# Bi-dimensional Fourier transform with irregular spatial sampling 

D. Lugano ${ }^{1}$, A. Mazzotti ${ }^{1}$ and E. Stucchi ${ }^{2}$<br>${ }^{1}$ Earth Science Department, Geophysics, University of Pisa, Italy<br>${ }^{2}$ Earth Science Department, Geophysics, University of Milan, Italy

(Received: February 16, 2009; accepted: September 8, 2009)
ABSTRACT Seismic data acquisition is frequently carried out at irregular sampling intervals along spatial coordinates. This causes problems in the subsequent multi-trace data processing which often requires equi-spaced traces and thus the data must be first regularised. To this end, we illustrate a frequency domain method to transform 2D data irregularly sampled in the spatial direction to equivalent equi-spaced data. We follow a probabilistic inversion where the a-posteriori model is the desired (correct and noise free) frequency spectrum, the a-priori model is computed through the Non Uniform Discrete Fourier Transform (also known as Riemann sum) and the noise introduced by the irregular sampling is described empirically, on the basis of the distances between samples. All the variables are assumed to have Gaussian distributions and are described by their means and covariances. Once the optimum frequency spectrum is estimated, a Fourier anti-transform brings the data back into the time-space at constant spatial intervals. The proposed method is applied to synthetic and real seismic data, with various degrees of sampling irregularities and offset gaps, and with different noise contaminations and dips of events. The results are satisfactory and are improved with respect to those obtained by applying a previously developed method.

## 1. Introduction

Despite efforts to acquire seismic data through regularly sampled source and receiver positions, seismic processing, usually, has to be performed considering irregularly-sampled, spatial coordinates. This situation is not handled accurately by complex multichannel algorithms, such as pre-stack time or depth migration, since the data is by force interpolated on a regular grid before critical processing steps can be carried out. Even the simple and widely used Discrete Fourier Transform (DFT) suffers this limitation and leads to sub-optimal results when neglecting this issue. Besides the more or less complex interpolation schemes in the time domain (linear, spline, cubic), geophysical literature has addressed the problem of irregular sampling through different methods. These include, amongst others: the hyperbolic and linear Radon Transform (Thorson and Claerbout, 1985), the parabolic Radon Transform (Hampson, 1986; Kabir and Verschuur, 1995) and trace interpolation in the $f$ - $x$ domain (Spitz, 1991).

In this paper, we propose an extension of the method presented by Duijndam et al. (1999), which is based on the estimation in the Fourier domain of the best $f-k$ spectrum of the irregularly sampled seismic data along one spatial coordinate ( $x$ coordinate). This estimate is obtained by means of a probabilistic inversion approach (Tarantola, 2005) in the $f-x$ domain. We start
explaining the problems brought by irregular spatial sampling to the computation of the bidimensional frequency spectrum. Then, we describe the method for an optimum, noise-free, frequency spectrum estimation from which a regularly sampled time-space record can be derived.

We assume that the seismic data are band-limited and regularly sampled in the Fourier domain and the "noise" is represented by the frequency components that are out of the considered frequency bandwidth. Differently from the approach proposed by Duijndam et al. (1999), which assumes that there is no a-priori information available, we consider that the a-priori information can be derived from the Non Uniform Discrete Fourier Transform of the original data. Moreover, its appropriate manipulation can significantly improve the estimation of the optimum spectrum.

Finally, we demonstrate the applicability of our method on synthetic data where we have removed groups of traces to simulate gaps in the spatial direction and on real data with both missing traces and non-uniform spatial sampling.

## 2. Method

### 2.1. Description of the problem

The DFT computes the bi-dimensional spectrum of the data on regular grids in both the original and the transform domain. Eq. (1) describes the formula implemented in a typical DFT algorithm:

$$
\begin{equation*}
D(k, f)=\sum_{n=0}^{x-1} \sum_{n=0}^{T-1} d(n, m) e^{-j \frac{2 \pi l n}{x}} e^{-j \frac{2 \pi f n}{T}} \tag{1}
\end{equation*}
$$

where $d(n, m)$ are the input data as a function of space index $n$, and time index $m ; \mathbf{D}$ is the bidimensional spectrum as a function of the spatial $k$, and temporal $f$ frequencies; $X$ is the number of traces and $T$ is the number of temporal samples of the input data.

