1 **MULTISCALE MATCHED-FIELD PROCESSING FOR NOISE-SOURCE**

2 **LOCALIZATION IN EXPLORATION GEOPHYSICS.**

3

4 Margherita Corciulo[1], Philippe Roux [1], Michel Campillo [1], Dominique Dubucq[2] and

5 W. A. Kuperman[3]

6

7 **Right Running Head:** Multi-scale MFP for noise-source localization

8

9 [1] Institut des Sciences de la Terre, Université Grenoble 1, CNRS UMR 5275,

10 Grenoble, France.

11 E-mail: margherita.corciulo@obs.ujf-grenoble.fr; philippe.roux@obs.ujf-grenoble.fr;

12 michel.campillo@obs.ujf-grenoble.fr.

13 [2] TOTAL, Centre Scientifique et Technique Jean-Féger, Pau, France.

14 E-mail: dominique.dubucq@total.com.

15 [3] Marine Physical Laboratory, Scripps Institute of Oceanography, University of

16 California San Diego, La Jolla, USA.

17 E-mail: wkuperman@ucsd.edu.

18

19 Date of submission: 1 November 2011

20

21

1 **Abstract**

2

3       In the study we investigate the use of ambient-noise data to locate microseismic

4 sources at the exploration scale. We develop a multiscale matched-field processing

5 (MFP) approach to localize seismic sources at frequencies below 10 Hz. An application

6 to an actual dataset acquired over a hydrocarbon field is presented to determine the

7 reliability of the MFP procedure. The data used were continuously recorded over five

8 days, at a total of 397 stations on a 1-km-per-side square seismic network. The MFP

9 results show: (1) a dominant and stable surface source associated with human activities

10 (road and exploration platforms) around the reservoir; and (2) weaker sources at depth

11 below the seismic network that are related to the injection/ extraction process.

12

## 1 Introduction

2

3    In the seismic exploration context, the seismic-noise field at frequencies between

4    1 Hz and 10 Hz mainly relates to human activities. Hydrocarbon extraction has been

5    invoked to explain changes in stress, pore pressures and volumes of hydrocarbon

6    reservoirs that result in microseismic events. The source positions of these events can be

7    useful to investigate the evolution of the hydrocarbon reservoirs (McGarr and Simpson,

8    1997; Maxwell and Urbanic, 2001; Phillips et al., 2002). A central issue, therefore, is

9    the masking of these microseismic events by louder, anthropogenic sources of noise.

10   The method proposed here uses the coherence of the loud anthropogenic noise over a

11   large array for noise cancellation processing on smaller sub-arrays, thereby enhancing

12   the signal-to-noise ratio of the weaker microseismic events.

13   Classical approaches to locate seismic sources in geophysics exploration are

14   based on (micro)earthquake triangulation that are applied to earthquake-type events

15   using ballistic arrivals. However, microseismic events induced by hydrocarbon

16   exploitation (as hydraulic fracturation, for example) are observed on sensors in

17   boreholes but cannot be detected when data are acquired by surface array since ballistic

18   events are then very weak and covered by surface noise.

19   In this surface array configuration, different methods had been proposed to

20   locate sources of long-period events in volcanic areas and further applied to

21   microseismic events in hydrocarbon fields. Almendros et al. (2001) proposed a

22   probabilistic method to locate long-period events with no evident ballistic arrivals

23   through comparison of data and model-derived slowness vectors. Other studies have

24   proposed the use of the amplitude decrease as a function of the distance from the source,

25   after correction for site effects (Aki and Ferrazzini, 2000; Battaglia and Aki, 2003). This

1  last method fails when the amplitude decay function is more complex than a linear

2  decay.

3    Among other localization techniques, time-reversal is based on the reverse

4  propagation of the signals from the receivers into the medium, where they interfere

5  constructively for optimal refocusing at the original source location. Time-reversal

6  modeling is strictly dependent on the array geometry. In general, an array with a high

7  density of seismic stations with a wide aperture is needed for accurate wave-field

8  sampling that avoids artifacts in the final source localization. O'Brien et al. (2011)

9  compared the time-reversal approach with crosscorrelation and full-waveform moment

10 tensor methods to locate long-period events; they showed that all of these methods give

11 comparable results in terms of source position.

12   Time-reversal modeling was also proposed at the exploration scale by Steiner et

13 al. (2008), to locate microseismic events associated with a hydrocarbon reservoir in

14 Austria. Data recorded by densely sampled surface arrays on the Valhall oil field was

15 processed by Chambers et al. (2010), using a ray-based diffraction stack method. This

16 migration-style technique demonstrated that microseismicity can be located even when

17 several events overlap although this requires an accurate *a-priori* velocity model.

18   In recent years, continuous passive monitoring has been investigated as a tool to

19 study the reservoir extraction process. Several techniques have been proposed for the

20 localization of microseismic events. Most of these are based on analysis of continuous

21 data acquired in boreholes, as the signal amplitude at depth is higher than at the surface.

22 Verdon et al. (2010) proposed a technique that matched the observed P arrivals and S

23 arrivals by ray-tracing to localize microseismic events detected using a triggering mode.

24 When applied to a borehole dataset, this technique uses a 1D velocity model, and it fails

25 when the signal-to-noise ratio is low or when several events overlap.

1    Through the use of continuous noise recorded on a seismic array at the surface,

2    the methodology presented in this study is based on phase coherence estimation using

3    the so-called matched-field processing (MFP) technique. This technique was first

4    developed in underwater acoustics and seismology, and it is based on phase match

5    between the data and a model on a dense receiver network, for the production of high-

6    resolution source-position images. In recent years, the MFP technique has been used

7    successfully with seismic-noise data for the location of sources of hydrothermal activity

8    at frequencies above 10 Hz (Legaz et al., 2009; Vandemeulebrouck et al., 2010, Cros et

9    al., 2011).

10   However, recent frequency studies that were performed over volcanoes or

11   various oil and gas fields have shown that low-frequency anomalies below 10 Hz can

12   also be related to microseismicity (Dangel et al., 2003; Holzer et al., 2005; Lambert et

13   al., 2008; Steiner et al., 2008; Saenger et al., 2009; Ali et al., 2009). The development of

14   new methodologies to automatically localize seismic noise sources at low frequencies

15   from continuous recordings is of particular interest at the exploration scale, for an

16   understanding of the evolution of hydrocarbon reservoirs and to optimize the extraction

17   process.

