

COMPARISON OF THREE METHODS TO STUDY THE WAVEFIELD IN THE VALLEY OF MEXICO USING DATA FROM DENSE ARRAYS

Agathe Roullé⁽¹⁾ and Francisco J. Chávez-García⁽²⁾

SUMMARY

Observed seismic motion in the lakebed zone of the Valley of Mexico is complex because of the large epicentral distance, an irregular 3D crustal structure, large noise levels, strong site effects, and the possible presence of mixed modes and multipathing. These conditions tax the possibilities of the array-processing techniques. In this paper, we test three methods to determine the most appropriate to analyze the data from a new dense accelerograph array, recently installed in Colonia Roma of Mexico City: the conventional frequency-wavenumber ($f-k$) method, the maximum-likelihood $f-k$ method and the Multiple Signal Characterization technique (MUSIC). Tests with synthetic data show that the $f-k$ methods are more robust in presence of mixed signals but that MUSIC has better resolution and is less sensitive to noise. Results from data analysis show the superiority of the conventional $f-k$ method in case of data with small coherence and large spatial aliasing, while MUSIC is more useful in case of data with high coherence and small time delays. Our results show that the wavefield includes slow diffracted waves coming mainly from south to west at short periods and faster waves from the epicenter at long periods. These observations support the idea that the observed large duration of ground motion is explained by the interaction of different modes of surface waves, arriving at different times at Mexico City, with the resonance of the clay layer.

RESUMEN

El movimiento sísmico observado en el Valle de México es complejo debido a la distancia epicentral, una estructura irregular, efectos de sitio importantes, y la posible presencia de distintos modos de propagación. Estas condiciones dificultan el análisis de datos de arreglos sísmicos. Hemos evaluado tres métodos para analizar los datos de un nuevo arreglo acelerográfico, instalado en la Colonia Roma: frecuencia-número de onda ($f-k$) convencional, $f-k$ de alta resolución, y la técnica MUSIC (Multiple Signal Characterization). Las pruebas con datos sintéticos indican que los métodos $f-k$ son más robustos ante señales complejas pero MUSIC tiene mejor resolución y es menos sensible al ruido. MUSIC es una buena alternativa para datos coherentes con pequeños retrasos entre estaciones. Los resultados para datos reales muestran la superioridad del método $f-k$ cuando la coherencia es baja. El campo de ondas incluye ondas difractadas lentas del SO en periodos cortos, y ondas rápidas del epicentro en periodos largos. Nuestras observaciones apoyan la

Artículo recibido el 18 de junio de 2004 y aprobado para su publicación el 8 de junio de 2005. Se aceptarán comentarios y/o discusiones hasta cinco meses después de su publicación.

- (1) Instituto de Geofísica, UNAM; currently at BRGM, 3 avenue Claude Guillemin, 45060 Orléans Cedex 2, France a.roulle@brgm.fr
- (2) Instituto de Ingeniería, UNAM, Ciudad Universitaria, Apdo. Postal 70-472, Coyoacán, 04510 México, D.F. México paco@pumas.iingen.unam.mx

idea de que la gran duración del movimiento sísmico se debe a la interacción entre modos de ondas superficiales, llegando a distintos tiempos a las estaciones, con la resonancia de la capa de arcilla.

INTRODUCTION

Research in strong ground motion in Mexico City during the last 15 years has been motivated, to a great extent, by the devastating 1985 Michoacán earthquake. That great ($M_s=8.1$) subduction zone event, with epicenter more than 300 km away, caused unprecedented damage in the lakebed zone of that large city (the geotechnical zone defined by the presence at the surface of extremely soft, saturated clays, with shear wave velocities as low as 60 m/s). Many studies have dealt with analysis of the source (*e.g.*, Eissler *et al.*, 1986; Houston and Kanamori, 1986; Singh *et al.*, 1988), path (*e.g.*, Ordaz and Singh, 1992; Cárdenas *et al.*, 1997; Cárdenas and Chávez-García, 2003), and site effects (*e.g.*, Singh *et al.*, 1988; Chávez-García and Bard, 1994; Chávez-García *et al.*, 1995; Iida, 1999). The combination of efficient energy propagation in directions perpendicular to the Pacific coast, a regional amplification of ground motion that is already observed on hard rock in Mexico City, and large amplification caused by the soft superficial sediments resulted in more than 10,000 casualties and large economic losses. Strong motion records obtained in Mexico City for that event showed peak ground acceleration three times larger on the lakebed zone than on hard rock. In addition, lakebed zone records also showed an increased duration of strong ground motion, almost up to three times that observed on firm soil. Since 1985, many efforts have been directed towards understanding these phenomena.

In 1985, eight digital strong motion instruments were in operation in Mexico City. Those data were analyzed by Singh *et al.* (1988) using spectral ratios, while Chávez-García *et al.* (1995) tried to identify the wavefield associated to strong ground motion. The latter authors identified coherent wavetrains in the records in the period band 3 to 5 sec, and suggested that the large duration was the result of the interaction of deeply guided surface waves (with wavelengths in the order of km) with the very local resonance of the thin (40 to 60 m) soft clay layer. After 1985, a significant effort allowed the installation of more than 100 new digital strong motion stations in Mexico City. Data recorded by these stations have allowed large progress in the measurement of amplification, and have produced a large database for statistical analysis. However, in spite of a few works (Chávez-García and Salazar, 2002), these data have allowed only little progress towards understanding the wavefield that propagates in Mexico City during large earthquakes. This understanding is critical as current predictions of strong ground motion are based on statistical analysis, while a physical model of the seismic response of this valley is lacking. Data from the permanent accelerograph network installed in Mexico City are not very useful to this aim, as it is marred by two problems: lack of a common time base for the recording stations, and an average interstation spacing that is too large for the very short wavelengths that dominate ground motion in the lakebed zone. Data from a dense array of stations is required to analyze the wavefield in the lakebed zone and explore the validity of the hypothesis proposed by Chávez-García *et al.* (1995) and Chávez-García and Salazar (2002): large duration of ground motion results from the interaction of different modes of surface waves, arriving at different times at Mexico City, with the local resonance of the soft clay layer.

The wavefield propagating in the lakebed zone in Mexico City is complex. There are several reasons for this complexity. Large earthquakes occur far from Mexico City, along the subduction zone in the Pacific coast. The large propagation path, coupled with the heterogeneous crustal structure makes multipathing likely (Singh and Ordaz, 1993). In addition, the crustal structure of central Mexico is irregular in 3D, due to the presence of the TransMexican Volcanic belt, oblique to the subduction zone. Furumura and Kennett, (1998) showed that ground motion is significantly affected by this structure. Another factor is the 3D structure of Mexico City valley, where a very thin layer amplifies ground motion by a factor as large as 40. This large amplification makes it difficult to determine the driving motion that excites the resonance of the soft clay layer. Finally, a large population makes large urban noise and, consequently, records with a low signal-to-noise ratio. It also makes it extremely difficult to install a dense array with an optimal geometry, since its geometry will be dictated by the location of parks or other public spaces. All these issues suggest that great care is to be taken to select the more appropriate method for the analysis of the data from the array.

In this paper we present results of a research program directed to understand the wavefield that propagates in Mexico City during large earthquakes. Results consist of two parts. First, a methodological study is conducted with the aim of evaluating three different methods for the analysis of data from arrays of stations considering the local conditions of Mexico City. These methods are: the $f-k$ (frequency-wavenumber) power spectrum (Aki and Richards, 1980), the high resolution $f-k$ power spectrum (Capon, 1969) and the Multiple Signal Characterization, MUSIC, method (Schmidt, 1981, 1986). Secondly, the three methods are applied to data recorded by two arrays in Mexico City: a regular array that was installed temporarily in 1994 in the firm soil zone of Mexico City, and a new accelerograph array, installed in the densely urbanized area of Colonia Roma, a neighborhood that was heavily affected in 1985.

Methods were evaluated using synthetic signals generated with the geometry of the actual arrays. The purpose was to evaluate the performance of each method, given the geometry of our arrays, in case of records consisting of several simultaneous signals and in presence of noise. This evaluation was more detailed for MUSIC method, given its larger number of parameters. The resulting parameters were then used to analyze one event recorded at each of the two arrays. Results show that no single method is adequate for the local conditions of Mexico City. In the firm zone, characterized by high coherency among stations and high velocities, MUSIC is shown to be more efficient and to give exact results with low uncertainties. On the contrary, in the lakebed zone, characterized by poor coherency among stations and very low velocities in the uppermost clay layer, the conventional $f-k$ method give better results.

Results show that ground motion in the lakebed zone of Mexico City results from different contributions and the critical change suffered by the wavefield nature at the site period. The wavefield recorded at long periods in the firm zone is composed of a first wavetrain of fast waves (probably body waves or L_g waves) followed by slower surface waves, all of them coming from the epicentral direction. In the lakebed zone, at short periods, we observe slow diffracted waves traveling from multiple directions, mainly from southern to western directions, and probably generated by the clay layer resonance. At longer periods (longer than the site period),

we observe faster wavetrains coming from the epicentral direction. These wavetrains correspond to surface waves guided by the deep crustal structure of the basin and observed in the firm zone.

METHODS FOR THE ANALYSIS OF ARRAY DATA

Three methods were considered to process data from dense arrays in the Valley of Mexico: the conventional f - k power spectrum analysis, the high-resolution f - k power spectrum analysis and the MUSIC algorithm. Following, we briefly describe each of them.

Frequency-wavenumber power spectrum

This method is based on the determination of a frequency-wavenumber (f - k) power spectrum computed using Fourier transform. This spectrum gives a measure of the amount of energy distributed among different wave velocities and directions of propagation. The position of the main peaks of the spectrum characterizes the signals that propagate through the array.

Consider an array of K sensors and seismic records of L samples divided in M windows of N data points ($L=MN$). The frequency-wavenumber power spectrum, $P(\mathbf{k},f)$, is defined by (Aki and Richards, 1980)

$$P(\mathbf{k}, f) = \frac{1}{K^2} \sum_{j=1}^K \sum_{l=1}^K w_j w_l^* C_{jl}(f) e^{i\mathbf{k} \cdot (\mathbf{r}_j - \mathbf{r}_l)} \quad (1)$$

where the cross power spectral function C_{jl} is defined by:

$$C_{jl}(f) = \frac{1}{M} \sum_{n=1}^M S_{jn}(f) S_{ln}^*(f) \quad (2)$$

In these equations, f is the frequency, \mathbf{k} is the wavenumber vector, \mathbf{r}_i is the position vector of station i , $S_{jn}(j)$ represents the discrete Fourier transform of the n^{th} window of the signal recorded at station j , and w_j denotes the weight function used to control the shape of the wavenumber window. Asterisk represents the complex conjugate.