If the data are irregularly sampled, suppose along the $x$ direction only, the above formula is inaccurate for the computation of the theoretical spectrum; nevertheless, even considering the real spatial coordinates the problem will not be amended, as the following example demonstrates.

Fig. 1a shows a noise-free synthetic data set made of five linear dipping events. Each event is characterised by a Ricker wavelet with central frequency of 30 Hz . The data are irregularly sampled along the $x$ coordinates, while the time sampling rate is constant at 2 ms . We chose this data set because the harmonic components can be easily evaluated and the spectrum is well known.

If the $f$ - $k$ spectrum is computed with the Fourier transform given by Eq. (1), a blurring and smearing of the true spectral components along with the appearance of some noise contaminating the whole spectrum can be observed in Fig. 1b.

To get a result closer to the true one, one can make use of Eq. (2), which expresses the Fourier transform calculated in the true positions of the traces (Riemann sum) where the data has been 1D Fourier transformed from time to frequency by means of the classic DFT. Here $\mathbf{M}$ is the computed $f-k$ spectrum, $\mathbf{d}$ are the original data in the $f-x$ domain, $\Delta x_{s}$ is the distance between two


Fig. 1 - Input data in the $x-t$ domain (a), and its $f-k$ spectrum calculated with the FFT 2D (b).
traces, $X$ is the number of traces and $n$ is the frequency index that varies from $-N$ to $+N-1.2 \mathrm{~N}$ is thus the total number of frequencies we want to represent. Even in this case, the result is not satisfactory, as can be clearly observed in Fig. 2a where some blurring and noise are still evident.

$$
\begin{align*}
& n=-N, \ldots,+N-1=-\frac{X}{2}, \ldots,+\frac{X}{2}-1 \\
& M\left(k_{n}, f\right)=\sum_{s=0}^{X-1} d\left(x_{s}, f\right) e^{-j n \Delta x_{s}} \Delta x_{s} \Delta x_{s}=l_{s+1}-l_{s}=\frac{x_{s+1}-x_{s-1}}{2}  \tag{2}\\
& l_{s}=\frac{\left(x_{s}+x_{s+1}\right)}{2}
\end{align*}
$$

Thus, the poor spectrum estimation is related to the irregular spatial sampling we are using: the fact that we cannot define a Nyquist frequency and apply the sampling theorem as usual, causes a leakage of some spectral components that show up as alias noise in the spectrum and blur the true events.

A starting point, to describe the problem analytically, comes from the practical considerations that every real signal is band-limited and that the wavenumber separation is given by:

$$
\begin{equation*}
\Delta k<\frac{2 \pi}{L_{t o t}} \tag{3}
\end{equation*}
$$

where $L_{\text {tot }}=\sum_{s} \Delta x_{s}$ is the spread extension. On the basis of the above two assumptions, the frequency range that can be covered considering a constant $\Delta k$ step is related to the number of traces in the gather and to the effective spatial extension of the data (usually some percent greater than $L_{\text {tot }}$. The frequency components outside this bandwidth will contribute to the alias noise.


Fig. 2 - $f$ - $k$ spectrum calculated with the Riemann Sum (a) and same spectrum after filtering to remove noise (b).

### 2.2. Probabilistic approach to the $f$ - $k$ spectrum estimation

The $f$ - $k$ spectrum estimation, in case of irregular sampling, can be formulated as an inverse problem. Here, we review the basic steps of this methodology as explained by Duijndam et al. (1999), pointing out the main differences that characterise our approach. The original data $\mathbf{d}(x, t)$ can be related to the $f-k$ spectrum $\mathbf{m}(f, k)$ by the following linear relationship (Inverse Fourier Transform):

$$
\begin{equation*}
\mathbf{d}=\mathbf{G m} \tag{4}
\end{equation*}
$$

where $\mathbf{G}$ is the forward operator and $\mathbf{m}$ is the model we want to estimate by the inversion. Tarantola (2005) introduces an uncertainty factor $\boldsymbol{\varepsilon}$ to the data $\mathbf{d}$ affecting the observed measurements $\mathbf{d}_{\text {obs }}$ :

$$
\begin{equation*}
\mathbf{d}_{o b s}=\mathbf{d}+\boldsymbol{\varepsilon} \tag{5}
\end{equation*}
$$

