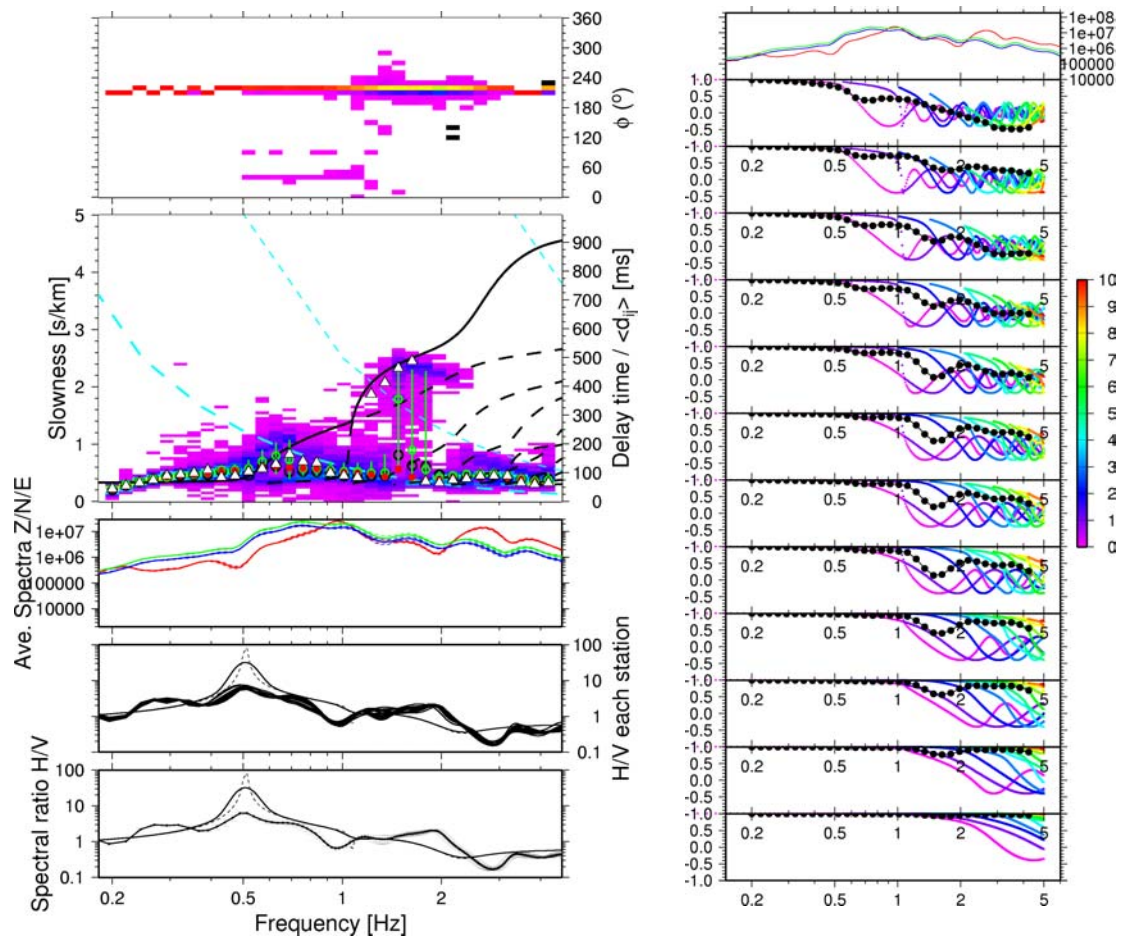


Figure 3-7 Results of array analysis for the “Single” data set including all existing Rayleigh wave modes (AM synthetic waveforms). The display of results is equivalent to previous figures (3-4 to 3-6). For details of interpretation, please see the text.

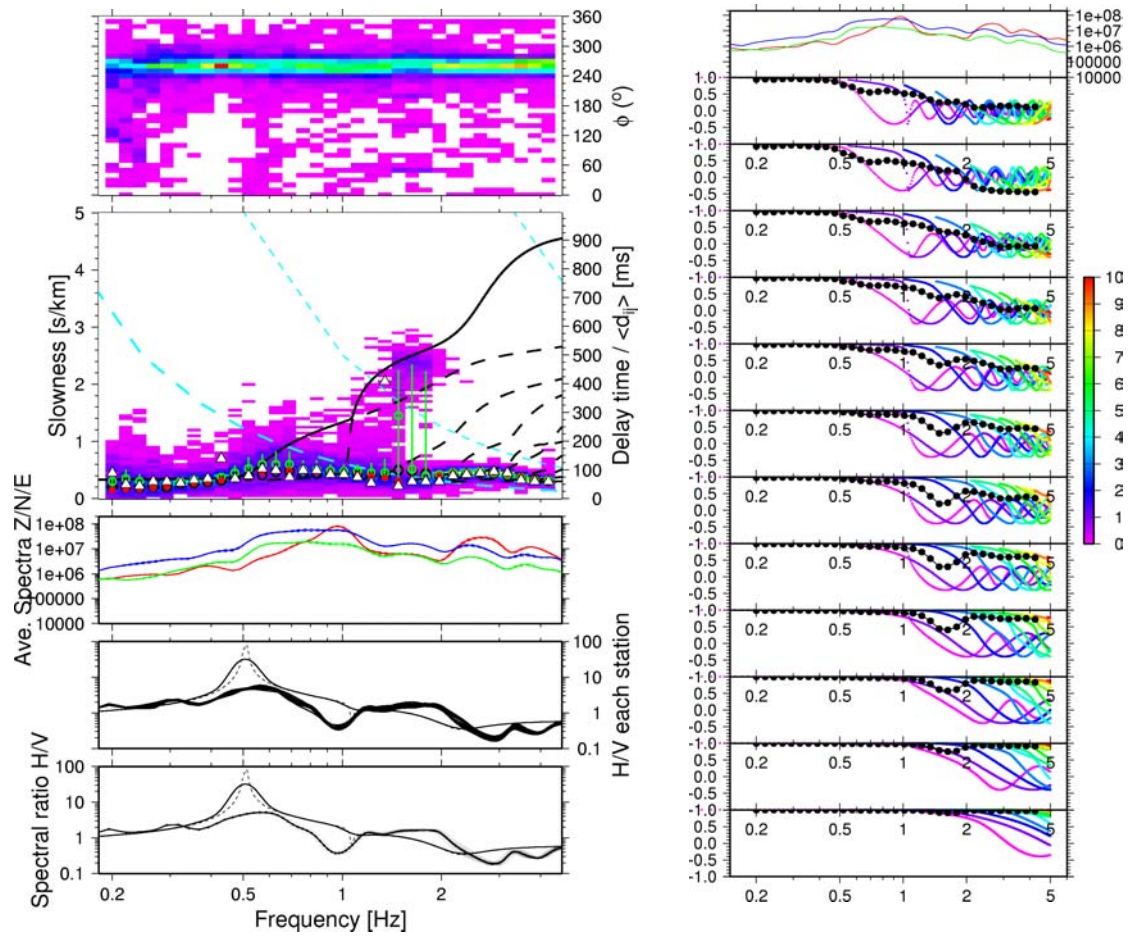


Figure 3-8 Results of array analysis for the “City” data set including all existing Rayleigh wave modes. The display of results is equivalent to previous figures (3-4 to 3-7). For details of interpretation, please see the text.