The differences between different estimates of the frequency-wavenumber spectra result from the choice of the weight functions. The conventional method uses fixed weights that do not depend on the wavenumber. In this case, resolution in the wavenumber domain is determined essentially by the beam pattern of the array (Aki and Richards, 1980). Thus, for an array with a limited number of sensors and limited aperture, for which the response of the array to the unit impulse has a broad main peak and secondary peaks, the precision of the wavenumber estimates will be limited.

Several high-resolution frequency-wavenumber methods have been developed to improve the conventional f - k method (Marcos, 1998). The more popular is the maximum-likelihood method by Capon (1969). This method computes a high-resolution estimate of the power spectrum as

$$P(\mathbf{k}, f) = \left[\sum_{j=1}^K \sum_{l=1}^K Q_{jl}(f) e^{i\mathbf{k} \cdot (\mathbf{r}_j - \mathbf{r}_l)} \right]^{-1} \quad (3)$$

where Q_{jl} represents the inverse of the cross spectral matrix C_{jl} .

Capon showed that $P(\mathbf{k}, f)$ given in eq. (3) could be considered as the power output of a maximum-likelihood filter whose shape varies with wavenumber. This filter does not distort monochromatic waves traveling at a velocity corresponding to the wavenumber \mathbf{k} and suppresses the power of other waves in an optimum least-square sense. Experimental results have shown significant improvements in the wavenumber resolution of this method relative to the conventional estimate (Capon, 1969, 1973).

Computation of f - k spectra using the previous expressions makes several assumptions. The wavefield is assumed to consist of a superposition of simple plane waves. Data are assumed to be stationary in both time and space. Finally, the different signals present in the record are considered to be uncorrelated. These assumptions could pose severe limitations in the case of Mexico City.

Multiple Signal Characterization (MUSIC)

The Multiple Signal Characterization method is a high-resolution technique based on the calculation of the eigenstructure of the cross-spectral matrix to determine signal parameters. It can be applied in time domain as well as in frequency domain. Following the signal subspace approach of Schmidt (1981, 1986), Goldstein and Archuleta (1987, 1991a, 1991b) applied MUSIC in seismology. They showed that it could resolve multiple, closely spaced sources. It works both with stationary and nonstationary signals and is more sensitive to the strongest sources. MUSIC assumes that there are a finite number of signals going through the array (smaller than the number of sensors, K).

The first step of the method is the calculation of the covariance matrix (in time domain) or cross spectral matrix (in frequency domain) for each frequency and time windows. Then, the eigenstructure of the covariance matrix is determined and divided into signal and noise subspaces. To do that, MUSIC determines the number of signals, q . The signal subspace is then defined by the eigenvectors corresponding to the q strongest eigenvalues, while the noise subspace is defined by the $K-q$ weakest eigenvalues.

The next step is the determination of the directions of the signal vectors using the property of orthogonality between signal and noise subspaces. A set of array manifold vectors $\mathbf{a}(\mathbf{k})$ is defined as:

$$\mathbf{a}(\mathbf{k}) = [e^{ik \cdot \mathbf{r}_1}, e^{ik \cdot \mathbf{r}_2}, \dots, e^{ik \cdot \mathbf{r}_K}] \quad (4)$$

Then, the vectors that give the minimum projection onto the noise subspace are determined through the search of the maxima of the directional function (the so-called MUSIC spectrum), $D(\mathbf{k})$,

$$D(\mathbf{k}) = \frac{1}{|\mathbf{a}(\mathbf{k}) \cdot \mathbf{V}_n|} \quad (5)$$

where \mathbf{V}_n is the matrix formed with the noise eigenvectors of the covariance matrix. Once the signal vectors are determined, we can compute the propagation direction and phase velocity of the signals propagating through the array, as functions of time and frequency.

The most critical point in the method is the estimation of the number of signals present in the records. This number determines the rank of signal and noise subspaces. Wax and Kailath (1985) proposed two estimators based on the information theory: the AIC (Aikake Information Criterion) and the MDL (Minimum Description Length) tests. These two methods are quite simple, but in the case of few receivers the probability of overestimating the number of signals is large (Marcos, 1998). If the number of signals is not correctly estimated, the resulting parameters of the signals (phase velocity and propagation direction) will be incorrect. A different way to determine the number of signals is a threshold criterion (Marcos, 1998; Cornou, 2003). It consists in searching a rupture in the eigenvalues profile by means of a statistical or empirical estimator. Two of the advantages of the latter criterion are that it does not require assuming that the noise included in the records has a white spectrum, and that it can be used even with few stations. In our study, we used a threshold criterion with an empirical estimator determined by trial and error tests on both synthetic and real data.

A refinement of MUSIC, proposed in Goldstein and Archuleta (1991a) improves its resolution in the case of correlated signals (a very relevant hypothesis for Mexico City, where we expect multipathing and mixed modes of propagation). However, this requires dividing the array of stations into linear subarrays, and then computing averages for those subarrays with the same orientation and interstation distance. Unfortunately, the geometry of the arrays where our data come from does not allow the implementation of this technique.

Estimation of uncertainties

We require an estimate of the uncertainty associated to the results obtained by the f - k methods and MUSIC. In the case of MUSIC, Goldstein and Archuleta (1991a) proposed a technique that requires the estimation of the signal-to-noise ratio for each frequency and each time window

which is quite impractical. For this reason, and in order to have a common estimation method for the f - k methods and MUSIC, we estimate the uncertainty of the results using the width of the spectrum peak (Cornou, 2002). We first estimate the width of the spectrum peak at a fixed value relative to the maximum value (*e.g.*, 90% of the maximum) in both radial and transverse directions. Then, we determine the velocity and azimuth uncertainties as

$$\frac{\Delta V}{V} = \frac{\Delta f}{f} + \frac{\Delta k_r}{k} \quad (6)$$

and

$$\Delta\theta = 2 \tan^{-1} \left(\frac{\Delta k_t}{2k} \right) \quad (7)$$

where $\Delta\theta$ represents the azimuth uncertainty, ΔV the velocity uncertainty, Δf the frequency sampling rate, Δk_r the width of the spectrum peak in the radial direction and Δk_t the width of the peak in the transverse direction. The sharper the spectrum peak, the more precise the results, and the smaller the uncertainty.

TEST OF THE METHODS USING SYNTHETIC DATA

Before their application to recorded data, the three methods described were tested using synthetic data to test their limitations. The array configuration used was the geometry of the Colonia Roma array that recorded the Coyuca 10/08/2001 earthquake; stars in Figure 1 represent the 6 stations that composed this array. We chose this geometry because of the poor spatial distribution of the records, giving rise to aliasing problems. The minimum spacing between stations is 58 m (between JPSK and RMAS) and the maximum spacing is 570 m (between PRJS and PROM).

For each station in this array, a synthetic seismogram was computed assuming a given direction of propagation, θ , and phase velocity V . The signal chosen for this was the Ohnaka pulse (Almendros, 1999)

$$s(t) = A \left(\frac{t}{t_0} \right)^B e^{-t/t_0} \sin(2\pi ft) \quad (8)$$

where A is related to the signal amplitude, t_0 describes the amplitude decrease of the envelope, B defines the shape of the signal (impulsive or emergent), and f the dominant frequency of the signal. This pulse allows the development of both short and long wavetrains.