18   In the following, a theoretical description of the Matched Field Processing and

19   Multi-Rate Adaptive Beamforming is presented. The data organization and data analysis

20   is then described for a five-day continuous recording performed on a dense network

21   deployed above an exploration reservoir. Dominant surface noise sources were first

22   located, mostly related to human activities, as well as secondary volume sources related

23   to the injection/extraction process at depth in the vicinity of the reservoir.

24

25   **The matched-field processing method**

1

2    The source-localization method proposed here was initially developed in ocean

3    acoustics to locate the depths and ranges of low-amplitude noise sources (Kuperman

4    and Turek, 1997) using an array-processing technique (Jensen et al., 2011).

5    Using an array of sensors, MFP is used to measure the phase coherence of

6    ambient noise data, to produce a map that shows the probable location of the dominant

7    noise source.

8    The MFP technique is the 3D generalization of the conventional plane-wave

9    beamformer, which was introduced by Bucker (1976) and Hinich (1979), and was

10   further discussed by Fizell (1987) and Baggeroer et al. (1988). The plane-wave

11   beamforming process consists of the summing of signals on a receiver array using

12   various phase shadings to isolate incident plane waves. Plane-wave beamforming

13   simply determines the azimuth direction of the noise source, whereas the 3D MFP

14   technique gives the spatial coordinates of a point source, which depend on the array

15   geometry and on the complexity of the medium.

16   The MFP technique is based on computation of a model-based synthetic field

17   (the replica vector) of a point source at each candidate point in a search grid of the

18   medium, which has its phase and/or amplitude matched to the data. The maximum of

19   the correlation, i.e. the maximum of the MFP output, is obtained when the surrogate and

20   actual point source co-localize (Kuperman and Turek, 1997).

21   MFP algorithms are array-processing techniques that are typically implemented

22   in the frequency domain. From the recorded time-domain signal $d_j(t)$, which is

23   associated with array element j (where j varies from 1 to N, the total number of

24   receivers), the cross-spectral density matrix (CSDM) $K$ is calculated at frequency $\omega$

25   from the Fourier transform $d(\omega) = [d_1(\omega),...,d_N(\omega)]$ as:

1

$$K(\omega) = \langle d(\omega)d^*(\omega) \rangle \qquad (1),$$

3

where * indicates the Hermitian transpose operation. The bracket in equation 1 refers to an ensemble averaging that is performed on different successive noise recordings, which allows a full-rank CSDM to be obtained.

The replica vector $d(\omega, a) = [d_1(\omega, a), ..., d_N(\omega, a)]$ is the model-based Green's function at frequency $\omega$ for a surrogate source position $a$ in the medium. The homogeneous velocity model is the simplest model for use in three dimensions when low frequencies are analyzed such that the large seismic wavelength $\lambda$ blurs out the medium spatial heterogeneity. In three dimensions, each element of the replica vector is defined for a homogeneous velocity model as:

13

$$d_j(\omega, a) = \frac{1}{4\pi a_j} \exp\left(\frac{i\omega a_j}{c}\right) \qquad (2),$$

15

where $c$ is the medium velocity and distance $a_j$ contains the distance between the array element j and the surrogate point source $a$. When seismic noise is dominated by surface waves, the 2D homogeneous velocity model is used to approximate the replica vector components as:

20

$$d_j(\omega, a) = \sqrt{\frac{2}{\pi a_j}} \exp(-i\pi/4) \exp\left(\frac{i\omega a_j}{c}\right) \qquad (3).$$

22

1 Further, more complex Green's functions can be used as replica vectors; e.g. in the case

2 of a layered medium.

3 The source location is retrieved by matching the CSDM matrix $K(\omega)$ against the

4 replica vector $\tilde{d}(\omega, a)$ using a linear processor, which is known as a Bartlett processor,

5 as defined for each surrogate point source as:

6

7
$$B_{Bart}(a) = \sum_{\omega} \left| \tilde{d}^*(\omega, a) K(\omega) \tilde{d}(\omega, a) \right| \qquad (4).$$

8

9 To avoid anomalous contributions due to spurious frequencies in the noise power

10 spectrum, $K(\omega)$ is traditionally normalized at each frequency. Thus, the final Bartlett

11 processor $B_{Bart}$ is averaged over a set of discrete frequencies to enhance the contrast of

12 the probability function. The range of frequencies used for this averaging must be

13 limited to a frequency band where $c$ is assumed to be constant. Physically speaking, the

14 Bartlett algorithm is equivalent to the crosscorrelation between the wave-field data and

15 the model. According to diffraction laws, the spatial resolution is thus limited by the

16 acoustic wavelength $\lambda$, which means that the optimal focal spot cannot be smaller than

17 $\lambda / 2$.

18 Adaptive MFP techniques can be used to obtain higher-resolution results. The

19 minimum variance distortionless response (MVDR) is one of the adaptive algorithms

20 that is based on a maximum likelihood match between the data and the model (Capon,

21 1969). The adaptive methods are high-resolution, non-linear techniques, the efficiency

22 of which depends on the signal-to-noise ratio in the data and the fidelity of the replica

23 model.

24 Compared to the Bartlett processor, the MVDR output is computed as follows:

1

$$B_{MV}(a) = \sum_{\omega} \left\| \frac{1}{\tilde{d}^*(\omega,a)K^{-1}(\omega)\tilde{d}(\omega,a)} \right\| \qquad (5),$$

3

where $K^{-1}$ is the inverse of the CSDM.

The choice between the linear Bartlett and the non-linear MVDR depends on the trade-off between robustness and spatial resolution. At low frequencies, where the spatial coherence and signal-to-noise ratio are generally high, MVDR will provide a high-resolution source localization that is not limited by the seismic wavelength. On the other hand, MVDR might not work any longer at higher frequencies, where the homogeneous velocity model is too simple to describe the medium complexity. In this case, the Bartlett may still provide satisfactory results.

The difficulty with all MFP algorithms is the need for *a-priori* knowledge of the velocity model, to generate the correct replica vectors. The non-linear MVDR is much more sensitive to environmental uncertainties than the Bartlett processor. When no *a-priori* information about the environment is available, the direct use of the MVDR can lead to incorrect source localization.