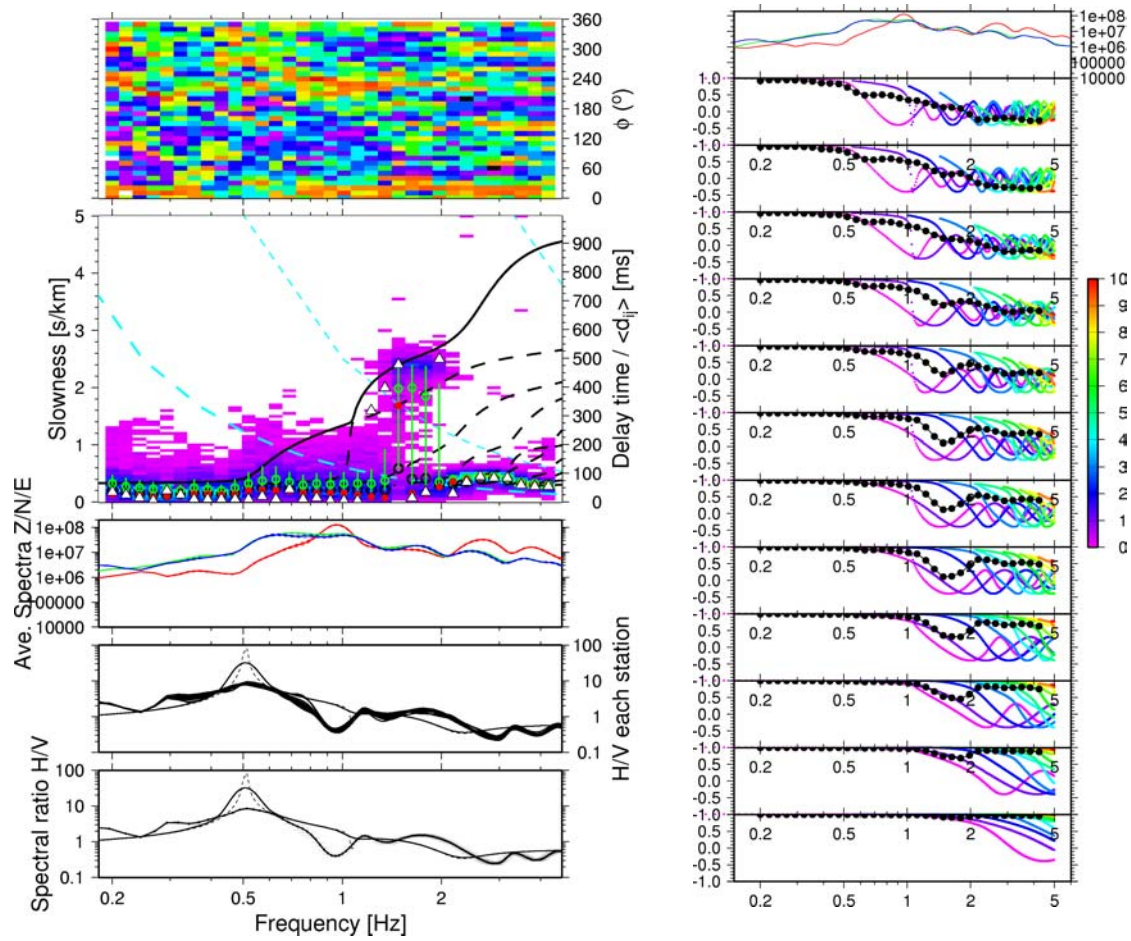


Figure 3-9 Results of array analysis for the "Random" data set including all existing Rayleigh wave modes. The display of results is equivalent to previous figures (3-4 to 3-8). For details of interpretation, please see the text.

4 Timing accuracy issues

The requirements regarding the necessary timing accuracy in field experiments for small aperture array analysis are high. Considering e.g. inter-station distances as small as 10 to 50 m and apparent phase velocities of 500 m/s, we must be able to measure arrival time differences between individual sensors of 20-100 ms with high accuracy. Allowing a 5% error in the phase velocity estimate for this example requires a timing error less than 1-5 ms.

There are mainly two instrumental effects that are critical for measuring the relative arrival times of seismic wave fronts: the clock stability and clock synchronization of the data acquisition system and the phase delays of seismometers used in field experiments.

4.1 Timing accuracy of data acquisition systems

Today's data acquisition systems usually make use of either GPS or DCF77 time synchronization. The timing accuracy of a GPS-synchronized data acquisition system depends on the quality of the GPS-receiver, the stability of the internal clock, and the software implementation of the time stamping of data blocks. The achievable timing accuracy of GPS-synchronized data loggers is considered to be usually below 1 ms. Systems based on DCF77-time decoding, however, show significant time delays (several tens of ms) when compared to GPS-systems. Those time delays are caused from the travel time of the radio signal captured by the DCF-antenna and passed through the analogue electronic of the DCF77-receiver. The time delays of individual DCF77 antennas differ significantly one from another (0-10ms); those time delays must be accounted for in array analysis applications. Fortunat Kind from ETHZ, Zurich has reported on laboratory measurements between time shifts of DCF77-clocks in comparison to a reference GPS-clock for the DCF77-synchronized Mars88 data loggers used at ETHZ. The report of Fortunat Kind is included as Appendix 1 to this deliverable.

The main results of Fortunat Kind's report have been important during the analysis of the data recorded at the test site in Otterbach (Weil am Rhein, Swiss-German border region, compare Deliverable D06.05). Equipment both from IGUP, Germany and ETHZ, Switzerland has been used in this array deployment in order to achieve a larger array configuration and to test sub-array processing techniques as well as to investigate structural 2D effects. The IGUP-equipment consisted in 13 Marslite data loggers with GPS-timing system, whereas the ETHZ-stations used DCF77-time synchronization with Mars88 data loggers. By duplicate recording of one of the sensor positions with both acquisition systems, we were able to verify the delay times of ca. 40 ms between GPS- and DCF77 timing for the particular station as has been determined in the laboratory at ETHZ.

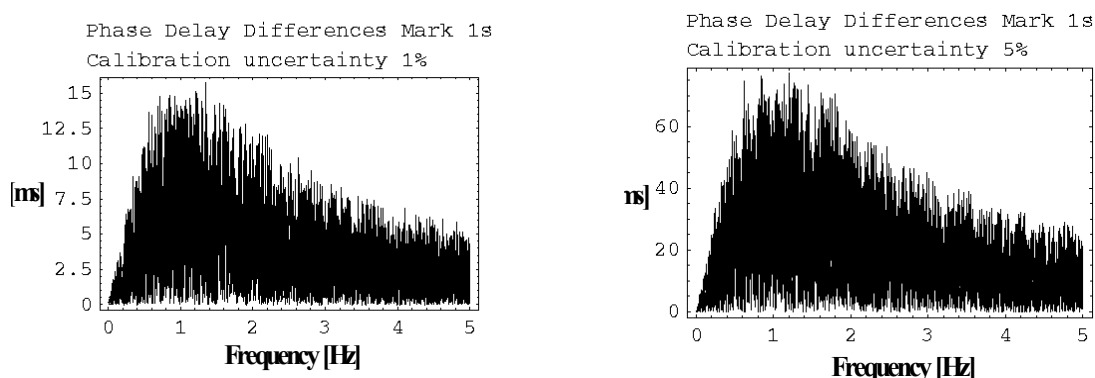
During the analysis we have recognized, that we were able to reduce the scatter of phase velocity estimates significantly when correcting the DCF77-time delays for the ETHZ-stations. However, even applying the time corrections, we have found that the results obtained by GPS-synchronized stations show narrower distributions for the phase velocities estimated from the same recording times. From this experience we suggest to use GPS-synchronized data acquisition systems for small aperture array analysis applications. If DCF77-time synchronized systems are used, it is essential to verify the DCF77-antenna time delays in the

laboratory and to apply the appropriate time corrections for estimating phase velocities with small apertures arrays.