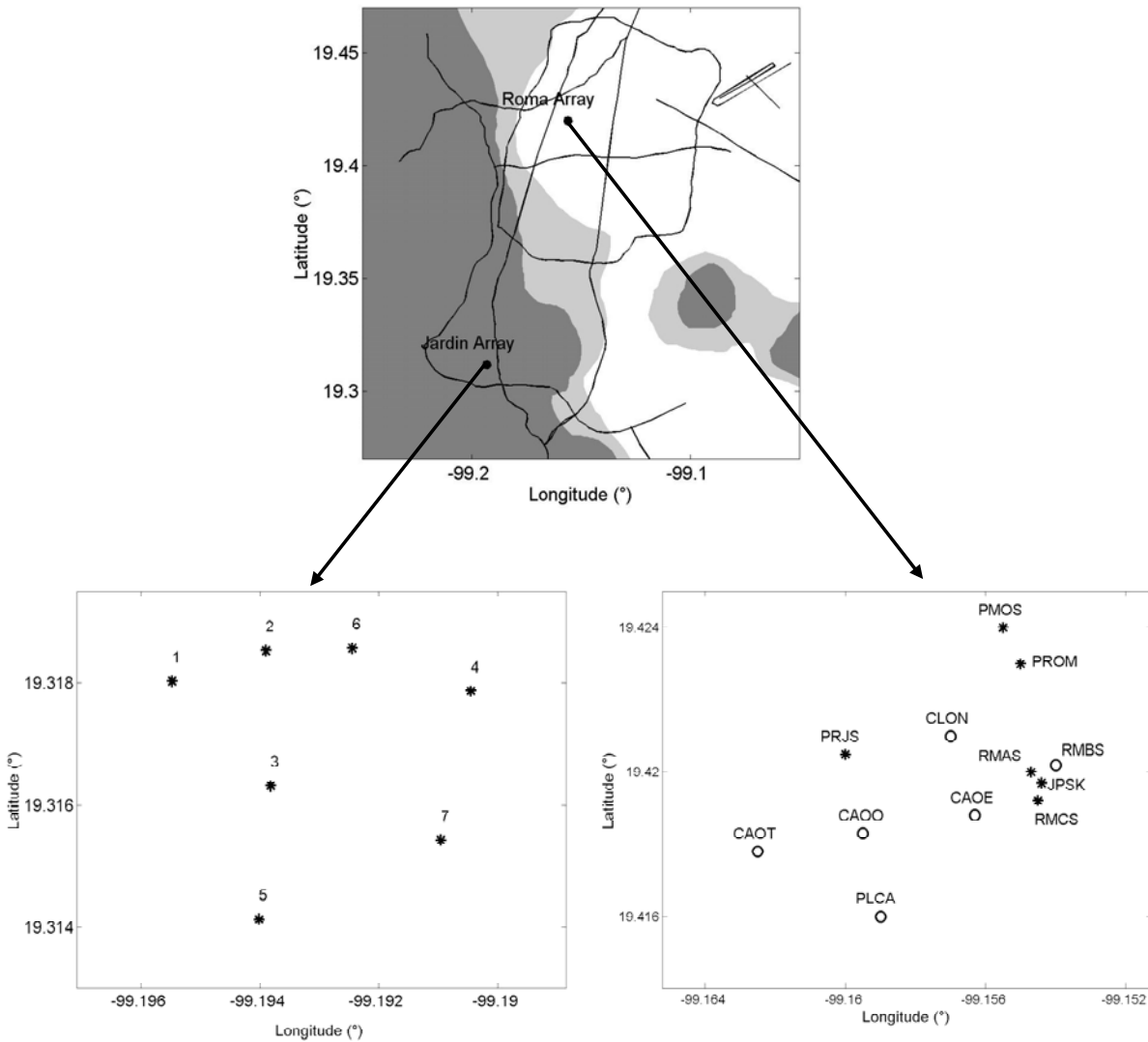


Figure 1. (Top): Map of Mexico City showing the location and geometry of Jardin Array and Colonia Roma array. Different shadings indicate the geotechnical zonation: dark grey corresponds to the firm zone, light grey corresponds to the transition zone and white corresponds to the lakebed zone. The solid lines indicate main streets in the city. The map for each array (bottom) shows with stars the stations that recorded the analyzed event. Open circles are used to indicate the stations that did not record the studied event.

Case of non-simultaneous signals

Goldstein and Archuleta (1987, 1991a) showed experimental tests comparing Capon's high-resolution method and MUSIC using a linear array with equally spaced stations. With this

configuration, they showed that both methods have similar resolution capabilities when a single signal is present. We repeated this test for our very irregular array, using two successive signals with the same dominant frequency and different propagation velocity and back azimuth (tests 1 and 2, Table 1). We first used signal parameters out of the aliasing zone, with a dominant frequency of 0.25 Hz and propagation velocities of 1000 and 2000 m/s (test 1, Table 1). The back azimuth values (200° and 250°) correspond to the expected back azimuth for earthquakes from the Guerrero subduction zone in Mexico City, and to the back azimuth with the strongest secondary peaks of the spectrum for a unit impulse excitation, respectively. In this case, the three methods give good results both for the velocity and the propagation direction. The precision, however, is better for the high-resolution $f-k$ method and MUSIC. The conventional $f-k$ method tends to underestimate the velocity of the second signal. Thus, results for the linear array tested in Goldstein and Archuleta (1991a) can be applied to the case of a very irregular array.

Table 1. Parameters for the synthetic signals used in the test.

Test number	Signal 1				Signal 2				Mixed signals?
	Freq. (Hz)	Max. Amp.	Velocity (m/s)	Baz ($^\circ$)	Freq. (Hz)	Max. Amp.	Velocity (m/s)	Baz ($^\circ$)	
1	0.25	100	1000	200	0.25	75	2000	250	no
2	1	100	50	200	1	75	150	20	no
3	0.25	100	1000	200	0.25	75	2000	250	yes
4	0.25	100	1000	200	0.33	75	2000	250	yes

We also test the case of very slow signals with propagation velocities of 50 y 150 m/s and a dominant frequency of 1 Hz (test 2, Table 1). This test comes in order because of the low propagation velocities of the clay layer in Mexico City's lakebed zone. In this case, the signal wavelengths are shorter than half the maximum spacing between stations: aliasing becomes a problem. We have to use a large time analysis window of 12 cycles and a maximum wavenumber of 0.2 rad/m otherwise none of the methods will give acceptable results, because of aliasing. Under these conditions, the $f-k$ methods provide a good estimation of propagation velocities and of the second signal back azimuth (Figure 2). For the first signal, the conventional $f-k$ method does not give a good estimate of the back azimuth. On the contrary, the high-resolution $f-k$ results are centered on the correct back azimuth but show a large dispersion. MUSIC is unable to solve this problem.

In case of aliasing, the high-resolution $f-k$ method, and in a smaller proportion the conventional $f-k$ method, is able to determine the propagation parameters of the signals but it requires the use of large time windows and large grids. This is not possible in a real case with non-stationary signals. Moreover, the appearance of the $f-k$ spectra (not shown here) with an unclear primary peak and numerous secondary peaks contributes to the uncertainty of the results: in a more complex case with noise and possible mixed signals, the results would not be very reliable. Complementary tests showed that the minimum velocity that can be solved by a reasonable time window length (the optimum length is about 3 cycles) corresponds to a maximum wavenumber of 2.5 km^{-1} .

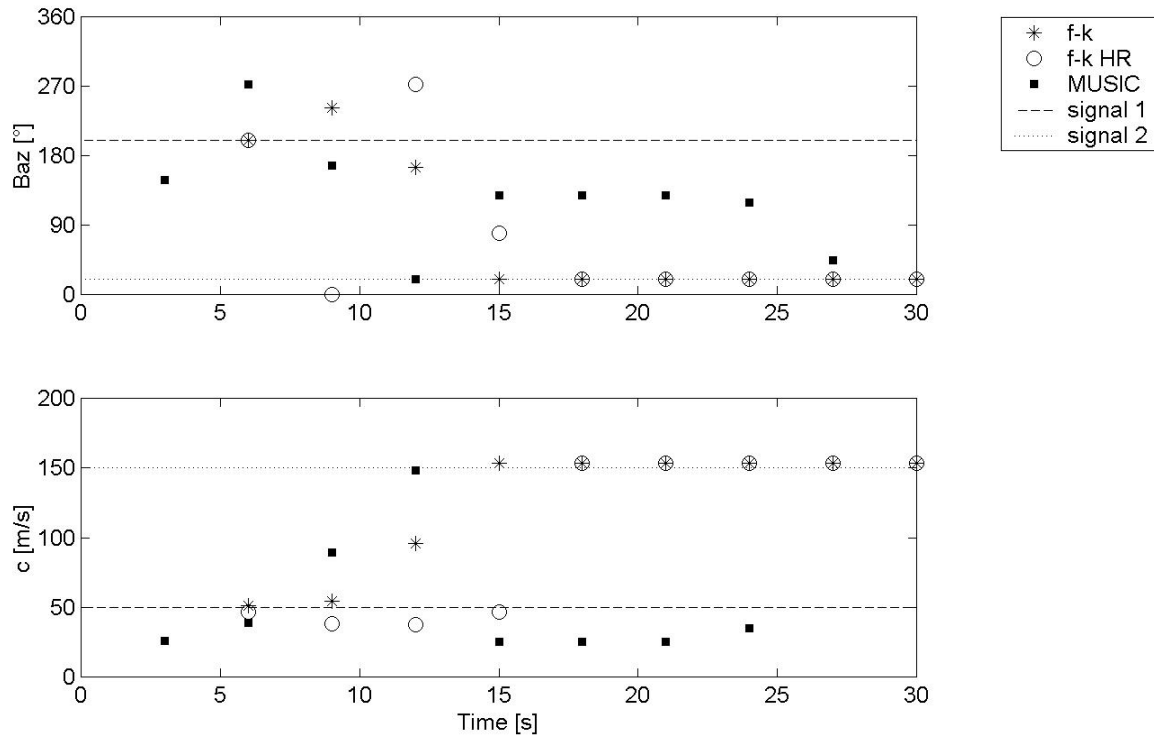


Figure 2. Results of the analysis using synthetic signal for test 2 (Table 1) for our three methods. The synthetic input consists of two simultaneous signals, for the case of strong spatial aliasing. The top diagram shows the back azimuth as a function of time. The bottom diagram shows the phase velocity as a function of time. The dashed line represents the input parameters for signal 1; the dotted line shows the input parameters for signal 2.

Case of simultaneous signals

We evaluated the three methods of analysis in the case of simultaneous arrival of two signals at the array. These tests correspond to cases 3 and 4 in Table 1. The first case corresponds to two signals with the same frequency content and different amplitude. The second case corresponds to two signals with different frequency content and amplitude.

Results for the first of these cases (test 3 in Table 1) are shown in Figure 3. This figure displays the phase velocity and back azimuth determined by each method as a function of time along the seismograms. The three methods give similar results. They cannot separate the two signals and solve the problem identifying only one signal with parameters that are an average of the parameters of the two original signals. As the first input signal has larger amplitude, the results are closer to its parameters.

The second test (test 4 in Table 1) consisted in two signals of different frequency and amplitude. Results are shown in Figure 4, again as phase velocity and back azimuth against frequency. Figure 4 shows that both f - k methods correctly identify the two different wavetrains

present in the seismograms. In terms of phase velocity resolution, the three methods obtain the correct tendency but the results obtained from the f - k methods are more precise than those from MUSIC. The high-resolution f - k method slightly underestimates the velocity values (about 750 m/s instead of 1000 m/s for the lowest frequency signal and about 1750 m/s instead of 2000 m/s for the second signal) and give more disperse results than the conventional f - k method. MUSIC is not able to distinguish the two directions of arrival; it only “sees” the back azimuth of the lowest frequency signal. Thus, in the case of mixed signals, f - k methods are more stable than MUSIC, and the conventional f - k seems to be the best method.