Similarly, if several noise sources overlap in time, with the linear Bartlett processor, there is superposition of the probability of the presence of the different sources, which leads to an overall ambiguity of the solution. Indeed, the background noise in the exploration context is often dominated by loud sources that are due to human activities (i.e. pumps or exploration platforms) and that mask the low-amplitude sources that are related to the reservoir itself. More precisely, the dominant sources mostly radiate surface waves that are recorded on the whole seismic network whereas the weaker sources are generated at depth (eventually close to the gas reservoir), and

1    suffer from a faster amplitude decay, which means that they might be recorded on a part

2    of the array only.

3        Such "loud source *versus* weak source" configurations are well treated in a

4    manner similar to that of multi-rate adaptive beamforming (MRABF), which was

5    developed for underwater acoustics by Cox (2000). MRABF originally dealt with the

6    different times scales that are associated with the accumulation of the CSDM for slow-

7    moving and fast-moving sources. However, for the case that we are considering here,

8    the different scales are associated with the spatial coherence of the different noise

9    sources at the array. MRABF localizes weak noise sources that are masked by strong

10   undesired sources through the use of a projection algorithm that is performed from the

11   eigenvector decomposition of the CSDM. In general, MRABF aims to: (1) cancel the

12   wave field of the strongest source measured on the whole network by applying an

13   orthogonal projection to the data matrix; and (2) accurately locate the weaker source on

14   the appropriate subset of sensors.

15       More precisely, in the first step of MRABF, the CSDM is decomposed into its

16   dominant eigen components through a Singular Value Decomposition (SVD) as:

17

18   $$K(\omega) = U\,SV^* = U_m\,\Lambda_m V_m^* + U_l\,\Lambda_l V_l^* \tag{6}.$$

19

20   The SVD algorithm classically projects the CSDM into a set of eigenvectors (matrix $U$

21   and $V$) and singular values (the elements of the diagonal matrix S).  From the SVD, $\Lambda_m$

22   indicates the diagonal matrix that contains the largest eigenvalue(s) and [$U_m$, $V_m$] are the

23   matrices of the corresponding eigenvector(s). Similarly, $\Lambda_l$ consists of the diagonal

24   matrix with all the remaining (lower) singular values and [$U_l$, $V_l$] are the corresponding

25   eigenvectors. The number of dominant eigen components selected in $U_m\,\Lambda_m V_m^*$ is

10

1    arbitrary and will be detailed later in the data analysis. From the decomposition in

2    equation 6, the matrices $[U_m, V_m]$ are used to project the largest eigenvectors out of the

3    data $d(\omega)$, i.e. to annul the dominant interferers. To this goal, the projector $I - U_m V_m^*$

4    is applied to the data $d(\omega)$ in order to built a "projected" data vector $z(\omega)$, as defined

5    in Cox (2000):

6

7
$$z(\omega) = \left[ I - U_m V_m^* \right] d(\omega) \tag{7},$$

8

9    where $I$ is the identity matrix. As $d(\omega)$ is the original data dominated by the loud

10   source, $z(\omega)$ is the "projected" data from which a "projected" CSDM matrix is defined

11   as:

12
$$P(\omega) = \left\langle \bar{z}(\omega) \bar{z}^*(\omega) \right\rangle \tag{8},$$

13

14   where $\bar{z}(\omega)$ refers to a subset of stations where the weak source at depth is eventually

15   recorded. The MVDR or Bartlett processor is then applied to $P(\omega)$, according to

16   equation 4 or 5, and using the 3D Green's function as the replica vector (equation 2) on

17   the same geophone subset.

18          Of course, the choice of the geophone subset must be made with great care to

19   allow for the detection of the weak source. The objective of MRABF is to use the whole

20   network to accurately identify the dominant "surface" source eigenvectors $[U_m, V_m]$ and

21   to cancel them on a sub-network that is likely to be sensitive to a weaker source at

22   depth.

23          As the main objective of the present study was to develop an innovative

24   methodology to localize noise sources in the context of exploration geophysics, a

11

1   multistep procedure is now introduced. This procedure begins in its early stages with the

2   Bartlett MFP (and the high-resolution MVDR, when appropriate) to locate the dominant

3   source. Then the algorithm benefits from the MRABF technique to iteratively annul the

4   dominant source and to localize the secondary noise sources in or around the reservoir.

5

6   **Data processing**

7

8        Data was acquired on an array with a 1-km-squared area with a total of 397

9   stations laid out in 13 lines, and with continuous 5-day acquisition. The distance

10  between lines was fixed at 60 m, while the inter-station distances varied from 10 m to

11  40 m, depending on the line (Figure 1).

12        The seismometers were 10-Hz geophones that were buried in 9-m-deep holes,

13  and the data were acquired using a 1000 Hz sampling rate. The array layout allowed for

14  highly redundant noise acquisition, while the sampling rate allowed for a wide range of

15  data frequency content. The data were originally gathered in 30-s time segments, as

16  shown in Figure 2a. The frequency analysis was performed on 2.5-min recordings that

17  allowed short to long wavelength analysis.

18        The time–frequency spectrogram averaged over the whole seismic network and

19  computed for a total of 18 h (Figure 2b) shows that the seismic noise was mostly

20  distributed below 20 Hz. Moreover, the time-averaged spectrogram (Figure 2c)

21  confirms a significant level of seismic intensity below 10 Hz confirming the sensitivity

22  of the geophones below their nominal frequency.

23        The MFP techniques require spatial coherence on a large set of seismic stations.

24  Crosscorrelation results performed for different frequency bands confirmed that a

25  stronger spatial coherence was retrieved at low frequencies (Figure 3a). Taking station

1  #185 as the reference station, spatial coherence is clearly visible from 2.5-min-long

2  recordings that were pre-filtered between 1.8 Hz and 3.5 Hz. Each crosscorrelation

3  function for any particular station is dominated by one peak, the arrival time of which

4  depends on the distance between the dominant noise source, the reference station #185,

5  and the station location. As the reference station chosen is located in the middle of the

6  array, high-amplitude coherence is noticeable for the whole seismic network. As

7  expected, the spatial coherence can be easily followed along the thirteen lines of the

8  seismic network.