4.2 Influence of seismometer calibration uncertainties

In order demonstrate the importance of taking into account seismometer phase delays for small aperture measurements the following numerical simulation has been performed. We have repeatedly calculated the phase delay for a standard 1 s seismometer (Mark 1s). For each computation we have selected randomly values for the corner frequency and damping following a normal distribution with 1% (5%) standard deviation and mean values according to the manufacturer's specifications ($f_0 = 1\text{ Hz}$, $h = 0.707$). Relative time delays between individual phase delay curves have been calculated for any pair of computed phase delay curves. The results of this simulation are displayed in Figure 4-1. The relative timing errors caused by uncertainties in the seismometer calibration are maximal around the corner frequency of the instrument. Relative time differences of up to 15 ms (60 ms) are found for this example, a value that exceeds the desired measurement accuracy of 1 to 5 ms for the aim to measure phase velocities with 5% accuracy for a typical small aperture array setup (see example given above). For this reason we consider it as very important to work with well calibrated seismic sensors in any measurement campaign.

At IGUP we use 13 Lennartz 3D-5s sensors for recording ambient vibration array data. All sensors have been tested and calibrated by the manufacturer at time of purchase (1997) and the uncertainty estimate given by Lennartz is better than 1%. In order to re-calibrate the sensors before each experiment (changes of system parameters due to aging and/or transport damages) we have purchased the portable broadband reference sensor KS2000 by Geotech Instruments, LCC. Unfortunately it turned out, that the Geotech Instruments did not offer a motorized mass centring device for the KS2000 at the time of purchase. As the stabilization time of the sensor after installation and manually centring the masses is very long, it is not feasible to use this reference sensor for fast installation in a field environment. We have therefore carried out a calibration test of the Lennartz 3D-5s sensors relative to the reference instrument KS2000 in the laboratory installations of IGUP.



□

Figure 4-1 Relative phase delay differences (in ms) calculated for a standard 1s seismometer with random uncertainties of 1% and 5% in both eigenfrequency as well as damping.

4.2.1 Calibration experiment setup

We have performed a huddle test at the laboratory installations of IGUP between the Lennartz 3D-5s seismometers and the reference broadband sensor KS2000 for a full day recording time. The data acquisition instruments available at IGUP are 13 Lennartz Marslite data loggers. For this reason, we could not record all 13 Lennartz 3D-5s together with the reference sensor KS2000 in a single experiment. We have performed the calibration test for two subsets subsequently, with parallel recordings of seven Le3D-5s (GP01-GP07) and the KS2000 on one day (24 h) and the remaining 6 Lennartz seismometers (GP08-GP13) together with the KS2000 at the following day (24 h). The KS2000 has been thermally isolated and protected from direct air circulation by wrapping the sensor with cotton fleece and covering it by a plastic basket. After installation we repeatedly adjusted the mass positions manually until we found that the KS2000 has been sufficiently stable. This procedure took two full days before being able to carry out the huddle test. In Figure 4-2 and Figure 4-3 we show the experimental setup at the installations of IGUP. The ground plate for all seismometers is decoupled from the building. When controlling the waveforms and amplitude spectra after the experiment, we were able to observe both resonance effects of the ground plate as well as electromagnetic noise coupling into the seismometer cables and/or electronics of the data acquisition system deteriorating the signal at higher frequencies above 8-10 Hz.



Figure 4-2 Huddle test of first seven Lennartz 3D-5s seismometers (GP01-GP07) together with the KS2000 broadband seismometer (thermally isolated below blue basket) in the laboratory installations of the Institute of Geosciences, University of Potsdam).



Figure 4-3 Huddle test for remaining six seismometers Le3D-5s (GP08-GP13) together with the reference instrument KS2000 (thermally isolated below blue basket).

4.2.2 Calibration evaluation

In order to obtain an estimate of the true transfer functions for each individual Lennartz 3D-5s seismometers, we have used a cross-spectral technique as described in Haskov and Alguacil (2002). Let $T_{1,theo}(j\omega)$ be the known complex transfer function of the reference sensor, $P_{21}(j\omega)$ an robust estimate of the cross spectra between the unknown sensor 2 and the reference sensor 1 and $P_{11}(j\omega)$ denoting the power spectra estimate of the reference sensor, then we can give an estimate of the unknown transfer function $T_{2,est}(j\omega)$ for sensor 2 as:

$$T_{2,est}(j\omega) = T_{1,theo}(j\omega) \frac{P_{21}(j\omega)}{P_{11}(j\omega)}$$

We have selected two hours of data recorded during night time to compute estimates of the complex cross spectra $P_{21}(j\omega)$ between each Lennartz 3D-5s seismometer (above: sensor 2) and the reference sensor KS2000 (above: sensor 1) by averaging over a number M of segments:

$$P_{21}(j\omega) = \frac{1}{M} \sum_{m=1}^M X_{2,m}(j\omega) X_{1,m}^*(j\omega)$$

where $X_{2,m}(j\omega)$ denotes the complex Fourier spectra of the respective Le3D-5s sensor and $X_{1,m}^*(j\omega)$ is the conjugate complex Fourier spectrum of the KS2000 reference sensor. For each hour of data we have averaged 11 segments of 600s length each with 50% overlapping

windows. The power spectrum estimate for the reference sensor ($P_{11}(j\omega)$) has been obtained equivalently by segment averaging for the same time windows.

The known transfer function of the reference sensor K2000 is given by the manufacturer as:

$$T_{1,theo}(j\omega) = \frac{K(j\omega)^2}{[(j\omega)^2 + j\lambda_L\varpi_L\omega + \varpi_L^2][(j\omega)^2 + j\lambda_H\varpi_H\omega + \varpi_H^2]}$$

with the following constants:

$K =$	$1.98 \times 1.e+8$	(amplification factor)
$\lambda_L =$	0.707	(damping for lower corner frequency)
$\omega_L =$	0.0629 radian/sec	(lower corner frequency)
$\lambda_H =$	0.707	(damping for upper corner frequency)
$\omega_H =$	314.6 radian/sec	(upper corner frequency)

Once we have obtained an estimate of the true transfer function $T_{2,est}(j\omega)$ for any of the Le3D-5s sensors, we have tried to fit those by adjusting the conjugate complex pole positions $a \pm jb$ and the generator constant G of the theoretical transfer function for the Lennartz 3D-5s seismometer $T_{2,theo}(j\omega)$ with a non-linear parameter space search technique, known as downhill-simplex algorithm (Press et al., 1994). The transfer function of the Le3D-5s is given as:

$$T_{2,theo}(j\omega) = \frac{G(j\omega)^3}{(j\omega - [a + jb])(j\omega - [a - jb])(j\omega - c)}$$

where the constant $c = -2\pi 0.068$ [Hz] is the pole position (on real axis) for an additional highpass filter of order 1 at the corner frequency 0.068 Hz stemming from the electronic feedback system after Lippmann (1982).