In the case of seismic data irregularly sampled, in the place of the uncertainty factor $\boldsymbol{\varepsilon}$, we introduce an error term $\mathbf{n}$ representing the frequency components outside the bandwidth of the data. $\mathbf{n}$ can be considered the "noise of the model". The constant time sampling allows us to apply a standard Fourier transform along the time dimension and to conceive the irregular sampling problem as a one-dimensional problem in the $x$ direction. Eq. (4) can thus be written:

$$
\begin{equation*}
\mathbf{d}=\mathbf{G m}+\mathbf{n} \tag{6}
\end{equation*}
$$

where $\mathbf{d}(x, f)$ is a function of space and frequency, $\mathbf{G}$ is the forward operator used which will be described later and $\mathbf{m}$ and $\mathbf{n}$ are respectively the model and the model noise as previously stated.

The simplest way to take into account all the possible realizations of $\mathbf{m}$ and $\mathbf{n}$ is to consider
their variations as described by Gaussian distributions, namely:

$$
\begin{equation*}
\mathbf{m} \in \mathbf{N}\left(\mathbf{m}_{\text {prior }}, \mathbf{C}_{M}\right) \quad \mathbf{n} \in \mathbf{N}\left(\mathbf{0}, \mathbf{C}_{N}\right) \tag{7}
\end{equation*}
$$

where $\mathbf{m}_{\text {prior }}$ and $\mathbf{C}_{\boldsymbol{M}}$ are respectively the mean and the covariance matrix of the model and $\mathbf{C}_{N}$ is the covariance matrix of the noise $\mathbf{n}$. The probability density function with mean $\mathbf{m}_{\text {prior }}$ and covariance matrix $\mathbf{C}_{\boldsymbol{M}}$ in Eq. (7) expresses the a-priori information on the model.

By means of the inverse problem theory (Tarantola, 2005), we can compute the solution that realizes the minimum of the least-square misfit function between the original data and the reconstructed data in the $x$ - $f$ domain through Eq. (8):

$$
\begin{equation*}
\tilde{\mathbf{m}}=\left(\mathbf{G}^{\mathrm{t}} \mathbf{C}_{N}^{-1} \mathbf{G}+\mathbf{C}_{M}^{-1}\right)^{-1}\left(\mathbf{G}^{\mathrm{t}} \mathbf{C}_{N}^{-1} \mathbf{d}_{\text {obs }}+\mathbf{C}_{M}^{-1} \mathbf{m}_{\text {prior }}\right) \tag{8}
\end{equation*}
$$

where $\mathbf{G}^{t}$ is the transposed complex-conjugate of the forward operator $\mathbf{G}$ and $\mathbf{d}_{\text {obs }}$ is the input data $\mathbf{d}(x, f)$. Thus $\tilde{\mathbf{m}}$ is the optimum estimate of the bi-dimensional Fourier spectrum of $\mathbf{d}$.

The main difference between Duijndam's approach and ours is the way the a- priori information is considered. Duijndam et al. (1999) states that there is no a-priori model for seismic data. Instead, we think that since the Riemann sum, $\mathbf{M}\left(k_{w} f\right)$ given by Eq. (2), does provide an approximate, although noisy, spectrum of the data, it can be used to derive the model a-priori information described by $\mathbf{m}_{\text {prior }}$ and $\mathbf{C}_{\boldsymbol{M}}$ in Eq. (8). Mathematically, $\mathbf{m}_{\text {prior }}$ corresponds to the mean value of $\mathbf{M}\left(k_{n} f\right)$ over all the temporal frequencies $F$ as a function of $k_{n}$ :

$$
\begin{equation*}
\mathbf{m}_{\text {prior }_{n}}=\frac{1}{F} \sum_{p=0}^{F-1} \mathbf{M}\left(k_{n}, f_{p}\right) \tag{9}
\end{equation*}
$$

while the matrix $\mathbf{C}_{M}$ is expressed simply as:

$$
\begin{equation*}
\mathbf{C}_{M_{n}}=\sigma_{k_{n}}^{2} \mathbf{I} \tag{10}
\end{equation*}
$$

where $\sigma_{k_{n}}^{2}$ is the variance of the spatial frequency $k_{n}$, and $\mathbf{I}$ is the identity matrix. However, due to the absence of a DC component in the seismic data, the contribution of $\mathbf{m}_{\text {prior }}$ is negligible and all the $a$-priori information is brought in by the covariance matrix.