9      Figure 3b shows the correlation functions stacked by 5-m intervals and sorted by

10 offset from the reference station instead of the station number. High-amplitude

11 coherence is still visible at offsets greater than 500 m. As classically observed in

12 seismic interferometry at both large and small scales [Sabra et al, 2005; Halliday et al.,

13 2008], this seismic-section display provides information on the nature of the continuous

14 noise-like wavefield illumination and on the wave velocity in the medium (Figure 3b,

15 red line). The mean apparent velocity retrieved was about 230 m/s, which suggests that

16 surface waves dominate the ambient noise recordings at low frequencies.

17     To investigate temporal changes in the spatial distribution of the noise sources,

18 the crosscorrelation function was calculated for one station pair along the five

19 acquisition days. Figure 4 shows the time-evolving correlation function for the station

20 pair #185 and #194, which are separated by a distance of 177 m and are located on array

21 line #7. Each correlation function was calculated from a 2.5-min time segment between

22 1.8 Hz and 3.5 Hz stacked over 1 hour. From the stability of the correlation function, it

23 can be deduced that the noise source remained stable over the five days of acquisition,

24 despite two anomalous time periods during days 1 and 3.

25

1    **Dominant noise-source localization**

2

3        As MFP is based on phase match between the data and a model on the geophone

4    array, the first step towards noise-source localization was to evaluate the spatial

5    coherence with respect to six reference stations in the network. For each reference, the

6    geophones with a coherence level above 0.7 were selected and used to perform MFP

7    localization. The six MFP outputs were then averaged to obtain the final localization

8    map of the dominant noise source.

9        With the seismic noise dominated by surface waves, the 2D grid search for MFP

10   was defined as a 5 km × 6.5 km grid, with a 100 m cell interval in both horizontal

11   directions. A medium velocity of 230 m/s was used to compute the replica vectors

12   according to equation 4. Figure 5a shows the probability of the presence of the noise

13   source obtained after averaging the Bartlett outputs, where the gray triangles represent

14   the geophones array, and the color scale represents the output amplitude of the Bartlett

15   MFP. By definition, the Bartlett output should be equal to one in the case of a perfect

16   match between the data and the replica vectors. The maximum value of ~0.7 observed

17   in Figure 5a corresponds then to a satisfactory match in phase and amplitude that

18   confirms the accuracy of the localization.

19       The dominant source clearly appears to be away from the array and the focal

20   spot is elongated in the North-East direction. The spatial dimension of the focus

21   retrieved with the Bartlett processor depends on the acoustic wavelength (~100 m at 2

22   Hz) and the angular aperture of the array viewed from the source location. To improve

23   spatial resolution, the MVDR was performed on a smaller 2.0 km × 2.0 km grid search

24   around this location, with a 100 m cell interval (Figure 5a, black dotted box). Figure 5b

25   shows the MVDR result in this expanded area. As expected, the MVDR shows higher

1    resolution for source localization. On the other hand, the MVDR algorithm being a non-

2    linear process requiring the inversion of the CSDM matrix, the MVDR output amplitude

3    does not carry any extra information compared to the Bartlett MFP.

4        Theoretical travel times associated with the noise-source location were

5    computed and superimposed on the crosscorrelation functions, using station #185 as

6    reference station (Figure 3a). In Figure 3a, the black dots appear to follow the phase of

7    the dominant wavefront, which confirms that MFP provides an excellent phase match

8    with the dominant noise source.

9        In a second step, the time evolution of the dominant noise-source location during

10   the five recording days was studied. A total of 2652 files of 2.5-min-long noise

11   recordings were processed, which provided six different MFP outputs that were

12   averaged out to define the source location for each time window. Finally, the retrieved

13   source locations are positioned on a geographic map (Figure 6). As expected, the

14   dominant noise sources are probably related to human activities. In particular, the north-

15   eastern source locations correspond to a straight road, while the southern source

16   locations are due to different human structures.

17

18   **Weaker noise-source location**

19

20       Figure 6 clearly shows that the dominant noise sources: (1) do not come from

21   the reservoir; and (2) might mask secondary sources that could be related to reservoir

22   activity (Draganov et al, 2010). To localize low-amplitude noise sources below the

23   geophone network, the MRABF method was used at higher frequencies, so in the 5 Hz

24   to 7 Hz frequency range.

1          As noise sources at depth suffer from a faster amplitude decay with respect to

2     distance than surface noise sources, they may only be recorded on a limited number of

3     geophones. A global search methodology was developed that applied the MRABF

4     algorithm to a set of sub-networks that covered the whole geophone array. Each sub-

5     network consisted of the set of geophones included in a 300 m × 300 m square area that

6     was progressively shifted by 100 m in the X and Y directions, to cover the entire

7     acquisition field. The lateral extension of each sub-network was chosen according to the

8     expected seismic wavelength (estimated at ~100 m). The 24 sub-networks were made

9     up of 65 to 90 geophones.

10          In the first step of the MRABF algorithm, the main eigenvector of the CSDM on

11     the whole network was calculated (equation 6). The orthogonal projector for the

12     dominant source was then defined in a second step for each sub-network, as given in

13     equation 7. An a priori model velocity of 600 m/s at the surface with a vertical gradient

14     of 1 m/s/m was used to compute the replica vectors, according to equation 3. Figure 7

15     shows the Bartlett output obtained after canceling the dominant source for one of the

16     sub-networks (Figure 7c, gray crosses). The focus is clearly visible underneath the sub-

17     network at a depth of 480 m, in the vicinity of one injection well (Figure 7c, central

18     black line). However, this source depth should not be trusted since diffraction physics

19     states that a surface array provides a good lateral resolution (in the X-Y plane, see

20     Figure 7c) but a poor axial resolution (along the Z axis, see Figures 7a, b and d).

21          This procedure is automatically repeated for the 2652 files of 2.5-min-long noise

22     recordings and for each sub-network. The noise-source location is recorded when the

23     Bartlett output shows a maximum value greater than a fixed threshold and

24     corresponding to a clear focus with no sidelobe in the X-Y plane. When this is not the

25     case, the orthogonal projector is calculated again to include the next loudest eigenvector

1    in the matrices ($U_m$, $V_m$) (see equation 7). The presence of the dominant source in more

2    than one eigenvector might be due to source motion or to any environment changes

3    during the 2.5-min-long time window. A maximum number of 10 eigenvectors were

4    then iteratively canceled, depending on the output of the Bartlett processor computed at

5    each step of this iterative procedure.