We have defined four different cost functions in order to investigate the stability of the non-linear fitting procedure. L1A and L2A as given below compute the residual according to the L1-norm and L2-norm, respectively, for the logarithm of the amplitude transfer functions between the observed ($T_{2,est}(j\omega)$) and fitted ($T_{2,theo}(j\omega)$) complex transfer functions over a limited frequency band at K discrete frequencies:

$$L1A = \frac{1}{K} \sum_{k=1}^K |\log(|T_{2,est}(j\omega)|) - \log(|T_{2,theo}(j\omega)|)|$$

$$L2A = \sqrt{\frac{1}{K} \sum_{k=1}^K |\log(|T_{2,est}(j\omega)|) - \log(|T_{2,theo}(j\omega)|)|^2}$$

For L1C and L2C (see below), the fit functions are defined as the complex transfer function differences between observed and fitted transfer function (L1-norm and L2-norm, respectively) for K discrete frequencies:

$$L1C = \frac{1}{K} \sum_{k=1}^K |T_{2,est}(j\omega) - T_{2,theo}(j\omega)|$$

$$L2C = \sqrt{\frac{1}{K} \sum_{k=1}^K |T_{2,est}(j\omega) - T_{2,theo}(j\omega)|^2}$$

In Figure 4-4 we display the result of the fit obtained for the four distinct cost functions for one of the Le3D-5s seismometers (the complete fitting results are given in Appendix 2). The frequency band has been limited to the range 0.15 to 1.0 Hz. For the visualization we have chosen to display both the amplitude transfer functions (lower box) as well as the phase delay curves $\phi(f)/2\pi f$ (upper box in each of the four sub-figures). The scattered data points in the figures are the values of the estimated transfer function which is to be fitted. The transfer functions computed for the starting parameters of the downhill-simplex search are shown as red curves. The final values of the conjugate complex pole position are given at the bottom whereas the eigen-frequency f_0 , the damping value h and generator constant G computed from the complex pole position are displayed at the upper part of the amplitude transfer function plots. The theoretical transfer function for the best fit result is displayed as green curve in all plots.

By comparing the fit results for the different cost functions, we find significant differences in the estimated pole position and the visual match of the fitting curves to the input data. Whereas for the L1A and L2A cost functions we are highly satisfied with the visual match of the fitted curve for the amplitude transfer plots, we see severe differences for the phase delays computed for the fitted seismometer parameters when compared to the recorded data. On the other hand we observe a very nice fit of the phase delays when using the L1C and L2C cost functions in the fitting procedure. However, smaller deviations between fitted and observed amplitude transfer functions are visible for lower frequencies in this case. Comparing the results obtained for all individual components of all seismometers and the two hours of data, we have observed very consistent results for each of the cost functions – however the absolute values of the pole positions differ significantly between the logarithmic amplitude fits and the complex transfer function fits. An obvious difference of the final derived seismometer parameters between the L1-norm and L2-norm fit functions could not be recognized (compare Appendix 2, complex pole position plots).

The main reason for performing this calibration test has been the aim to evaluate the magnitude of phase delay differences for our instruments with respect to the task of determination of accurate phase velocity curves from ambient vibration data. We will therefore give preference to the results which have been obtained by using the fitting functions L1C and L2C, providing a better fit to the observed phase delay curves for our seismometers.

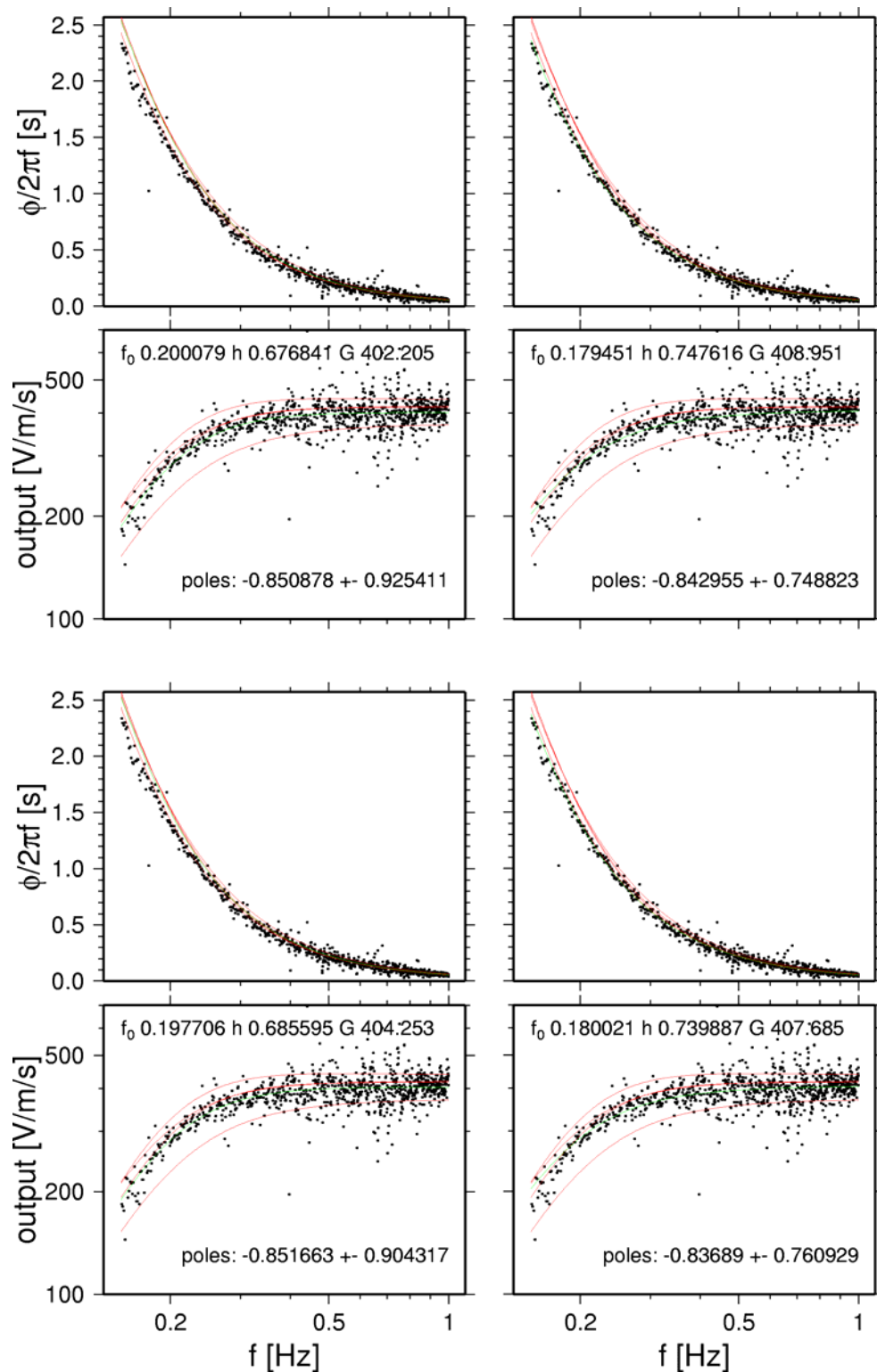


Figure 4-4 Display of downhill simplex inversion results for varying cost functions. Two panels on left bottom show the resulting model fits when considering the L1-norm logarithmic amplitude transfer cost function L1A. The red curves show the random starting models for the start of the downhill simplex. To judge the quality of the fit, the observed and modelled phase shifts are shown above the amplitude transfer fits. The results obtained by applying the cost function L2A are shown in the two panels on top left. The fitting results for L1C and L2C cost functions are displayed on the right bottom and top figures, respectively. For more details, please consult the text.