To formulate Eq. (10), we have supposed that the spatial frequencies are uncorrelated. In this case, $\mathbf{C}_{\boldsymbol{M}}$ becomes a diagonal matrix where the elements are the variances of each spatial frequency of the Riemann sum.

In the forward problem described by Eq. (6), the operator $\mathbf{G}$ allows us to pass from the $f-k$
domain of the model $\mathbf{m}$ to the $f-x$ domain of the data $\mathbf{d}$. $\mathbf{G}$ is defined as follows:

$$
G_{n s}=\frac{\Delta k}{2 \pi} e^{j n \Delta x_{s}} \quad \begin{align*}
& s=0, \ldots, X-1  \tag{11}\\
& n=-N, \ldots,+N-1=-\frac{X}{2}, \ldots,+\frac{X}{2}-1
\end{align*}
$$

where $\Delta k$ is the sampling rate for the spatial frequency $k$ in the Fourier domain, and $x_{s}$ represents the true positions of the samples; as an example for the seismic data, we can use the source to receiver distance (offset) as $x_{s}$. We call $\mathbf{G}$ the exponential matrix operator.

Using this approach, we consider a bandwidth limited from $-N \Delta k$ to $(+N-I) \Delta k$. The choice of $\Delta k$ [see Eq. (3)] is very important and is a trade off between two opposite conditions. On the one hand, if $\Delta k$ is too small, for a given $N$, we decrease the bandwidth. On the other hand, if $\Delta k$ is too large, we risk introducing the aliasing artefacts associated to the largest trace intervals.

Tests based on synthetic data, show that optimal results are achieved by increasing the spread length used for determining $\Delta k$ by a value of around $30 \%$ [Eq. (12)], thus considering a sampling rate in the Fourier domain somewhat smaller than the one used for regular sampling :

$$
\begin{equation*}
\Delta k \approx \frac{2 \pi}{1.3 L_{\text {tot }}} \tag{12}
\end{equation*}
$$

Note that to preserve the complex-conjugate symmetry of the Fourier Transform, the exponential matrix operator is computed for $n$ by varying between $-N$ to $+N-1$, differently from the DFT, where it usually varies from 0 to $2 N-1$.

Finally, detailed information is not available for a quantitative description of the noise term. As suggested by Duijndam et al. (1999), a practical approximation can be taken replacing the covariance matrix $\mathbf{C}_{N}$, by a diagonal matrix of weights $\mathbf{W}^{-1}$. The diagonal elements of $\mathbf{W}$ are expressed by the difference $\left(l_{s-1}-l_{s}\right)$ where $l_{s}=\left(x_{s+1}-x_{s-1}\right) / 2$ :

$$
\begin{equation*}
\mathbf{C}_{N}^{-1}=\mathbf{W}=\mathbf{I} \Delta x_{s}=\mathbf{I}\left(l_{s+1}-l_{s}\right)=\mathbf{I}\left(\frac{x_{s+1}-x_{s-1}}{2}\right), \tag{13}
\end{equation*}
$$

weighting the widely spaced samples more than those densely spaced in the misfit function.

## 2.3. f-k spectrum estimate

We can now evaluate $\widetilde{\mathbf{m}}$, the optimum estimate of the bi-dimensional Fourier spectrum of the input data, by introducing the terms previously described in Eqs. (9) and (10):

$$
\begin{equation*}
\tilde{\mathbf{m}}=\left(\mathbf{G}^{\mathbf{t}} \mathbf{W} \mathbf{G}+\mathbf{C}_{M}^{-1}\right)^{-1}\left(\mathbf{G}^{\mathbf{t}} \mathbf{W} \mathbf{d}_{\text {obs }}+\mathbf{C}_{M}^{-1} \mathbf{m}_{\text {prior }}\right) \tag{14}
\end{equation*}
$$

where, to summarize, $\mathbf{G}$ and $\mathbf{G}^{\mathbf{t}}$ are, respectively, the exponential matrix operator [Eq. (11)] and its transposed complex-conjugate, $\mathbf{W}$ is the weighting matrix [Eq. (13)], $\mathbf{C}_{M}$ is the covariance matrix [Eq. (10)] of the model a-priori information (Riemann sum), $\mathbf{d}_{\text {obs }}$ is the original data in