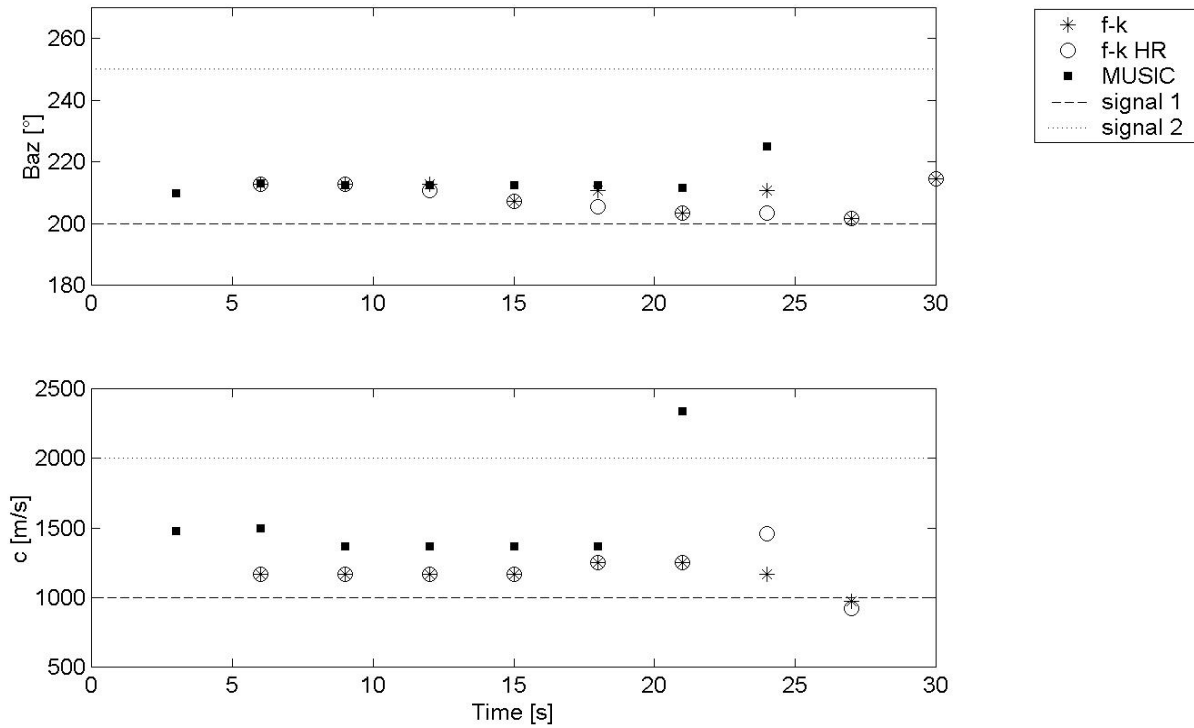


Figure 3. Results of the analysis using synthetic signal for test 3 (Table 1) for our three methods. The synthetic input consists of two mixed signals of different amplitude and the same central frequency. The top diagram shows the back azimuth as a function of time. The bottom diagram shows the phase velocity as a function of time. The dashed line represents the input parameters for signal 1 and the dotted line shows the input parameters for signal 2.

Influence of the noise

The last test using synthetic data consisted of introducing noise in the synthetic seismograms of test 1. The frequency-wavenumber spectra obtained from the three methods when the seismograms are noise free are shown in Figure 5 (left column). The results show important differences between the resolutions of the three methods. Conventional f - k shows a very broad

single peak whereas the high-resolution $f-k$ and MUSIC show sharper peaks. MUSIC is able to estimate without uncertainty the location of the propagating wave.

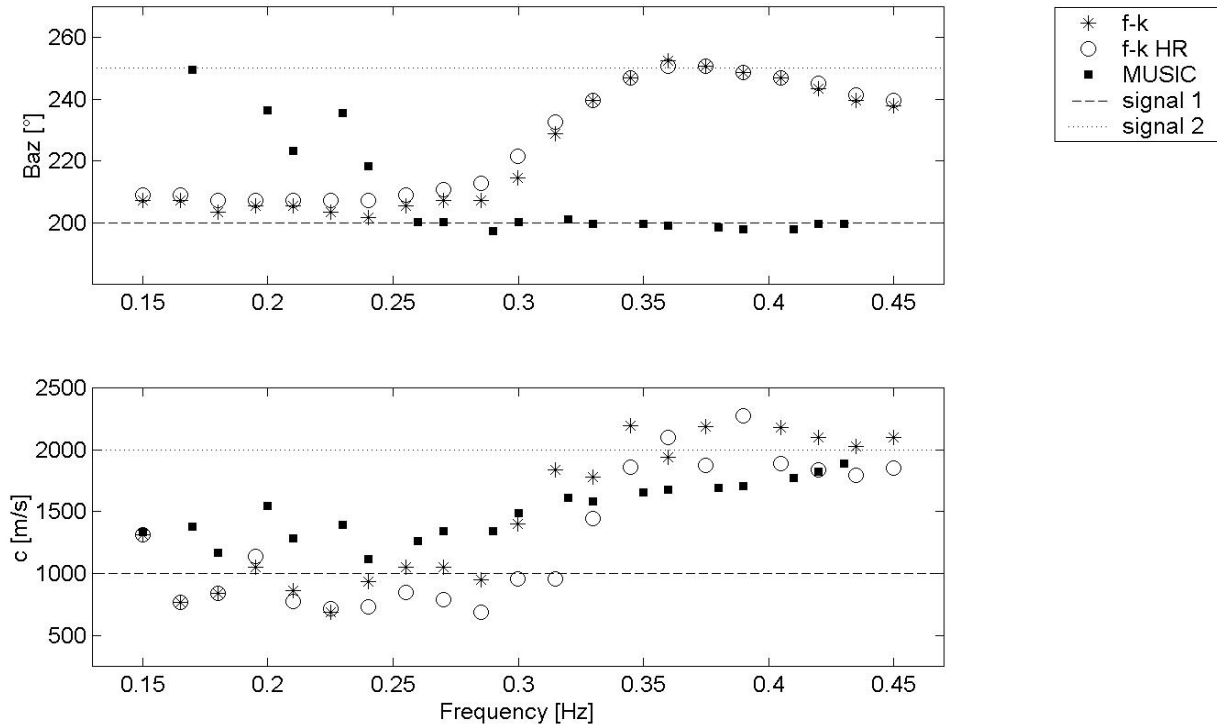


Figure 4. Results of the analysis using synthetic signal for test 4 (Table 1) for our three methods. The synthetic input consists of two mixed signals of different amplitude and the different central frequency. The top diagram shows the back azimuth as a function of time. The bottom diagram shows the phase velocity as a function of time. The dashed line represents the input parameters for signal 1 and the dotted line shows the input parameters for signal 2.

We have repeated this computation, adding white noise characterized by its variance amplitude equals to 25% of the maximum amplitude of the signal. This test (Figure 5, right column) shows that the presence of noise increases the uncertainties of the results of the high-resolution $f-k$ method, whereas the resolution of the MUSIC and conventional $f-k$ methods is not affected.

APPLICATION TO REAL DATASETS

The methods we have described and tested have been used to analyze two different datasets, recorded by two different arrays. The first one comes from a temporary network installed on firm soil in Mexico City, the Jardin array (see Figure 1). The second was recorded by a new, semi permanent, dense array installed in Colonia Roma in downtown Mexico City. For each array we analyze only one event. The data from Jardin array has been analyzed in Barker *et al.* (1996).

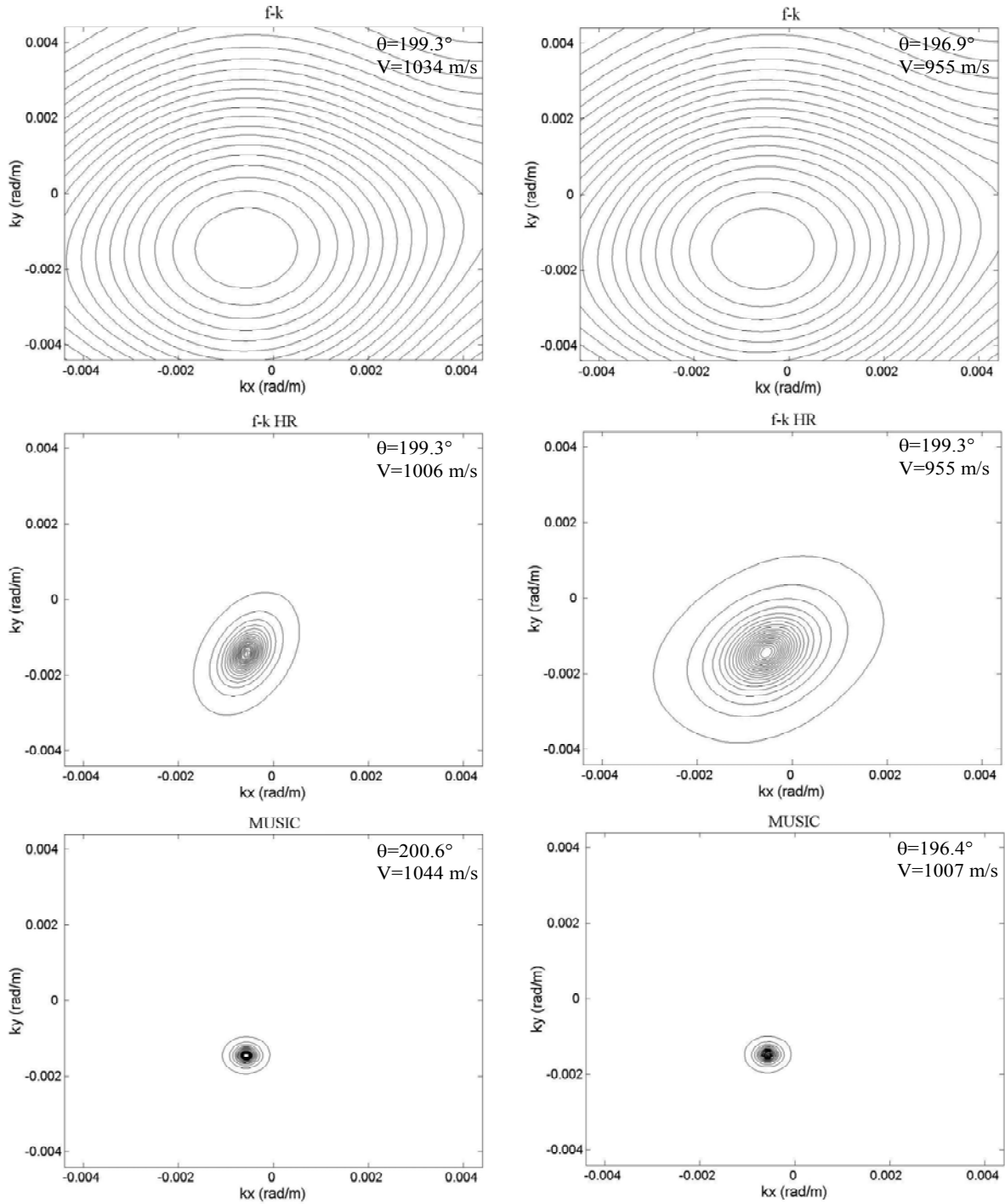


Figure 5. Amplitude spectra obtained in the horizontal wavenumber domain for our three different methods. The synthetic input motion consisted of only the first signal for test 1 (Table 1). The diagrams on the left column were computed for noise-free signals. The diagrams on the right column correspond to the same signals, to which white noise has been added (variance amplitude equals to 25% of the maximum amplitude of the signal). Numerical results are indicated on the top of each diagram.