From the display of the fitted complex pole positions, shown for seismometers GP01 to GP05 in Figure 4-5, we can recognize that the scatter of the derived pole estimates is relatively low. In general we find too low eigenfrequencies and too high damping values if we compare the fitted parameters ($f_0 \sim 0.18$ Hz, $h \sim 0.74$) to the manufacturer specifications ($f_0=0.2$ Hz, $h=0.707$). We have found just a single exception: the fitted conjugate complex pole pair for the east component of seismometer GP05 (top right in Figure 4-5) is located at a very different position in the complex plane. The complete display of fitted pole positions in the complex plane for all seismometers is given in Appendix 2.

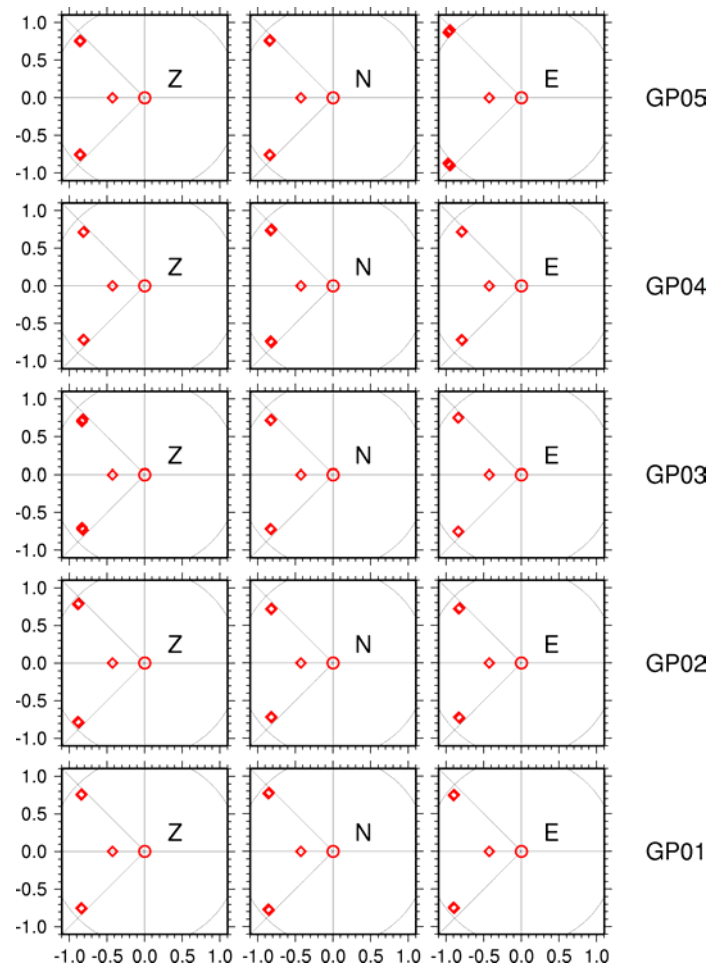


Figure 4-5 Display of pole position fits (red diamonds on left) in the complex plane for seismometers GP01 to GP05, all components. The pole describing the electronic high-pass filter of the feedback system lies on the real axis and has not been fitted. Crossing points of black circle and lines indicate the expected pole positions according to the manufacturer's specifications. Besides GP05 east component all seismometers seem to have lower corner frequencies and higher damping than expected.

The consistently too low values for the eigenfrequencies as well as the too high values for the damping seem to be far off the expected results for the well-known quality of instruments manufactured at Lennartz Electronic GmbH. This lets us wonder whether the absolute values should be actually interpreted. Taking into account, that the estimates of the observed transfer functions are obtained as a relative measure in relation to the accuracy of the known system, we rather suspect, that the KS2000 reference system shows small deviations from the manufacturer's specifications. We are currently thinking of a repetition of this experiment

with a STS2 broadband sensor and within an improved experimental setup (better controlled environment, i.e. seismological observatory site).

However, we still think that the relative scatter of the observed results for the instruments can be interpreted in terms of expected phase delay differences between the individual sensors. For this reason we have compared the phase delay curves computed for the fitted conjugate complex pole positions in order to get an estimate of the expected time shifts introduced by inaccuracies of instrument calibration. The phase delay curves are shown for all instruments and components in Figure 4-6. To our surprise, the differences of the phase delay curves show even larger times as expected for a 5% calibration uncertainty (~ 100 ms at the corner frequency of the instruments). This result stresses the need for performing an additional calibration experiment which provides more accurate calibration results.

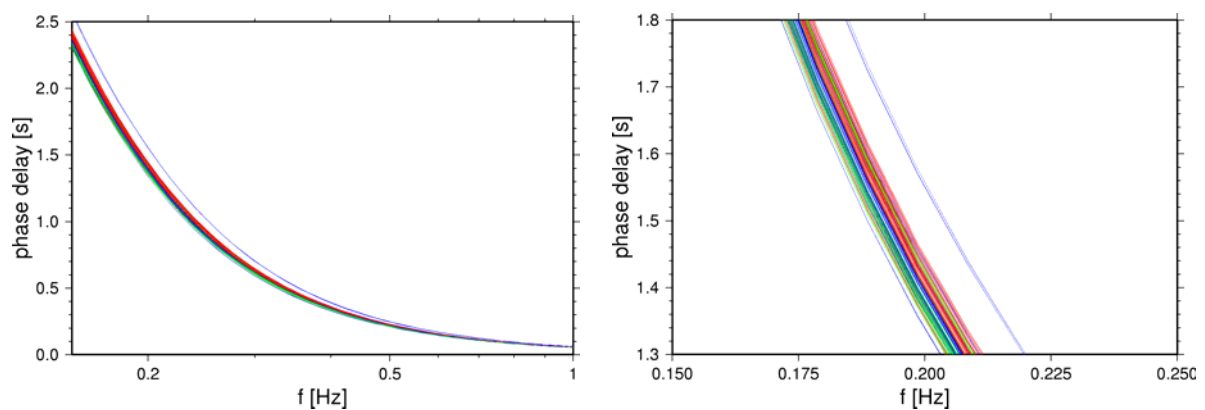


Figure 4-6 Phase delay curves for all components of all 13 seismometers for two frequency ranges. The vertical component phase delay curves are displayed in red, north components in green and east component curves in blue. In the left panel we have displayed the phase delay curves for the frequency band corresponding to the frequency band used for transfer function fitting. On the right, a detailed view close to the eigen-frequency of the instruments is shown, as the largest phase differences are expected for this frequency range. Even neglecting the strong deviations of the east component of GP05 (upper blue curves), we observe phase delay differences up to 100 ms between individual sensors.

5 Comparison to examples for real data sets

The following preliminary analysis has been performed for real data sets recorded within the scope of SESAME project. A detailed description of the data sets, measurement campaigns and recording sites is given in the deliverable D06.05. We present examples for two test sites: “Baviere” (Liege, Belgium), and “TST” (Euro-Seistest, Volvi-Graben, Greece) in the following. A complete overview of preliminary array analysis results for all test sites specified within the SESAME project is given in Appendix 3.