Fig. 3-f-k spectrum calculated with our method (a) and with the method proposed by Duijndam et al. (1999) (b) .
the $f$ - $x$ domain irregularly sampled and $\mathbf{m}_{\text {prior }}$ is the mean of the Riemann sum [Eq. (9)].
A few words can be said about the right hand side of Eq. (14). The term $\mathbf{G}^{\mathbf{t}} \mathbf{W} \mathbf{d}_{\text {obs }}$ of the second factor corresponds to the Riemann sum scaled by $\Delta k / 2 \pi$ [see Eqs. (2) and (15)]:

$$
\begin{equation*}
\left[\mathbf{G}^{\mathbf{t}} \mathbf{W} \mathbf{d}_{o b s}\right]=\frac{\Delta k}{2 \pi} \sum_{s=0}^{X-1} d\left(x_{s}, f\right) e^{-j n \Delta x_{s}} \Delta x_{s} . \tag{15}
\end{equation*}
$$

This term is added to $\mathbf{C}_{\mathbf{M}}{ }^{-1} \mathbf{m}_{\text {prior }}$, a factor that, as previously discussed, turns out to be negligible. The overall effect of the whole procedure is thus embedded in the first factor $\left(\mathbf{G}^{t} \mathbf{W G}+\mathbf{C}_{M}^{-1}\right)^{-1}$, which does the task of improving, in the least squares sense, the raw estimate of Eq. (15). It depends on the diagonal covariance matrix $\mathbf{C}_{M}^{-1}=\sigma_{k_{n}}^{2} \mathbf{I}$, whose spatial frequency variances $\sigma_{k_{n}}^{2}$ can be evaluated through the Riemann sum in Eq. (2). Better, $\mathbf{C}_{\boldsymbol{M}}{ }^{-1}$ can be computed on a Riemann sum spectrum where the alias noise, due to irregular sampling, has been attenuated by means of an appositely designed $f-k$ filter (Fig. 2b). The more the noise is reduced by the filter, the more the $\sigma_{k_{n}}^{2}$ variances are related to the dispersion of the signal around each spatial frequency $k_{n}$, allowing a more efficient "deconvolution" expressed by the probabilistic generalized inverse $\left(\mathbf{G}^{\mathbf{t}} \mathbf{W} \mathbf{G}+\mathbf{C}_{\boldsymbol{M}} \mathbf{}^{-1}\right)^{-1}$. Fig. 3a shows the optimum spectrum $\widetilde{\mathbf{m}}$ estimated by the Riemann sum of Fig. 2a with the covariance matrix computed on its filtered version (Fig. 2b). Clearly, the five events of the spectrum appear now less contaminated by the noise, enabling a more accurate data reconstruction and regularization.

Differently from our approach, Duijndam et al. (1999) do not consider any a-priori information and set the inverse of the covariance matrix $\mathbf{C}_{M}{ }^{-1}$ equal to the identity matrix I scaled by a stabilisation factor $k^{2}$ :

$$
\mathbf{C}_{M}^{-1}=k^{2} \mathbf{I} .
$$

No discrimination among the spatial frequencies is made by the covariance matrix. The $k$


Fig. 4 - Reconstructed signal (black traces) on a regular grid with our method (a); difference between the reconstructed signal and the regularly sampled original signal (b).
factor is dependent by ratio between the energies of the noise and the signal in the input data and is determined by trial and error.