6          The X-Y projection of all of the secondary noise sources is finally shown in

7    Figure 8 as the local number of detection per hour for a 2D grid with a 50m x 50m unit

8    cell. These results show that low-amplitude noise sources originate from the injection

9    wells and the extraction process in the reservoir. When summed over the whole area, the

10    temporal evolution of the source localization process can finally be plotted (Figure 9a).

11    When compared to noise source detection at specific places (for example the tip of the

12    injection wells as circled in Figure 8), this reveals periods of calm and strong spatially

13    independent microseismic activities (Figures 9b and 9c), which should be related to the

14    injection/extraction process at the different well locations.

15          As shown by Saenger et al. (2009), seismic noise around 5 Hz appears to be

16    related to microseismic events located in the hydrocarbon reservoir. They suggested that

17    the source mechanism is related to poro-elastic effects due to injection or withdrawal of

18    fluids. Another source mechanism could be the bubble collapse associated to the fluid

19    injection at the tip of the wells, such bubbling noise being classically observed in a

20    broadband frequency range around 10Hz.

21          Unfortunately, no independent data that measured this activity were available for

22    comparison during the five days of recording.

23

24    **Conclusions**

25

1    In the present study, the use of ambient noise to locate seismic sources related to the

2    extraction process was investigated at the exploration scale. Taking advantage of the

3    density of seismic geophones deployed in the exploration context, the localization of

4    microseismic events was performed according to the MFP technique, based on the

5    spatial coherence between close sensors at low frequency (typically below 10 Hz).

6           As the seismic noise is dominated in this frequency band by human-generated

7    surface noise, the MFP and MRABF are applied in two steps: first, to detect and

8    localize the dominant surface source; and secondly, after canceling out the surface noise

9    source, to solely focus on the low-amplitude noise sources at depth around the injection

10   wells.

11          The robustness of this methodology is showed in Figure 8, where consistent

12   source locations were retrieved using different unrelated sub-networks. However, the

13   non-uniqueness of the problem associated with the medium complexity might have a

14   role in the ambiguity of the source localization at depth. In other words, performing a

15   non-ambiguous inversion for both source position and velocity model without any *a-*

16   *priori* information is a hard task at this scale of analysis.

17          Nevertheless, if some information regarding the velocity field in the medium is

18   available, the MFP algorithm has the advantage of an automatic procedure for source

19   localization with satisfactory resolution along the X-Y lateral dimensions of the seismic

20   array. Therefore, MFP techniques represent a noninvasive and rapid method to monitor

21   the reservoir seismic activity during the exploitation process.

22

1

## Acknowledgements

6

1

2 **References**

3

4 Aki, K., and V. Ferrazzini, 2000, Seismic monitoring and modelling of an active

5         volcano for prediction: Journal of Geophysical Research, **105**, 16617-16640.

6 Ali, M. Y, K. A. Berteussen, J. Small, B. Barkat and O. Pahlevi, 2009, Results from a

7         low frequency passive seismic experiment over an oilfield in Abu Dhabi: First

8         Break, **27**, 91-97.

9 Almendros, J., B. Chouet and P. Dawson, 2001, Spatial extent of a hydrothermal system

10         at Kilauea volcano, Hawaii, determined from array analyses of shallow long-

11         period seismicity: Journal of Geophysical Research, **106**, 13565-13580.

12 Baggeroer, A. B., W. A. Kuperman and H. Schmidt, 1988, Matched field processing:

13         source localization in correlated noise as an optimum parameter estimation

14         problem: Journal of Acoustical Society of America, **83**, 571-587.

15 Battaglia, J., and K. Aki, 2003, Location of seismic events and eruptive fissures on the

16         Piton de la Fournaise volcano using seismic amplitudes: Journal of Geophysical

17         Research, **108**, 2364, doi:10.1029/2002JB002193.

18 Bucker, H. P., 1976, Use of calculated sound fields and matched-field detection to

19         locate sound sources in shallow water: Journal of Acoustical Society of

20         America: **59**, 368-373, doi:101121./1.380872.

21 Capon, J., 1969, High resolution frequency-wavenumber spectrum analysis: Proceeding

22         IEEE, **57**, 1408-1418.

23 Chambers, K., J. M. Kendall and O. Barkved, 2010, Investigation of induced

24         microseismicity at Valhall using the Life of Field Seismic array: The Leading

25         Edge, **29**, 290-295.

Cox, H., 2000, Multi-rate adaptive beamforming (MRABF): *Sensor Array and Multichannel Signal Processing Workshop*, Proceeding of the 2000 IEEE, 306-309.

Cros, E., P. Roux, J. Vandemeulebrouck and S. Kedar, 2011, Locating hydrothermal acoustic sources at Old Faithful Geyser using Matched Field Processing: Geophysical Journal International, **187**, 385-393. doi:10.1111/j.1365-246X.2011.05147.x.

Dangel, S., M. E. Schaepman, E. P. Stoll, R. Carniel, O. Barzandji, E. D. Rode, and J. M. Singer, 2003, Phenomenology of tremor-like signals observed over hydrocarbon reservoirs: Journal of Volcanology and Geothermal Research, **128**, 135-158.

Draganov, D., X. Campman, J. Thorbecke, A. Verdel, and K. Wapenaar, 2010, Event-Driven Seismic Interferometry With Ambient Seismic Noise: 72nd EAGE Conference and Exhibition, Barcelona, Spain, G013.

Fizell, R. G., 1987, Application of high-resolution processing to range and depth estimation using ambiguity function methods: Journal of Acoustical Society of America, **82**, 606-613, doi:10.1121/1.395462.

Halliday, D., A. Curtis, and E. Kragh, 2008, Seismic surface waves in a suburban environment: active and passive interferometric methods: The Leading Edge, **27,** 210-218, doi:10.1190/1.2840369

Hinich, M. J., 1979, Maximum likelihood estimation of the position of a radiating source in a waveguide: Journal of Acoustical Society of America, **66**, 480-483, doi:10.1121/1.383099.

Holzner, R., P. Eschle, H. Zücher, M. Lambert, R. Graf, S. Dangel and P. F. Meier, 2005: Applying microtremor analysis to identify hydrocarbon reservoir: First Break, **23**, 41-46.

Jensen, F., W. A. Kuperman, M. Porter and H. Schmidt, 2011, Computational Ocean Acoustic (2$^{nd}$ Edition): Springer-Verlag New York Inc.