5.1 Baviere (Liege, Belgium)

The test site Baviere resides in the urban area of Liege, Belgium located on a small island in the river Meuse. The site is an open space of approximately 120 m diameter surrounded by streets and buildings. The ambient noise conditions are mainly dominated by the river Meuse and sporadic traffic on the surrounding streets. The supposed thickness of sedimentary soft layer over the first strong impedance contrast is estimated to be around 10 m for site Baviere.

In cooperation with the University of Liege and University of Grenoble we have measured several hours at high sampling rates (250 Hz) and varying array configurations. This allows us to compare the performance of different array geometries, i.e. for different array apertures and inter station distances. The results of this comparison are shown in Figure 5-1. In the upper left figure we show the results for the first array deployment, a 9 station configuration (the 10th station did not record properly) consisting of three tilted triangles with increasing diameter surrounding a central station. The station distribution for this setup is indicated by red and black circles in the station plot (lower right figure). After ca. 1 hour recording time the outer stations (red circles) have been re-deployed to the positions marked with the green circles. The radius of this inner triangle configuration is 10 m. The results for the second configuration are displayed in the upper left panel. Due to the narrower station spacing in the second array configuration, aliasing occurs at much higher frequencies, allowing for an extended recovery of the dispersion curve. The loss of resolution capabilities for the smaller aperture is recognized by the up-going dispersion curve at the lower frequency end.

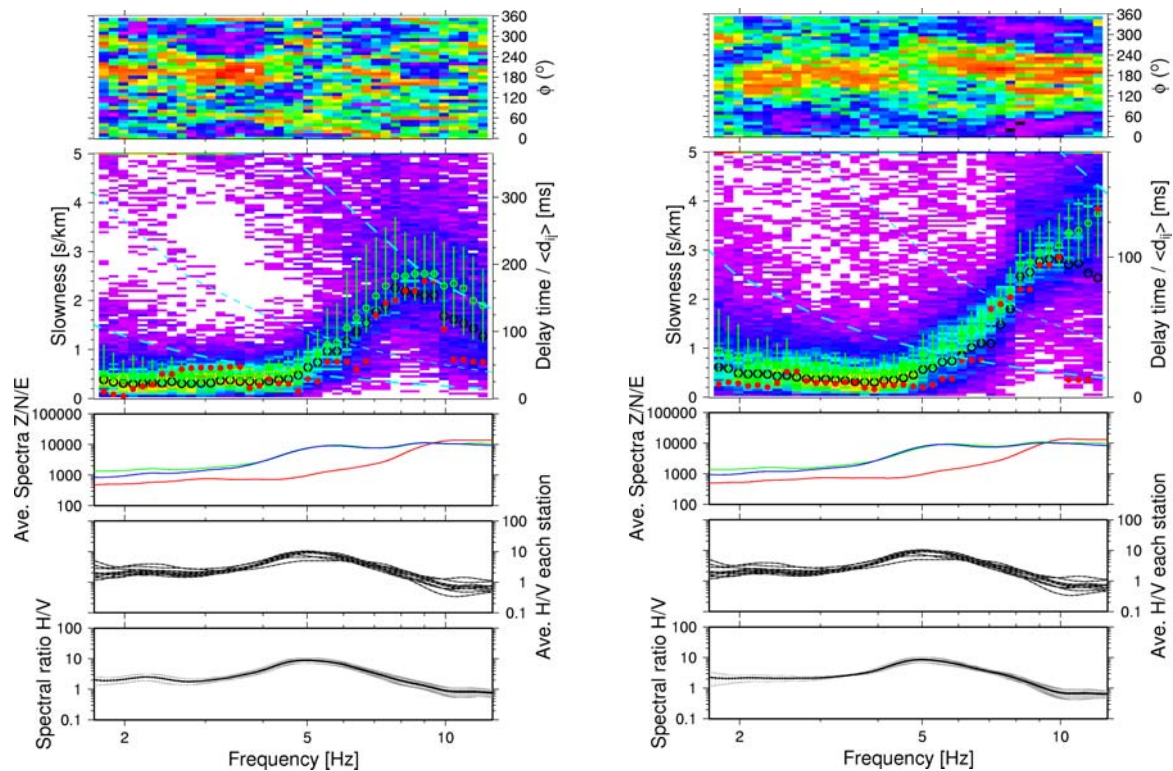
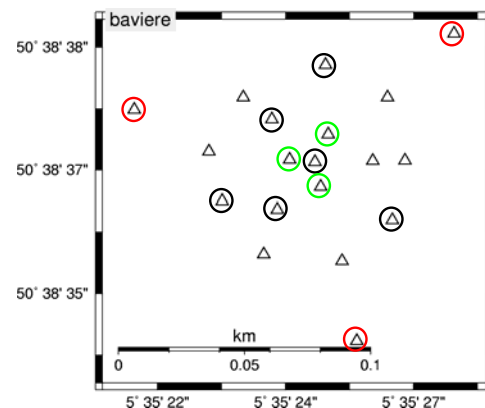


Figure 5-1 Comparison of the array results for different array apertures and inter station distances at site Baviere. Results displayed on top left have been obtained for the 9 station configuration indicated by red and black circles in the station plot (to the right). In a second step the outer stations (red circles) have been re-deployed to the positions marked with the green circles. The results for the second configuration are displayed in the upper left panel. Due to the narrower station spacing in the second array configuration, the aliasing occurs at much higher frequencies, allowing for an extended recovery of the dispersion curve. The loss of resolution capabilities for the smaller aperture is recognized by the up-going dispersion curve at the lower frequency end.



5.2 TST (Euro-Seistest, Volvi graben, Greece)

The Euro-Seistest site located in the Volvi graben close to Thessaloniki in NE Greece has been subject of a number of site amplification studies (e.g. Riepl et al., 1998, Dimitriu et al., 1998, Raptakis et al., 2000, Chavez-Garcia et al., 2000, Marrara and Suhadolc, 2001). The Volvi graben structure is well known from both geophysical as well as geotechnical investigations (Jongmans et al., 1998, Pitilakis, et al., 1999). During the field campaign in August 2002 several small aperture arrays have been deployed both within the valley as well as on the valley shoulder (Profitis generic rock site) in cooperation between the SESAME partners ITSAK (Thessaloniki) and IGUP (Potsdam). Within the context of instrumental

layout for ambient vibration array measurements we focus on the location of the Euro-Seistest, as we were able to deploy several array configurations with varying aperture, inter-station distances, and number of sensors.