Fig. 3b shows the result of the Duijndam et al. (1999) approach for the spectrum estimation of the data in Fig. 1a. In this case $\widetilde{\mathbf{m}}_{\text {Duij }}$ is given by:

$$
\begin{equation*}
\tilde{\mathbf{m}}_{\text {Duij }}=\left(\mathbf{G}^{\mathbf{t}} \mathbf{W} \mathbf{G}+k^{2} \mathbf{I}\right)^{-1}\left(\mathbf{G}^{\mathbf{t}} \mathbf{W} \mathbf{d}_{o b s}\right) . \tag{16}
\end{equation*}
$$

Comparing the spectra in Figs. 3 a and 3 b it can be observed that the introduction of the $\sigma_{k_{n}}^{2}$ variances of the filtered Riemann sum in the covariance matrix $\mathbf{C}_{M}$, has improved the spectrum estimation by reducing the dispersed alias noise and by enhancing the energy content of the five events.


Fig. 5 - Reconstructed signal (black traces) on a regular grid with the method proposed by Duijndam et al. (1999) (a); difference between the reconstructed signal and the regularly sampled original signal (b).


Fig. 6 - Zoom of the overlap of the input data (red traces) and the signal reconstructed with our methodology (black traces (a); difference between the reconstructed signal and the regularly sampled original signal (b).

To return to the $x$ - $t$ domain the Inverse Discrete Fourier Transform of Eq. (17) is applied where the exponential matrix $\mathbf{G}_{\text {reg }}$ at regular spatial sampling $\Delta x_{\text {new }}$ is computed by :

$$
G_{r e g_{p n}}=\frac{\Delta k}{2 \pi} e^{j m \Delta k\left(p \Delta n_{n e x}\right)} \quad \begin{align*}
& p=0, \ldots, P-1  \tag{17}\\
& n=-N, \ldots,+N-1=-\frac{X}{2}, \ldots,+\frac{X}{2}-1
\end{align*}
$$

where $P$ is the number of traces we want to reproduce and $\Delta x_{\text {new }}$ is their constant spatial sampling. The $f$ - $x$ regularized data are given by:

$$
\mathbf{d}_{r e g}=\mathbf{G}_{r e g} \tilde{\mathbf{m}}
$$

from which, with a simple 1D Inverse Fourier Transform along the frequencies, we come back into time.

Fig. 4a shows, in red, the synthetic model of Fig. 1a and in black the reconstructed traces that were missing, while Fig. 4b displays the differences between the original, regularly sampled model and the reconstructed one. From the analysis of these pictures, we can observe that a satisfactory reconstruction and regularization of the data is achieved, with errors varying from a maximum of $8 \%$ to about $2 \%$. The corresponding results obtained with the Duijndam et al. (1999) method show a somewhat inferior performance, as illustrated in Fig. 5, for the data and the differences, respectively.

## 3. Application and results

We have carried out many tests on synthetic and real data to check the limits of the proposed method (how many missing traces can be reconstructed) and to verify its applicability when noise is contaminating the data.


Fig. 7 - Zoom of the overlap of the input data (red traces) and the signal reconstructed with the Duijndam et al. (1999) methodology (black traces) (a); difference between the reconstructed signal and the regularly sampled original signal (b).

Using synthetic data, we removed groups of traces and we checked how the reconstructed traces would match the original ones. We started removing groups of 3 adjacent traces and we ended up removing groups of 8 adjacent traces from a total of 83 synthetic traces. Here, we illustrate the results of only one of these tests (Figs. 6 and 7). This noise-free synthetic data is made up of 83 traces with a spatial sampling of 1 m and shows 5 seismic events at different dips. Four gaps of five adjacent traces have been produced to create the input data for the reconstruction algorithms (see Figs. 6a and 7a, red traces).

Figs. 6a and 6 b show the results obtained by applying our method to the same synthetic test data. Now, the missing traces have been properly reproduced (black traces in Fig. 6a), matching both the amplitudes and dips of the original signals.

The black traces in Fig. 7a, represent the results obtained by applying the previous method of Duijndam et al. (1999). Differences with respect to the original traces are plotted in Fig. 7b to


Fig. 8 - Real data (common receiver gather) used as input data for testing the different methods.


Fig. 9 - Zoom of the overlap of the input data (red traces) and the signal reconstructed with our methodology (black traces) (a); zoom of the overlap of the input data (red traces) and the signal reconstructed with the Duijndam et al. (1999) methodology (black traces) (b).
show the errors in the reconstruction. It can be seen that although the gaps have been somehow filled, there are significant errors in the amplitude of the events and some inconsistency in their dips.

This is clearly demonstrated in Fig. 6b where the errors are much lower than in the previous case.