Kuperman, W. A. and G. Turek, 1997, Matched field acoustics: Mechanical Systems and Signal Processing, **11**, 141-148, doi:10.1006/mssp.1996.0066.

Lambert, M. A., S. M. Schmalholz, E. H. Saeger and B. Steiner, 2008, Low-frequency microtremor anomalies at an oil and gas field in Voitsdorf, Austria: Geophysical Prospecting, doi:10.1111/j.13652478.2008.

Legaz, A, A. Revil, P. Roux, J. Vandemeulebrouck, P. Gouédard, T. Hurst and A. Bolève, 2009, Self-potential and passive seismic monitoring of hydrothermal activity: A case study at Iodine Pool, Waimangu geothermal alley, New Zealand: Journal of Volcanology and Geothermal Research, **179**, 11-18.

Maxwell, S. C. and T. I. Urbanicic, 2001, The role of passive microseismic monitoring in the instrumented oil field: The Leading Edge, **20**, 636-639.

McGarr, A. and D. Simpson, 1997, Keynote Lecture: A broad look at induced and triggered seismicity, rockbursts and seismicity in mines: Gibowicz & Lasocki, Balkema, Rotterdam, 437pp.

O'Brien, G. S., I. Lokmer, L. De Barros, C. J. Bean, G. Saccarotti, J. P. Metaxian and D. Patane, 2011, Time reverse location of seismic long-period events recorded on Mt Etna: Geophysical Journal International, **184**, 452-462.

Phillips, W. S., J. T. Rutledge, L. S. House and M. C. Fehler, 2002, Induced microearthquake patterns in hydrocarbon and geothermal reservoirs: six case studies: Pure and Applied Geophysics, **159**, 345-369.

1    Sabra K. G., P. Gerstoft, P. Roux, W.A. Kuperman and M. Fehler, 2005, Extracting

2           time domain Green's function estimates from ambient seismic noise:

3           Geophysical Research Letters, **32,** L03310.

4

5    Saenger, E. H., S. M. Schmalholz, M. A. Lambert, T. T. Nguyen, A. Torres, S. Metzger,

6           R. M. Habiger, T. Müller, S. Rentsch and E. Méndez-Hernandez, 2009, A

7           passive seismic survey over a gas field: Analysis of low-frequency anomalies:

8           Geophysics, **74**, 029-040.

9    Steiner, B., E. H. Saenger and S. M. Schmalholz, 2008, Time reverse modelling of low-

10          frequency microtremor: Application to hydrocarbon reservoir localisation:

11          Geophysical Research Letters, **35**, L03307, doi:10.1029/2007GL032097.

12    Verdon, J. P., J. M. Kendall, D. J. White, D. A. Angus, Q. J. Fisher and T. Urbancic,

13          2010, Passive seismic monitoring of carbon dioxide storage at Weyburn: The

14          Leading Edge, **29**, 200-206.

15    Vandemeulebrouck, J., P. Roux, P. Gouédard, A. Legaz, A. Revil, T. Hurst, A. Bolève

16          and A. Jardani, 2010, Application of acoustic noise and self-potential

17          localization techniques to a buried hydrothermal vent (Waimangu Old Geyser

18          site, New Zealand): Geophysical Journal International, **180**, 883-890,

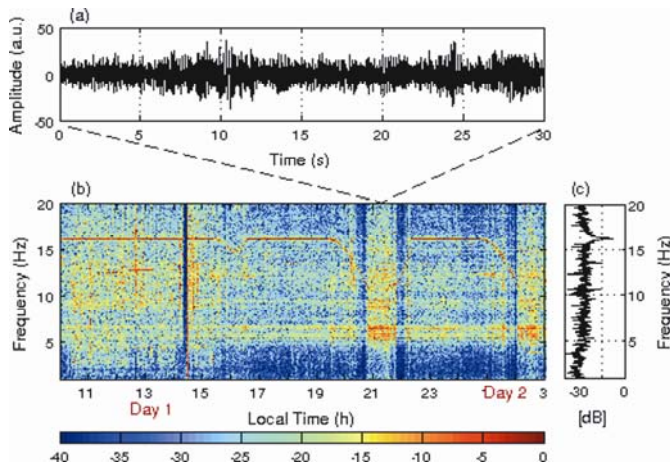19          doi:10.1111/j.1365.246X.2009.04454.x.

20

1 **Figure**

2
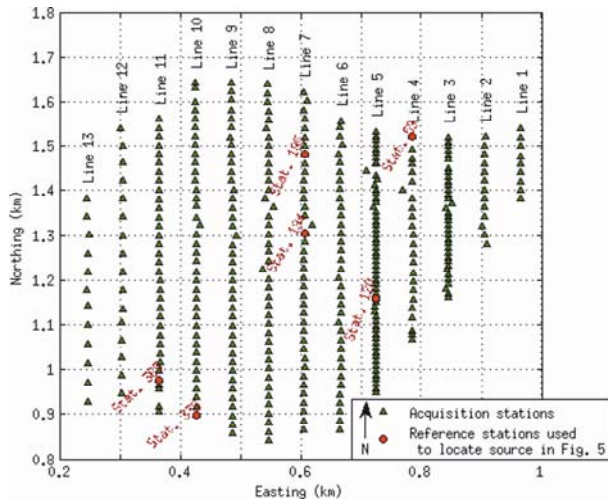


4 **Figure 1:** Data acquisition geometry. Green triangles, location of the 397 vertical-component geophones.

5 Red spots, reference stations used to localize the dominant source (see Figure 5) through MFP. The

6 network dimension is approximately 800 m × 900 m.

7



9 **Figure 2:** (a) Time recording (30 s long) of ambient seismic noise acquired at station #185 (see Figure 1).