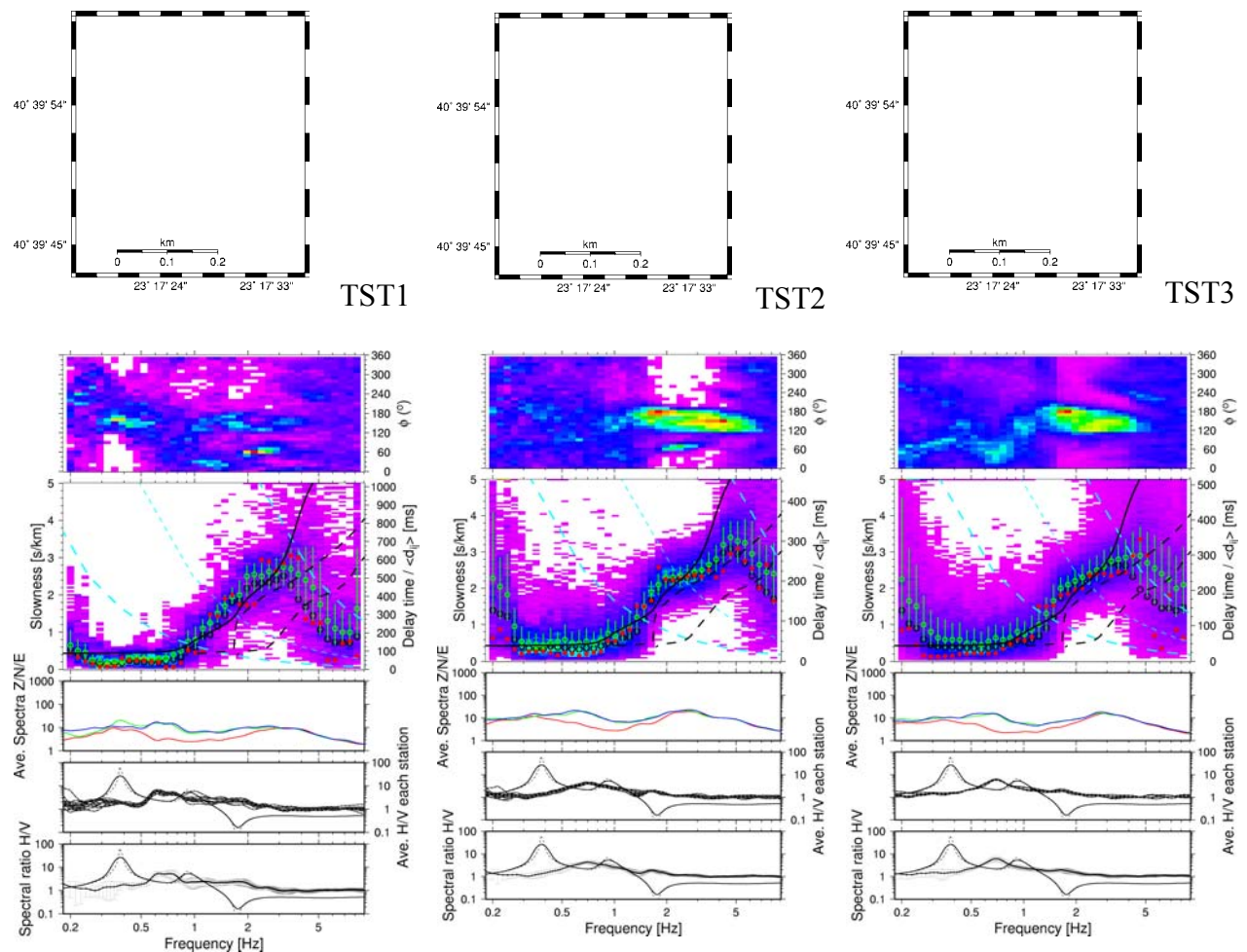


Figure 5-2 Results of FK-analysis for sites TST (Euro-Seistest, Volvi graben, Greece). Successively we have deployed three array configurations (TST1, TST2, and TST3) as shown on the top. For the first two settings we have used 13 sensor locations, whereas for the third configuration, only 7 stations were in use. In general the results of the dispersion curve analysis are similar for the frequency band from 1 to 3 Hz. The influence of inter-station distance on the occurrence of aliasing effects can be clearly observed.

In Figure 5-2 we show the results of FK-analysis for the three array deployments at site TST. The aperture of the first array setup is approximately 500m, and has been reduced to ca. 220m for the second geometry by relocation of the four outer stations closer to the centre. In the third setup, we have only used 7 stations out of 13, the 6 remaining instruments have been used in a parallel array recording at site GA in the Volvi basin.

The main features that can be observed from the FK-analysis results are the following: from setting TST1 to TST2, the upper frequency limit at which aliasing effects are clearly recognized is shifted from ~3 Hz to ~4 Hz, whereas the slowness distribution below 1 Hz seems to be much sharper for the larger aperture in the TST1 setup. Regarding the frequency band between 1 to 4 Hz, the second array configuration provides a more detailed picture of

the run of the dispersion curve. However, it has to be noted, that also the azimuthal distribution of coherent wave arrivals is quite distinct for the second experiment TST2. A change of wavefield characteristics to a more confined region of ambient noise sources might also explain the apparently better resolution of the structural information contained in the dispersion curve for this specific frequency band. Comparison of TST2 and TST3 configuration shows that the reduction of the number of stations used in the array analysis clearly deteriorates the quality of the measured dispersion curve. However, for a narrower frequency band (1 to 2 Hz) the results obtained from the 7 station array setup TST3 are very similar to the ones obtained with the 13 station configuration in TST2.

6 Conclusions

Within the context of WP05 of the SESAME project, the instrumental layout of array measurements for the determination of dispersion curves from ambient vibration recordings, we have investigated the resolution capabilities of various array geometries, processing methods and measurement errors due to limited calibration accuracies of seismic sensors.

If we just consider the frequency wavenumber filter response of a specific array configuration and further assume that the ambient vibration wavefield is best described as a superposition of plane wave arrivals from surface waves, we must recognize, that for a limited number of stations and feasible array apertures, we are not able to estimate high-quality dispersion curves for a broader frequency band of interest. The superposition of array response functions prohibits to reasonably estimate the true slowness value by picking that absolute slowness value, which is connected to the apparently most coherent plane wave arrival in the wavenumber grid. For this general situation (arbitrary azimuth of propagation and higher mode contributions) the obtained estimates of the true slowness are heavily biased in almost all cases.

When trying to determine the resolution capabilities from a synthetic waveform data set for a realistic geologic situation (generic deep basin model), we find that the distribution of sources as well as the presence of higher mode contributions has a severe influence on the estimated dispersion curve. Only in situations where mainly fundamental mode (Rayleigh-type) surface waves are propagating from a preferential source azimuth, an acceptable dispersion curve could be derived from the synthetic data. The choice of analysis method (FK versus HRFK) seems not to matter in such a situation. Comparing the array analysis results obtained for real data sets (see also Appendix 3 to this report), the situation of preferential source azimuths seems to be a feature which is relatively often observed in real data. At least for certain frequency bands we have observed dominant directions of signal arrivals, and the presence of higher modes could not be confirmed in any measurement from the preliminary analysis performed until now. The little influence of clearly recognizable higher mode contributions may be caused either by irregularities in the topography of subsurface layers, which rapidly destroys the higher mode wavefield in nature or because we interpret the obtained dispersion curves wrongly as the fundamental mode branch although we are in reality observing an “apparent dispersion curve” with mixed contributions of fundamental and higher mode surface waves (Tokimatsu et al., 1992).

As it is not feasible to perform ambient vibration measurements with several tens of stations within urban environments in order to increase the wavenumber resolution of array measurements for a broad frequency band, we suggest the use of an adaptive deployment

strategy for the determination of partially better dispersion curves for narrower frequency bands (wavelength ranges). However, this implies that we need to perform various array measurements subsequently with different aperture and station spacing in order to cover the complete frequency band of interest (wavelength range of interest). As a result, ambient vibration noise measurements will become more expensive as the time for field campaigns will increase significantly and additionally the logistics of field campaigns will be more complicated. We feel therefore that there is a strong need to develop tools for fast array deployment and online-processing capabilities to reduce the time and cost of field measurements.