We have chosen one of the events they analyzed, to test our methods from a dataset with a large coherency among the records. In the case of Colonia Roma array, we use the first dataset recorded by this new array. The two analyzed events originated in the Pacific coast of Mexico, with a very similar azimuth and epicentral distance to Mexico City.

Data from the Jardin array

From April 1 to May 14, 1994, a temporary array of seven three-component, five second Lennartz seismometers was installed in the botanical gardens of the National Autonomous University of Mexico. Data were recorded at a rate of 25 samples per second, in 16 bit digital seismographs. The array was located on firm soil, on the lava flows originated in the Ajusco range, responsible for closing the valley and transforming it into a closed basin approximately 700,000 years ago. The distance between stations was 150 m, and its aperture was 500 m in the NS and EW directions (see Figure 1). The event we analyze was recorded on May 5, 1994, and its relevant data are given in Table 2. The short distance between stations, and the large propagation velocity of the lava produced extremely coherent records. Average coherency is larger than 0.95 among all stations.

Table 2. Data of the two analyzed earthquakes.

Location	Date	Origin time (GMT)	Distance (km)	Baz (°)	Depth (km)	Magnitude
Guerrero	05/05/1994	12:18:50.7	337.4	205.0	41	4.4
Coyuca	08/10/2001	03:39:20	289.7	201.7	16	6.1

The very large coherency between stations made it very difficult to the determine phase differences among records. Barker *et al.* (1996) suggested that $f-k$ analysis did not have the necessary resolution and they devised a method to measure phase differences among stations smaller than the time sampling step. We have analyzed these data using our three methods in the period range 5 to 10 sec. The traces, bandpass filtered in this period range, are shown in Figure 6. To process the records, we computed wavenumber spectra using time windows of 20 seconds duration with a 50% overlap.

Results are shown in Figure 7 in terms of phase velocity and back azimuth as a function of time. MUSIC (black squares) and the conventional $f-k$ method (stars) give similar coherent results, but the high-resolution $f-k$ method (open circles) offers poor results in terms of velocity (many values were obtained from peaks close to the origin, where a velocity of 7 km/s is equivalent to infinity) and back azimuth (the values are incoherent and do not correspond to the expected trend described in Barker *et al.*, 1996). The results for the conventional $f-k$ method are quite reasonable, contrary to the hypothesis of Barker *et al.* (1996) that $f-k$ methods would be ineffectual. Nevertheless, it presents too large uncertainties that make the results unreliable. On the contrary, MUSIC presents coherent results both for propagation velocity and back azimuth, and agrees very well with the previous study of these data. MUSIC can be considered as a useful alternative to $f-k$ methods when time delays between stations are very small.

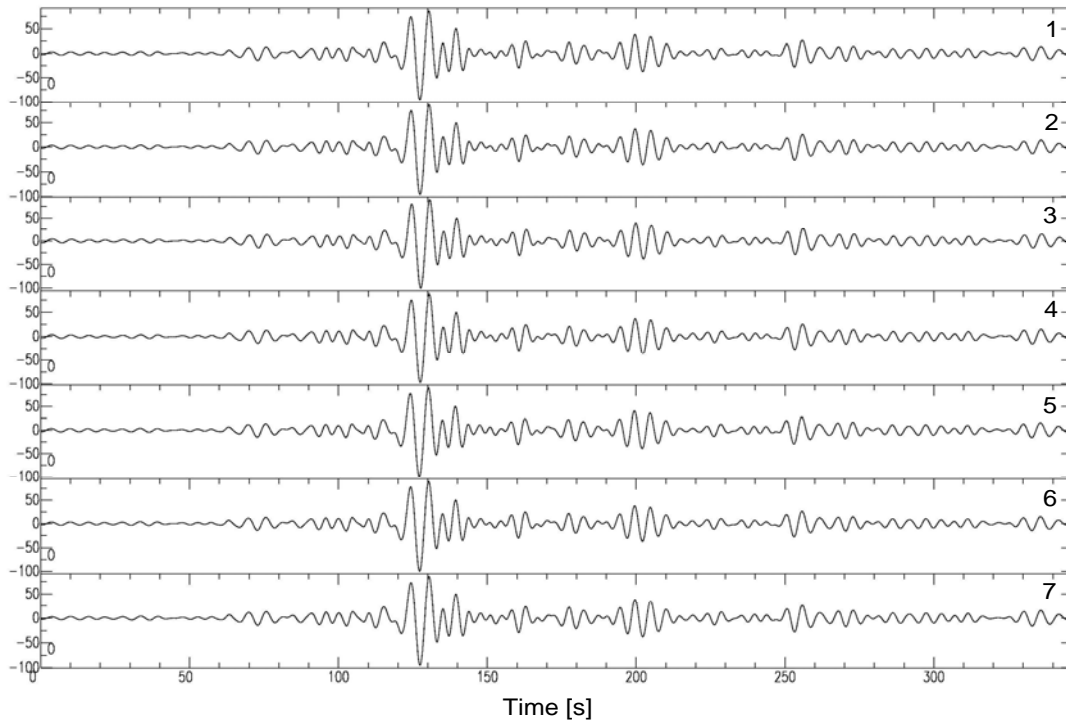


Figure 6. Traces recorded by the Jardin array, in UNAM's botanical garden, for the 05/05/94, Guerrero, earthquake. The traces have been band-pass filtered in the period band 5 to 10 sec.

At large periods, the signal comes mainly from the epicenter (back azimuth of 205°) with no clear evidence of multipathing for the first 200 sec. After 200 sec, we observe a greater dispersion of the back azimuth which values between 160° and 220° . In terms of propagation velocity, the first 80 sec of signal show large phase velocities (larger than 4000 m/s) with smaller values after 80 sec (between 2000 and 3000 m/s). These results agree with the previous study of Barker *et al.* (1996).

MUSIC gives very stable results for the back azimuth with uncertainties smaller than 10° , except for the last 100 s of the record, where the uncertainties attain 35° . The trend of phase velocity as a function of time computed from MUSIC is similar to that calculated by conventional f - k method, but shows larger values for the first 100 sec. The values of the uncertainty for the phase velocity are quite large (up to 3700 m/s for the first arrivals and up to 1500 m/s for the rest of the trace). They indicate that the absolute values are not reliable, even if the synthetic tests showed us we can trust the main trend.

Results obtained from the analysis of this dataset, recorded on the firm zone of the Valley of Mexico, south of the lakebed zone, give important information on the incident wavefield at this valley for earthquakes occurring in Guerrero coasts. The first 200 sec of the trace are composed of signals arriving from the epicenter with rapid early waves (interpreted as P or L_g waves by Barker *et al.*, 1996) and slower surface waves in the second part of the records. Late

arrivals (after 200 sec) show significant dispersion of their direction of propagation, with variations up to 40° from the epicenter.

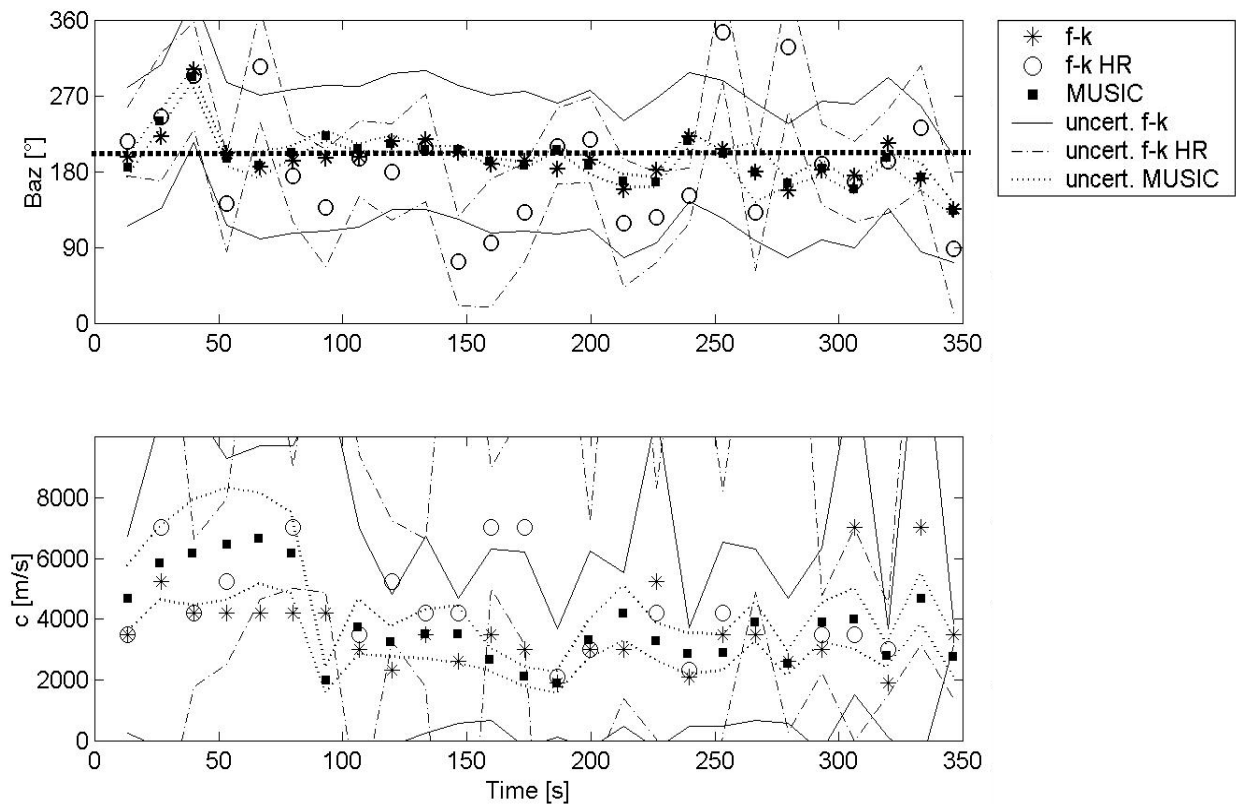


Figure 7. Results of the analysis for the traces shown in the previous figure. The top diagram shows the back azimuth as a function of time, while the bottom diagram corresponds to the phase velocity. Results of the conventional f-k method are shown with stars, open circles correspond to the high-resolution f-k method, while results using MUSIC are shown with black squares. The solid line indicates the uncertainty limits computed for the conventional f-k method, the dash-dotted line represents the uncertainty limits for the high-resolution f-k method, while the thin dotted line shows the uncertainty limits for MUSIC method. The thick dotted line (bottom) indicates the epicentral direction.