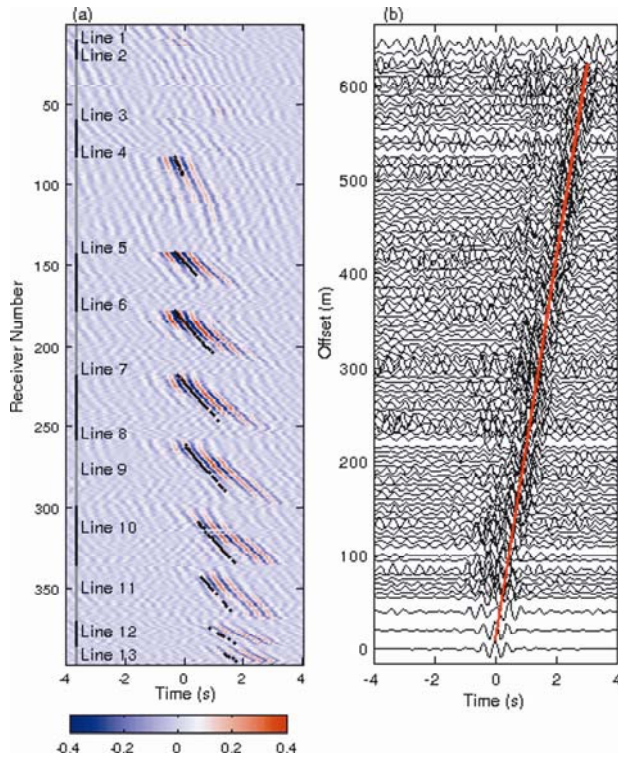10 (b) Spatially-averaged ambient noise spectrogram (18 h long) recorded from 5 p.m. of the first day, to 10

11 a.m. of the second day. Color scale represents the normalized spectral amplitude, in dB. The dotted lines

12 between (a) and (b) indicate the acquisition time of the 30 s signal. (c) Time average of the ambient noise

13 spectrogram in (b).

14

**Figure 3:** (a) Noise correlation functions with station #185 as reference in the 1.8 Hz to 3.5 Hz frequency band for a 2.5-min-long time recording. Color scale represents the amplitudes of the correlations. Black dots, theoretical travel-times calculated for the dominant source location retrieved from the MVDR analysis (see Figure 5). (b) Correlation functions stacked on 5-m intervals and sorted by offset from the reference station. Red line, group velocity of 230 m/s.
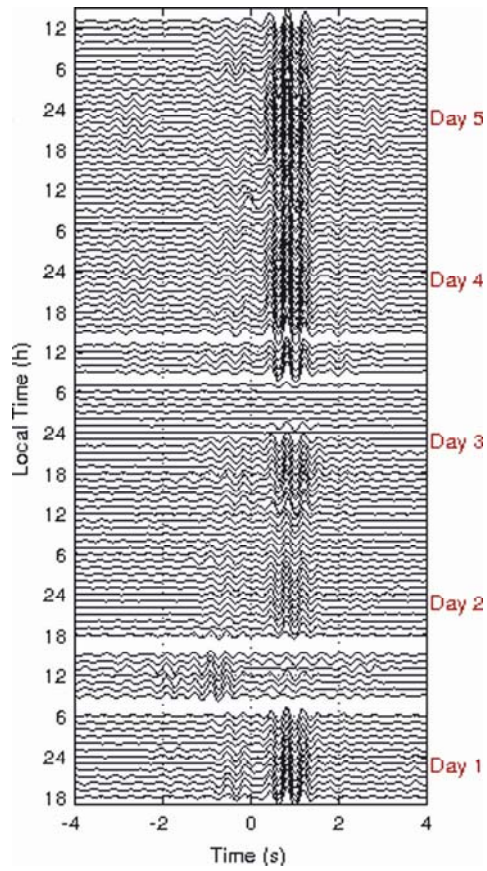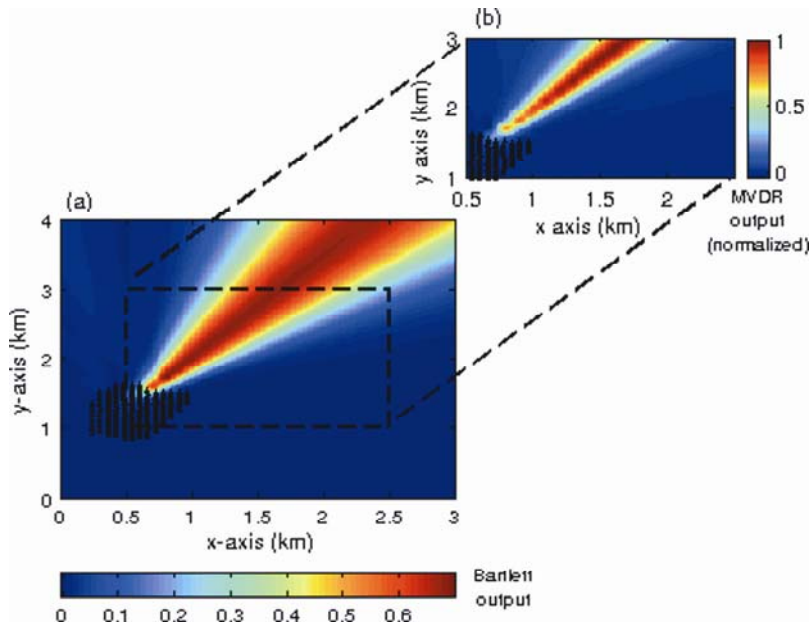
1

2 **Figure 4:** Time evolution of the 1 hour stacked cross-correlated function for stations #185 and #194 over

3 the five recording days. The inter-station distance is 177 m.

4

1



2

**Figure 5:** (a) Bartlett MFP output of the dominant noise source projected on a 100 m × 100 m search grid

using a 2.5-min time recording (pre-filtered between 1.8 Hz and 3.5 Hz) and a surface velocity model of

230 m/s. Color scale represents the amplitude of the MFP output. Gray triangles, geophone locations.