Additionally it is of great importance to address the following question: what is the necessary accuracy of measurements we need to obtain for the determination of dispersion curves with respect to the resulting site response estimates? For certain situations it might not be interesting to try to estimate phase velocities arbitrarily precise if the resulting site amplification factors for those frequency bands which are important in engineering problems are not sensitive to the specific phase velocity estimates. The measurement accuracy requirement will most certainly vary from one geologic situation to another. This suggests that we need to incorporate the inversion of dispersion curve measurements into structural models and subsequent site response modelling into the array measurement strategy. This may allow to put the focus of adaptive measurement schemes to those wavelength ranges that are of key interest in terms of engineering purposes and seismic hazard issues.

Acknowledgements

This project (Project No. EVG1-CT-2000-00026 SESAME) is supported by the European Commission – Research General Directorate. We thank all personnel from our partner institutions, who have helped in the preparations of field experiments, the measurement campaigns and data archiving.

References:

- Aki, K.,** *Space and time spectra of stationary stochastic waves, with special reference to microtremors*, Bull. Earthquake Res. Inst. Univ. Tokyo, Vol. 35, pp. 415-457, 1957.
- Bettig, B., Bard, P.-Y., Scherbaum, F., Riepl, J., and Cotton, F.,** *Analysis of dense array noise measurements using the modified spatial auto-correlation method (SPAC). Application to the Grenoble area*, submitted, 2001.
- Brüstle, W., & Stange, S.,** *Geologische Untergrundklassen zum Entwurf von Normspektren für DIN 4149 (neu), LRGB Baden-Württemberg, AZ: 3480.01/98-4764, 1999.*
- Capon, J.,** *High-Resolution Frequency-Wavenumber Spectrum analysis*, Proceedings of the IEEE, Vol. 57, No. 8, pp. 1408-1419, 1969.
- Chávez-García, F.J., Raptakis, D., Makra, K., and Pitilakis, K.,** *Site effects at Euroseistest – II. Results from 2D numerical modelling and comparison with observations*, Soil Dynamics and Earthquake Engineering, Vol. 19, pp. 23-39, 2000.
- Dimitriu, P.P., Papaioannou, Ch.A., and Theodulidis, N.P.,** *EURO-SEISTEST Strong-Motion Array near Thessaloniki, Northern Greece: A Study of Site Effects*, Bulletin of Seismological Society of America, Vol. 88, No. 3, pp. 862-873, 1998.
- Haskov, J., and Alguacil, G.,** *Instrumentation in Earthquake Seismology*, Monograph, 313 pp., prel. version, June 2002, <http://www.ifjf.uib.no/Seismologi/SOFTWARE/OTHER/instrument.pdf>.
- Herrmann, R.B.,** *Computer programs in seismology*, Version 3.0, 1996, St. Louis University, <http://www.eas.slu.edu/People/RBHerrmann/ComputerPrograms.html>.
- Jongmans, D., Pitilakis, K., Demanet, D., Raptakis, D., Riepl, J., Horrent, C., Tsokas, G., Lontzetidis, K., and Bard, P.-Y.,** *EURO-SEISTEST: Determination of the geological structure of the Volvi graben and validation of the basin response*, Bulletin of the Seismological Society of America, Vol. 88, pp. 473-487, 1998.
- Kværna, T., and Ringdahl, F.,** *Stability of various f-k estimation techniques*, in: Semiannual Technical Summary, 1 October 1985 - 31 March 1986, NORSAR Scientific Report, 1-86/87, Kjeller, Norway, pp. 29-40, 1986.
- Lippmann, E.,** *Umwandlung eines elektrodynamischen Seismometers in einen Beschleunigungs-Aufnehmer und Entwicklung eines Filters zur Simulation eines niederfrequenten Seismometers*, Diplomarbeit, Institut für Allgemeine und Angewandte Geophysik der Ludwig-Maximilians-Universität München, 1982.
- Marrara, F., and Suhadolc, P.,** *2-D Modeling of Site Effects Along the EURO-SEISTES Array (Volvi Graben, Greece)*, Pure and Applied Geophysics, Vol. 158, pp. 2369-2388, 2001.
- Ohrnberger, M. and Scherbaum, F., 2002,** *Derivation of surface wave dispersion curves from array analysis of ambient vibrations*, presented as poster in Subcommission SCF-3 „Seismological Investigations of Site Conditions and Site Response“ at: XXVIII General Assembly of the European Seismological Commission (ESC), September 2nd-6th, 2002, Genoa, Italy.
- Pitilakis, K., Raptakis, D., Lontzetidis, K., Tika-Vassilikou, Th., and Jongmans, D.,** *Geotechnical and geophysical description of Euro-Seistest, using field, laboratory tests and moderate strong motion recordings*, Journal of Earthquake Engineering, Vol. 3, No. 3, pp. 381-409, 1999.
- Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.R.,** *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edition, Cambridge University Press, New York, 1994.
- Raptakis, D., Chávez-García, F.J., Makra, K., and Pitilakis, K.,** *Site effects at Euroseistest – I. Determination of the valley structure and confrontation of observations with 1D analysis*, Soil Dynamics and Earthquake Engineering, Vol. 19, pp. 1-22, 2000.
- Riepl, J., Bard, P.-Y., Hatzfeld, D., Papaioannou, C., and Nechtschein, S.,** *Detailed Evaluation of Site-Response Estimation Methods across and along the Sedimentary Valley of Volvi (EURO-SEISTEST)*, Bulletin of Seismological Society of America, Vol. 88, No. 2, pp. 488-502, 1998.
- Rietbrock, A. and Scherbaum, F.,** *The GIANT analysis system*, Seismological Research Letters, Vol. 69, No. 6, pp. 40-45, 1998.

Reamer, Sh.-K., & Hinzen, K.-G., *Evaluation of the Phase Data Catalog for the Station Bensberg for the Years 1970-2000*, DGG Annual Meeting, Frankfurt am Main, 19.03.-23.03.2001.

Scherbaum, F. and Johnson, J., *PITSA, Programmable Interactive Toolbox for Seismological Analysis*, IASPEI Software Library, Vol. 5, 1992.

Scherbaum, F., Hinzen, K.-G., and Ohrnberger M., *Determination of shallow shear velocity profiles in the Cologne/Germany area using ambient vibrations*, Geophysical Journal International, in press, accepted August 21st, 2002.

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

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

Tokimatsu, K., Tamura S., and Kojima, H., *Effects of multiple modes on Rayleigh wave dispersion*, Journal of Geotechnical Engineering, ASCE, Vol. 118, No. 10, pp. 1529-1543, 1992.

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

Woods, J.W., and Lintz, P.R., *Plane waves at small arrays*, Geophysics, Vol. 38, No. 6, pp. 1023-1041, 1973.

Zywicki, D.J., *Advanced Signal Processing Methods Applied to Engineering Analysis of Seismic Surface Waves*, Ph.D. thesis, Georgia Institute of Technology, 1999.

List of Appendices:

- Appendix 1: Internal report on Timing differences of Mars88 DCF77 time synchronisation by Fortunat Kind (ETHZ, Zurich, Switzerland).**
- Appendix 2: 81 Figures on transfer function from the calibration test performed for all 13 Lenmartz 3D-5s sensors at IGUP.**
- Appendix 3: 30 Figures on the preliminary results of dispersion curve analysis for test sites.**

These appendices are available on the CD appended to the 18th month report.