Data from Colonia Roma dense array

As part of the instrumentation program conducted in Mexico City after 1985, CENAPRED, Mexico, installed a synchronized array of 5 stations in Colonia Roma, the Roma array. This array included 3 stations at the surface, with an average spacing of 115 m, plus two additional borehole accelerometers at 30 m and 102 m under one of the surface stations. The results obtained with this array (Cárdenas-Soto, 2000), prompted us to install a semi permanent array using Roma array as the starting point. An additional 9 surface stations and two borehole accelerometers, at 50 m depth, were installed. The spatial distribution of the new stations (Figure 1) was dictated by

the location of parks and public spaces in this densely urbanized neighborhood in Mexico City. The new array became operational during mid 2001. Its aperture is 1 km, average distance between stations is 313 m, while minimum interstation distance is 57 m. The first event recorded by several stations occurred in October 2001. The analysis of this dataset is presented in the next paragraphs.

The Coyuca event of October 8, 2001, was recorded by 10 stations (6 surface stations and 4 borehole ones), out of the 16 available. The triggering parameters of the new stations were still being calibrated when the earthquake occurred. For this reason, many stations did not record this event, and some of the records obtained are too short. The recorded traces, bandpass filtered in the period bands 1.29-2.14 sec (smaller than the dominant period at site) and 3.04-5.07 sec (larger than the dominant period at site), are shown in Figures 8 and 9. The sampling rate was 100 Hz. Figure 10 shows the average coherency between traces for the radial component, comprised mainly between 0.5 and 0.7. The results of the analysis will be presented for the radial component, although our conclusions are derived from the analysis of the three components of motion.

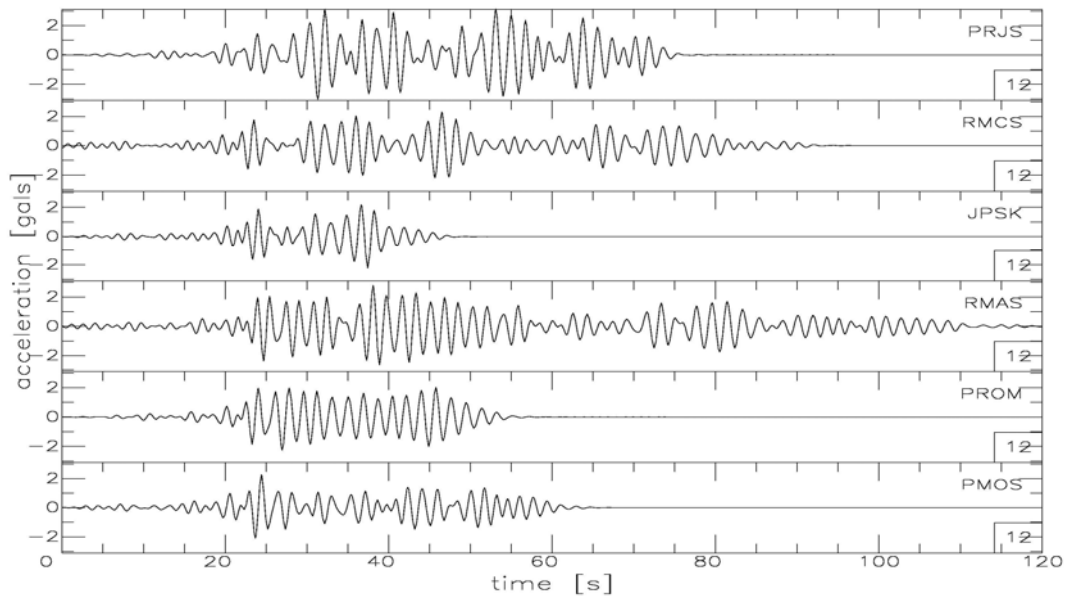


Figure 8. Radial component traces recorded by the Colonia Roma array for the 10/08/2001, Coyuca, earthquake. The traces have been band-pass filtered in the period band 1.29 to 2.14 sec.

To process the records, we used frequency dependent time windows with duration of 3 cycles, with an overlap of 50% between two successive windows. Figures 11 and 12 show the results for phase velocity and back azimuth as a function of period. At each period, each symbol shows the result for an individual time window. In terms of phase velocity, and for the periods smaller than the resonance period of the site, the $f-k$ methods give similar trends with small velocities between 200 m/s and 700 m/s. For longer periods, both $f-k$ methods give larger

velocities: between 400 m/s and 1500 m/s for the conventional $f-k$ method, and between 600 m/s and 2000 m/s for the high resolution $f-k$ method.

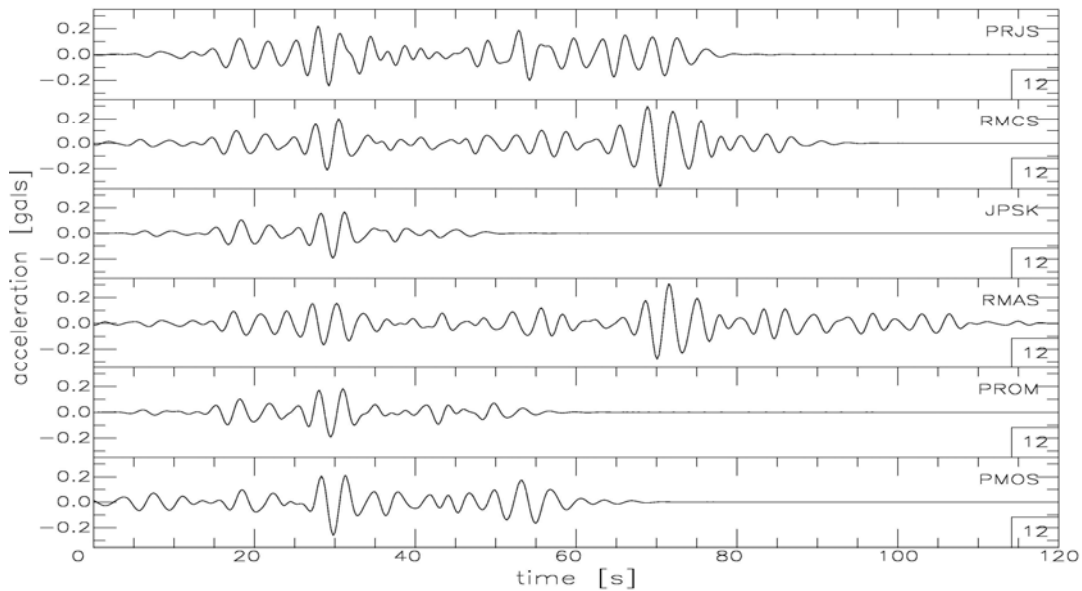


Figure 9. Radial component traces recorded by the Colonia Roma array for the 10/08/2001, Coyuca, earthquake. The traces have been band-pass filtered in the period band 3.04 to 5.07 sec.

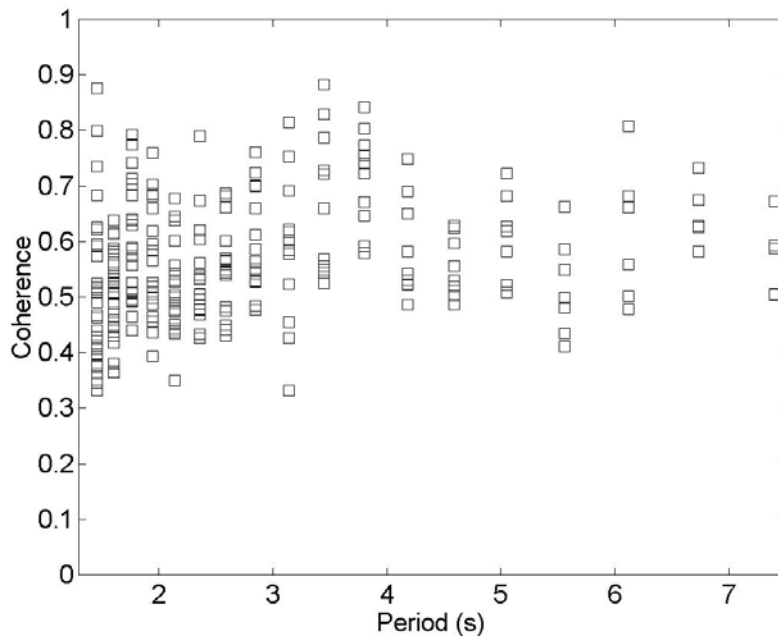


Figure 10. Average coherence between stations as a function of time and frequency for the radial component traces recorded by the Colonia Roma array for the 10/08/2001, Coyuca, earthquake. Each rectangle is centered at the time-frequency window used for the computation.