After many applications on synthetics, real data gathers (common shots, common receivers and CMPs) were used as input for the reconstruction algorithms. Since the same conclusions can be reached in any of the tests carried out, we show here, the results obtained for a common receiver gather characterised by large offsets, irregular spatial sampling and large gaps of missing traces (Fig. 8). This is one of the most difficult cases we experimented.

Figs. 9 a and 9 b show close-ups of the data in the region of the largest gap around a -4100 m offset, where six traces are missing. The original data are represented in red while the reconstructed traces are in black.

Fig. 9a shows our results and Fig. 9b illustrates results when applying the method proposed by Duijndam et al. (1999). While the dip of the events in the smaller gaps are reproduced rather well by both methods, based on the correlation with the existing adjacent traces, it seems that our method reproduces the missing signal amplitudes better. The reconstruction of the six missing traces forming the largest gap is difficult: the results obtained with the Duijndam et al. (1999) approach are completely unsatisfactory both in terms of dips and amplitudes of the reconstructed signals. Our method is able to somehow reproduce the wavelets at higher amplitude (at $1.9-2.0$ s) but it fails to reconstruct the deeper events (see around 2.2 s ).

## 4. Conclusions

We have discussed the problem of irregularly-spaced 2D seismic data and we have proposed a frequency domain method that is able to reconstruct the missing traces and to yield an evenly spaced data set. This method is an extension of a previously developed approach presented by

Duijndam et al. (1999).
Assuming a limited frequency band of the input data and a Gaussian distribution of all the variables (data, model and noise), we have used a probabilistic approach to estimate the optimum $f$ - $k$ spectrum where the noise components outside the signal bandwidth are removed and the signal is amplified.

Differently from the Duijndam et al. (1999) approach, which assumes that no a-priori information is available and in its place it introduces a constant factor (the noise/signal energy ratio) of difficult estimation, we consider the Riemann sum of the input data as very useful $a$ priori information. In fact, it can be conveniently filtered to emphasize the signal content, significantly improving the outcome of the inversion.

Many tests have been carried out on synthetic and on real data to verify the limits of our method and to verify the improvements brought by including a-priori information in the inversion. What we have shown here is limited to only one synthetic and one real data example but the results we achieved are quite representative of the whole set of tests. The method we propose always produced improved results when compared with the previous one. In the synthetic example shown, out of a total of 83 traces with a nominal spatial sampling of 1 m , we were able to reconstruct five gaps of five missing traces each.

The application to real data presents more difficulties due to the presence of noise which is additional to that due to irregular sampling. The test we showed on the common receiver is very difficult because there are many missing traces and large gaps. However, our results are satisfactory for most of the situations. Only for the largest gap of six traces is our method not able to properly reconstruct all the dipping events and particularly those at lower amplitudes.

Besides the extension to 3D, a future development could be the inclusion, in the inversion, of the full covariance matrix of the a-priori information instead of only the diagonal one. In our view, this should reinforce the information on the dip of the events present in the Riemann sum.

## REFERENCES

Duijndam A.J.W., Schonewille M.A. and Hindriks C.O.H.; 1999: Reconstruction of band-limited signals irregularly sampled along one spatial direction. Geophysics, 64, 524-538.
Hampson D.; 1986: Inverse velocity stacking for multiple elimination. J. Can. Soc. Expl. Geophysics, 22, 44-55.
Kabir M.M.N. and Verschuur D.J.; 1995: Restoration of missing offsets by parabolic radon transform. Geophysics, 43, 347-368.

Spitz S.; 1991: Seismic trace interpolation in the fx domain. Geophysics, 56, 785-794.
Tarantola A.; 2005: Inverse problem theory and methods for parameter estimation. Institut de Physique du Globe de Paris, Université de Paris 6, Siam.

Thorson J.R. and Claerbout J.F.; 1985: Velocity-stack and slant-stack stochastic inversion. Geophysics, 50, 2727-2741.

| Corresponding author: | Diletta Lugano |
| :--- | :--- |
|  | CGGVeritas SERvices (UK) LImited |
|  | Crawley RH10 9QN, United Kingdom |
| phone: +4401293 683894; e-mail: diletta.lugano@cggvertisas.com |  |