Black dotted box, expanded area in (b) Normalized MVDR, output obtained for the zoomed area in (a).
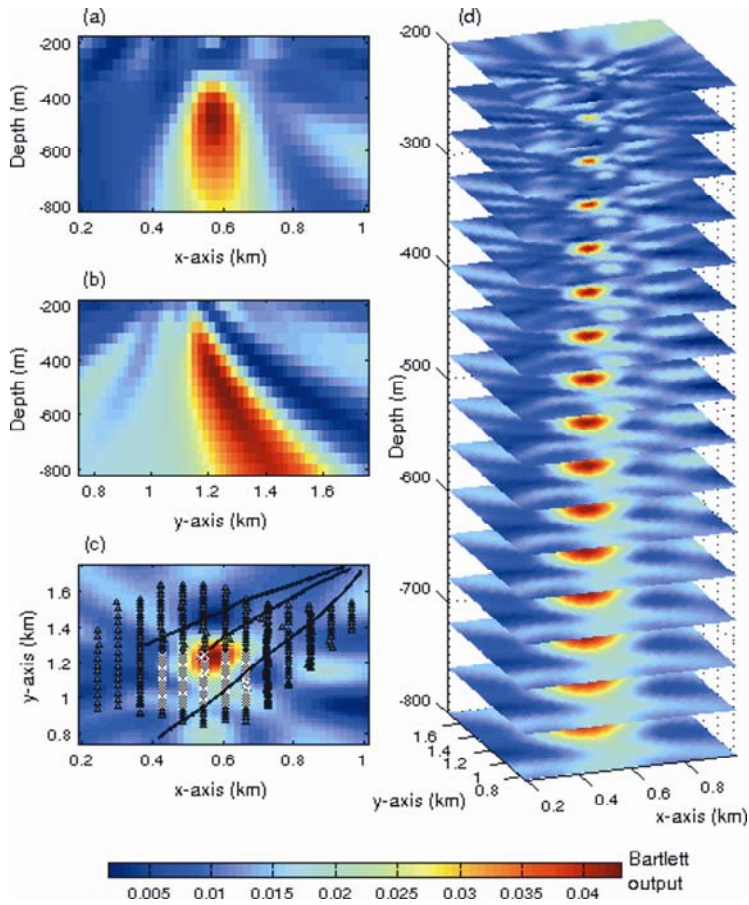
Both images are plotted using the same X-Y scale.

8
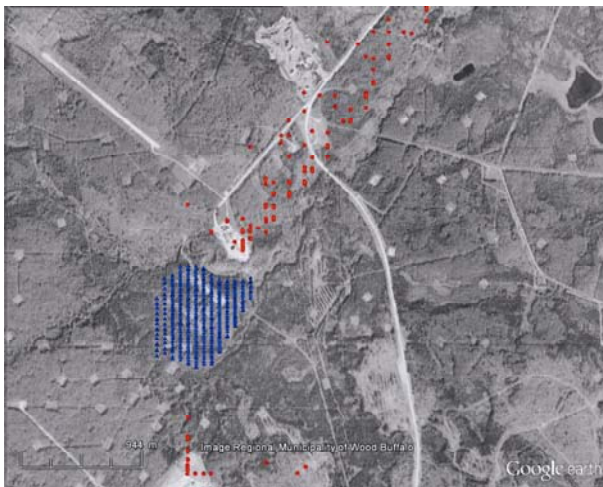


**Figure 6:** Geographic map of the dominant noise source positions retrieved at the surface using the

MVDR procedure. Red dots, noise source locations, which correlate well with human structures (road and

exploration platforms). Blue triangles, geophone locations.

1



2

3 **Figure 7:** Bartlett output obtained for the noise source at depth using the MRABF technique. The MFP

4 output is projected on a 20 m × 20 m × 40 m grid cell. (a), (b) and (c) X-Z, Y-Z and X-Y surface sections,

5 respectively, crossing the maximum of the MFP output. The source depth is estimated at 480 m. Black

6 triangles in (c), full network used to null the dominant surface source, gray cross, sub-network used to

7 localize the noise source at depth. Black lines, the well traces. (d) Three-dimensional MFP output

8 representation. Color scale represents the amplitude of the MFP output.

9
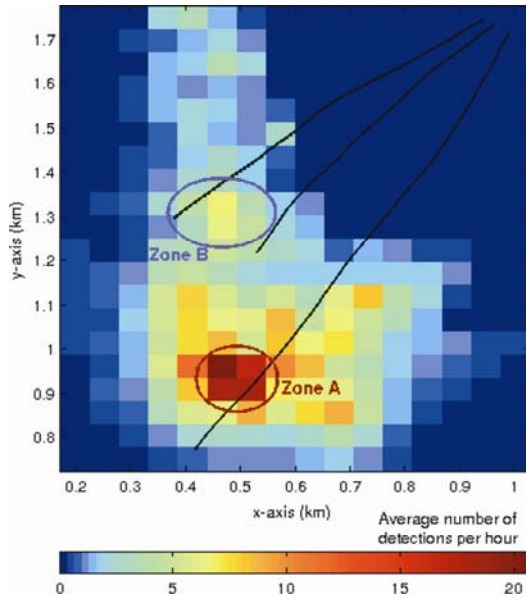
**Figure 8:** X-Y spatial distribution of the collection of low-amplitude sources at depth retrieved during the five recording days for a 2D grid with a 50m x 50m unit cell. Black lines, the X-Y projection of the injection/ extraction wells. Zone A and zone B (red and purple circles) correspond to areas of interest for the time evolution study plotted in Figure 9.
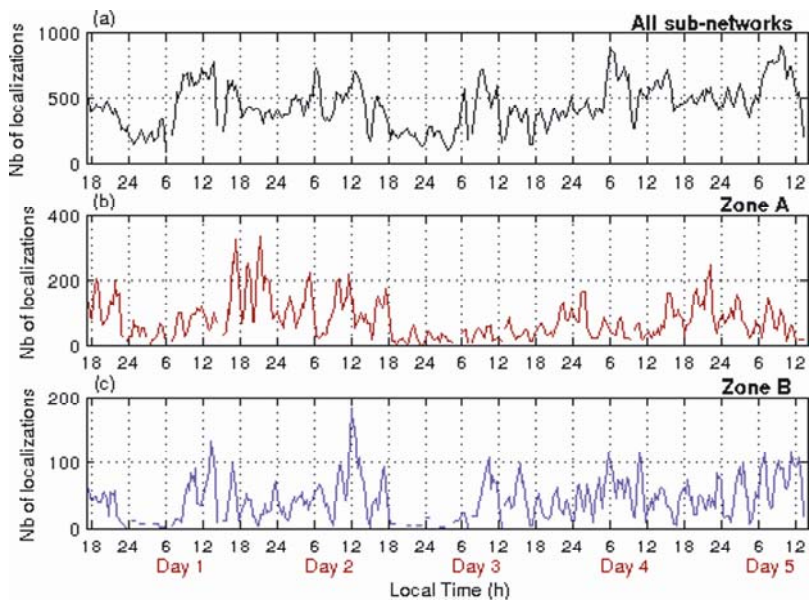


**Figure 9:** Temporal evolution of the source localizations at depth below the seismic array. (a) Number of localizations summed over the whole network. (b) and (c) Number of localizations summed over the circled areas in Figure 8 (zone A and zone B) at the tip of the extraction wells.