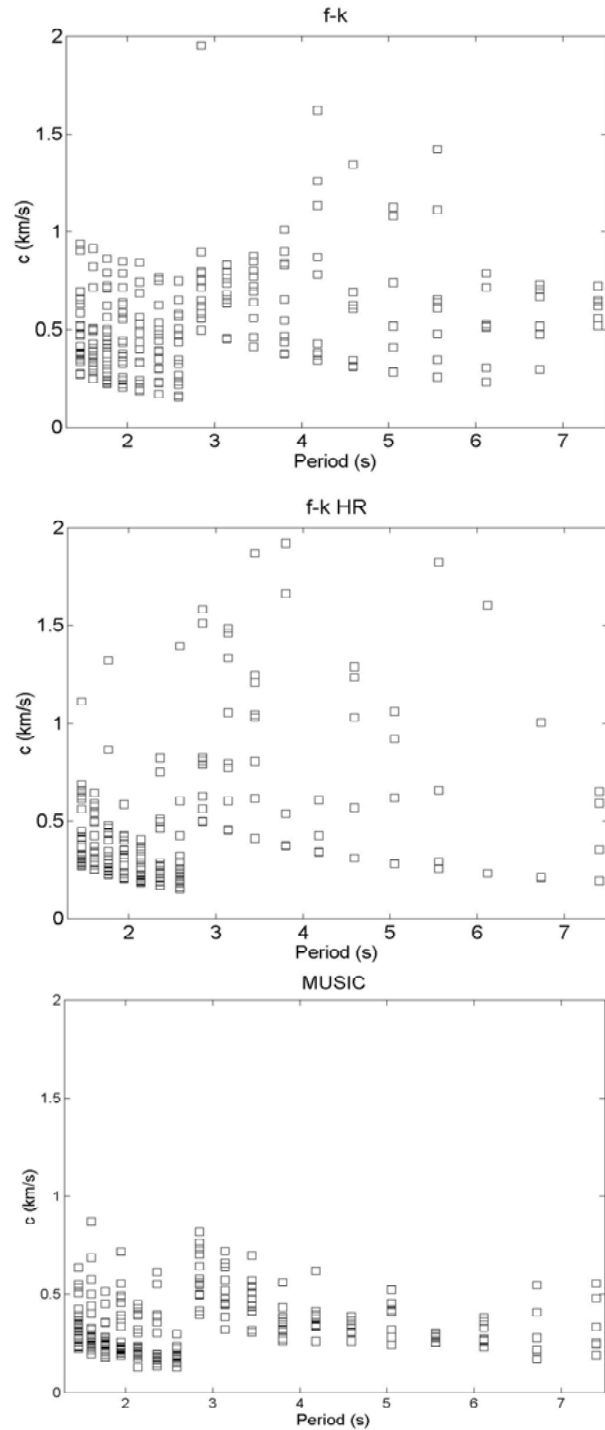


Figure 11. Phase velocity as a function of period computed for the radial component traces recorded by the Colonia Roma array for the 10/08/2001, Coyuca, earthquake. Top diagram: conventional f-k method. Middle diagram: high-resolution f-k method. Bottom diagram: MUSIC method. The different symbol for each period value corresponds to different time windows along the seismograms.

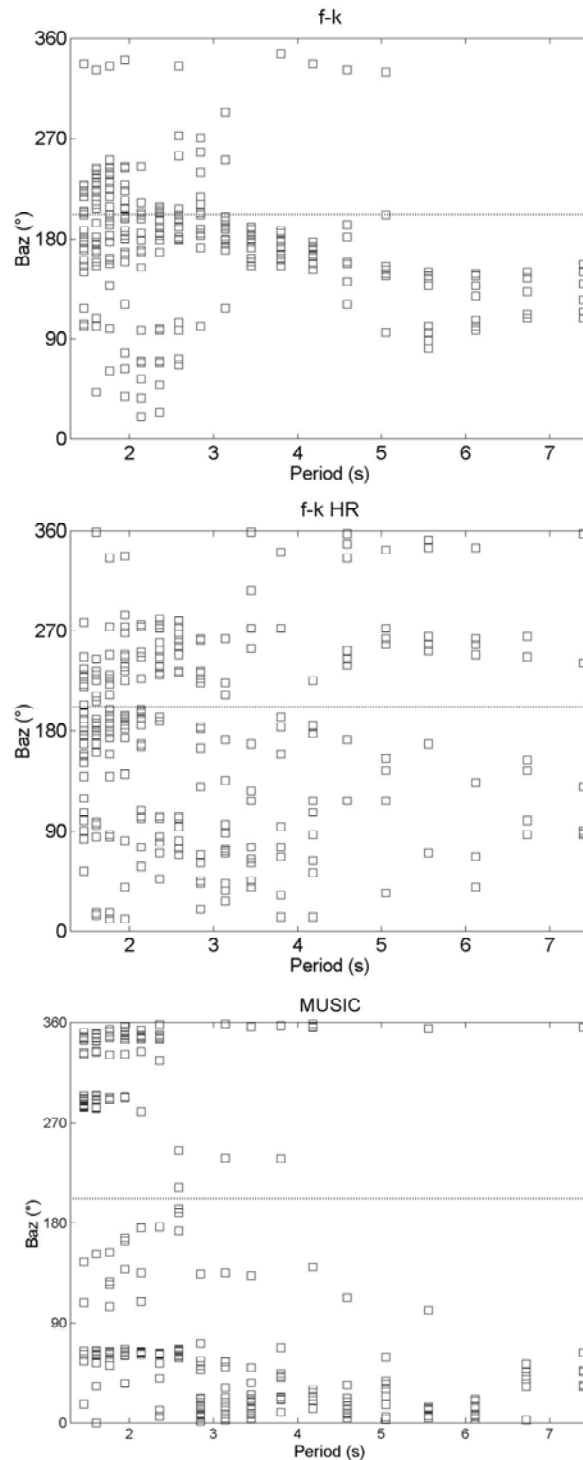


Figure 12. Back azimuth as a function of period computed for the radial component traces recorded by the Colonia Roma array for the 10/08/2001, Coyuca, earthquake. Top diagram: conventional f-k method. Middle diagram: high-resolution f-k method. Bottom diagram: MUSIC method. The different symbol for each period value corresponds to different time windows along the seismograms. The dotted line indicates the epicentral direction.

The results obtained by MUSIC do not show this discontinuity at 2.5 sec: all phase velocity values are comprised between 200 m/s and 700 m/s, even for the largest periods where we expect larger velocity values. At short periods, the velocity values are limited by aliasing. Synthetic tests showed that the smallest velocity values (200 m/s) could correspond in fact to phase velocities in the range 50-100 m/s (the shear wave velocity in the soft clay layer). The velocity values obtained by the high-resolution $f-k$ method and, to a smaller degree by the conventional $f-k$ method, agree with the lakebed zone velocity profiles described by Iida (1999). The velocities smaller than 800 m/s, obtained by the conventional $f-k$ methods for periods larger than 4 sec, do not agree with the expected velocities described by the velocity model of Campillo *et al.* (1996) or the precedent work of Chávez-García and Salazar (2002). They are not reliable and could be explained by the weakness of the signal at long periods or by an artifact of the method caused by the irregular geometry of the recording stations.

In terms of back azimuth, the three methods give very different results, as show in Figure 12. For periods smaller than 2.5 sec, the conventional $f-k$ method indicates that wavetrains come from a wide range of directions, mainly from the south-southwest (between 160° and 260°) with a secondary arrival from the east (about 90°). For periods between 2.5 and 3.5 sec, the back azimuth shows small dispersion and points to the epicenter, about 200° . For periods larger than 4 sec, we observe a rotation of back azimuth towards south-southeast (from 180° to 140°). The results using the high-resolution $f-k$ method are similar for short periods, but show larger dispersion at large periods and it becomes difficult to determine a coherent trend. The results using MUSIC are unacceptable, especially at large periods, where we expect wavetrains arriving mainly from the epicenter (Chávez-García and Salazar, 2002).

The comparison of the three methods allows the following observations. $f-k$ methods seem more adequate to process data tainted by low coherency and spatial aliasing. Under these conditions, MUSIC gives erroneous results for both phase velocities and back azimuth values. The conventional $f-k$ method gives good results in terms of back azimuth but underestimates the velocity values for large periods, due to aliasing. In contrast, the high-resolution $f-k$ seems to be more efficient in terms of velocity but cannot solve correctly the back azimuth for periods larger than the resonance period. Despite its problems, the conventional $f-k$ method was the only method to give coherent results both for the velocity and the propagation direction. It seems to be the best method for the processing of data recorded the lakebed zone, but we have to take into account the weaknesses of the method in the interpretation of the data.

Let us consider now the results in terms of their contribution to the understanding of the wavefield in the lakebed zone of Mexico City. The analysis of the records from the Coyuca event of October 8, 2001 allows the following conclusions. At periods smaller than 2.5 sec (the resonance period at the site of Roma array), the wavefield includes both slow (with velocities probably smaller than 100 m/s) and fast wavetrains (with velocities up to 800 m/s) coming mainly from the south-southwest (between 160° and 260°), with some arrivals from the east (about 100°). This can be interpreted as the result of the interaction between diffracted waves coming from the west of the lakebed zone and the resonance of the soft surficial clay layer. In the band between 2.5 and 3.5 sec period, we observe a significant increase in the phase velocity, with values between 700 m/s and 2000 m/s. In this period band, directions of propagation are very

coherent, with small dispersion, and come from the epicenter (about 200°). These arrivals are surface waves guided by the deep layers of the crust and coming from the epicenter, as described by Chávez-García and Salazar (2002). For periods larger than 3.5 sec, we observe a similar velocity trend, with values between 700 m/s and 2000 m/s, and a systematic rotation of back azimuth towards southeast (from 180° to 140°). This change of back azimuth was only observed in the data recorded in the lakebed zone. However, Barker *et al.* (1996) observed similar back azimuth rotations in the data from Jardin array for earthquakes coming along different paths (in particular for events with epicenters in Puebla, San Luis Potosí and Texcoco). This rotation of back azimuth could be interpreted as a path effect due to the irregular crustal structure in central Mexico.

CONCLUSIONS

We have tested three different methods, proposed in the literature, in terms of their suitability to analyze the wavefield that propagates in Mexico City: conventional and high resolution f - k spectra, and MUSIC. This evaluation was necessary because of the complex character of ground motion in the Mexico City basin, where local site effects are determinant in the generation of destructive ground motion. However, the very large amplification caused by the very soft clay sediments makes it extremely difficult to analyze the driving wavefield. We have analyzed the results by using those three methods with synthetic datasets. This allowed determining their advantages and disadvantages in terms of the expected characteristics of recorded ground motion in Mexico City.

The synthetic tests showed that the f - k methods are more robust than MUSIC in presence of simultaneous arrivals, but result in broader peaks, in particular in presence of noise, which limits their resolution. When spatial aliasing is large, none of the methods provides a correct estimate of the signals parameters in presence of multiple signals. In addition, our tests showed that the resolution is limited to a maximum wavenumber of 2.5 km^{-1} , in the case of the geometry of the array that recorded the Coyuca event.

The three methods were used to analyze two different datasets, obtained by two different arrays. Jardin array had a very regular geometry and a small interstation distance. In addition, it was installed on firm soil (ancient lava flows) with large propagation velocity. For these reasons, average coherency between stations was very large (larger than 0.95) and phase delay between stations very small. We showed that, in this particular case, MUSIC is the most adequate method to process these data, since it gives both accurate results and reasonable uncertainties. The conventional f - k method gives coherent results but with very large uncertainties, whereas the high-resolution f - k method cannot solve the problem neither for the velocity nor for the back azimuth. The results are in good agreement with the results presented in Barker *et al.* (1996) and give some useful indications about the wavefield incident on the lakebed zone. The first part of the trace (until 80 sec) is composed of fast waves (with velocities superior to 4000 m/s), probably L_g waves or P waves as interpreted by Barker *et al.* (1996), coming from the epicenter. After 80 sec, the wavetrains have smaller phase velocities (between 2000 and 3500 m/s) but still come from the epicentral direction with small dispersion (about 10° between 80 and 200 sec and up to

40° for the late part of the records). These wavetrains are surface waves generated or refracted between the epicenter and the valley of Mexico and guided by the deep layers of the basin (Chávez-García and Romero-Jiménez, 1998; Chávez-García and Salazar, 2002).

Finally, the three methods were applied to the first dataset recorded by the new dense array installed in Colonia Roma. The quality of this dataset was low because the network was still undergoing adjustments in the triggering and recording parameters of the stations when the earthquake occurred. In spite of these problems, the analysis demonstrated that the conventional $f-k$ method is the best tool to process such complex data. The high-resolution $f-k$ method gives good results for the phase velocity, but is not able to estimate precisely the propagation direction, a critical parameter for the comprehension of the wavefield. The low quality of the dataset and the aliasing prove to be fatal for MUSIC method, which is inefficient under these conditions. In terms of understanding the observed ground motion, the analysis of the results showed without doubt the critical change suffered by the wavefield at the site period. For periods smaller than the resonance period (2.5 sec at the site of the array), the wavefield consists of both slow waves (with phase velocities smaller than 100 m/s) and fast waves (with phase velocities about 800 m/s). These waves seem to arrive from multiple directions, mainly from the west (from 160° to 260°, as observed by Chávez-García and Salazar, 2002), with a secondary arrival from the east (about 100°). They could result from the interaction between diffracted waves generated between the coast and the lakebed zone and the local resonance of the soft clay layer. For periods between 2.5 sec and 3.5 sec, we observed waves with phase velocities between 700 and 2000 m/s, coming from the epicenter (200°) with small dispersion (smaller than 40°). These waves correspond to surface waves guided by the deep layers of the basin. At periods larger than 3.5 sec, phase velocities are similar but a clear rotation of back azimuth towards southeast (from 180° to 140°) appears in the data. This rotation could be similar to that described by Barker *et al.* (1996) in the firm zone of Mexico City. These observations support the hypothesis that the large duration of strong ground motion in the lakebed zone of Mexico City is the result of the interaction of different modes of surface waves with the local resonance of the soft clay layer (Chávez-García and Romero-Jiménez, 1998; Chávez-García and Salazar, 2002).

ACKNOWLEDGEMENTS

The Colonia Roma Dense Array was installed and is maintained by the Instrumentation Coordination of the Instituto de Ingeniería, UNAM. We would like to thank all the people which have participated to the installation and still participate in the maintenance of this array. The data from Roma array was provided by CENAPRED. Jeffrey Barker from State University New York, Binghamton, kindly provided the data recorded by Jardin array. This research was done while one of us (AR) was a PhD student at UNAM with support from the Secretaría de Relaciones Exteriores, Mexico, and the Ministère des Affaires Etrangères, France. The acquisition, installation and maintenance of the Colonia Roma Dense Array was supported by CONACYT under project number 32588-T. The research presented here was part of that project. The critical reading of two anonymous reviewers greatly improved this manuscript.

REFERENCES

- Aki, K and P G Richards (1980), *Quantitative seismology, theory and methods*, W.H. Freeman and Company.
- Almendros González, F J (1999), “Análisis de señales sismo-volcánicas mediante técnicas de array”, *Ph.D. thesis*, Departamento de Física Teórica y del Cosmos de la Universidad de Granada, Spain.
- Barker, J S, M Campillo, F J Sánchez-Sesma, D Jongmans, and S K Singh (1996), “Analysis of wave propagation in the Valley of Mexico from a dense array of seismometers”, *Bulletin of the Seismological Society of America*, Vol. 86, pp. 1667-1680.
- Campillo, M, S K Singh, N Shapiro, J Pacheco, and R B Herrmann (1996), “Crustal structure south of the Mexican Volcanic Belt based on group velocity dispersion”, *Geofísica Internacional*, Vol. 35, pp. 361-370.
- Capon, J (1969), “High-resolution frequency-wavenumber spectrum analysis”, *Proceedings of the IEEE*, Vol. 57, pp. 1408-1418.
- Capon, J (1973), “Signal processing and frequency wavenumber spectrum analysis for a large aperture seismic array”, *Methods in Computational Physics*, Vol. 13, pp. 1-59.
- Cárdenas, M, F J Chávez-García, and A Gusev (1997), “Regional amplification of ground motion in central Mexico. Results from coda magnitude data and preliminary modelling”, *Journal of Seismology*, Vol. 1, pp. 341-355.
- Cárdenas, M, and F J Chávez-García (2003), “Regional path effects on seismic wave propagation in central Mexico”, *Bulletin of the Seismological Society of America*, Vol. 93, pp. 973-985.
- Chávez-García, F J and P Y Bard (1994), “Site effects in Mexico City eight years after the September 1985 Michoacan earthquakes”, *Soil Dynamics & Earthquake Engineering*, Vol. 13, pp. 229-247.
- Chávez-García, F J, J Ramos-Martínez, and E Romero-Jiménez (1995), “Surface wave dispersion analysis in Mexico City”, *Bulletin of the Seismological Society of America*, Vol. 85, pp. 1116-1126.
- Chávez-García, F J and E Romero-Jiménez (1998), “Ground motion in Mexico City: contribution of large-scale geologic structure”, *The effects of surface geology on seismic motion; Proceedings of the 2nd International Symposium. on the Effects of Surface Geology on Seismic Motion*, K. Irikura, K. Kudo, H. Okada, and T. Sasatani (editors.), 1-3 December, Yokohama, A.A. Balkema, Vol. 2, pp. 413-420.
- Chávez-García, F J and L Salazar (2002), “Strong motion in central Mexico: a model based on data analysis and simple modeling”, *Bulletin of the Seismological Society of America*, Vol. 92, pp. 3087-3101.
- Cornou, C (2002), “Traitement d’antenne et imagerie sismique dans l’agglomération grenobloise (Alpes françaises): implications pour les effets de site”, *Ph.D. thesis*, Observatoire de Grenoble et Laboratoire de Géophysique Interne et Tectonophysique, Université Joseph Fourier, France.

- Eissler, H, L Astiz, and H Kanamori (1986), "Tectonic setting and source parameters of September 19, 1985 Michoacan, Mexico earthquakes", *Geophysical Research Letters*, Vol. 13, pp. 569-572.
- Furumura, T and B L N Kennett (1998), "On the nature of regional seismic phases-III. The influence of crustal heterogeneity on the wavefield for subduction earthquakes: the 1985 Michoacán and 1995 Copala, Guerrero, Mexico earthquakes", *Geophysical Journal International*, Vol. 135, pp. 1060-1084.
- Goldstein, P and R J Archuleta (1987), "Array analysis of seismic signals", *Geophysical Research Letters*, Vol. 14, pp. 13-16.
- Goldstein, P and R J Archuleta (1991a), "Deterministic frequency-wavenumber methods and direct measurements of rupture propagation during earthquakes using a dense array: theory and methods", *Journal of Geophysical Research*, Vol. 96, pp. 6173-6185.
- Goldstein, P and R J Archuleta (1991b), "Deterministic frequency-wavenumber methods and direct measurements of rupture propagation during earthquakes using a dense array: data analysis", *Journal of Geophysical Research*, Vol. 96, pp. 6187-6198.
- Houston, H and H Kanamori (1986), "Source characteristics of the 1985 Michoacan, Mexico, earthquake at periods of 1 to 30 seconds", *Geophysical Research Letters*, Vol. 13, pp. 597-600.
- Iida, M (1999), "Excitation of high-frequency surface waves with long duration in the Valley of Mexico", *Journal of Geophysical Research*, Vol. 104, pp. 7329-7345.
- Marcos, S (1998), *Les méthodes à haute résolution: traitement d'antenne et analyse spectrale*, Editions Hermès, Paris, 783 pp.
- Ordaz, M and S K Singh (1992), "Source spectra and spectral attenuation of seismic waves from Mexican earthquakes, and evidence of amplification in the hill zone of Mexico City", *Bulletin of the Seismological Society of America*, Vol. 82, pp. 24-43.
- Schmidt, R O (1982), "A signal subspace approach to multiple emitter location and spectral estimation", *Ph.D. thesis*, Stanford University, Stanford, California.
- Schmidt, R O (1986), "Multiple Emitter Location and Signal Parameter Estimation", *IEEE Transactions on antennas and propagation*, Vol. AP-34, pp. 276-280.
- Singh, S K , E Mena, and R Castro (1988), "Some aspects of the source characteristics and ground motion amplifications in and near Mexico City from acceleration data of the September, 1985, Michoacan, Mexico earthquakes", *Bulletin of the Seismological Society of America*, Vol. 78, pp. 451-477.
- Singh, S K and M Ordaz (1993), "On the origin of long coda observed in the lake-bed strong motion records of Mexico City", *Bulletin of the Seismological Society of America*, Vol. 83, pp. 1298-1306.
- Wax, M and T Kailath (1985), "Detection of signals by Information Theoretic Criteria", *IEEE Transactions on acoustics, speech, and signal processing*, Vol. ASSP-33, pp. 387-392.