# Source depth estimation using modal decomposition and time-frequency representations

Grégoire Le Touzé, Barbara Nicolas, Jean-Louis Lacoume and Jérôme Mars
Laboratoire des Images et des Signaux
INPG/ENSIEG
Grenoble, France
Email: letouze@lis.inpg.fr

Dominique Fattaccioli
Centre Technique des Systèmes Navals
Toulon, France

***Abstract*— Source localization in shallow water environment is a crucial issue in underwater acoustics. Many methods for the source depth estimation use modal decomposition. Most of them use a multi-sensor array (vertical or horizontal). We propose a method based on time-frequency representations and Matched Mode Processing. It needs thus only one sensor. Modes amplitudes are estimated on time-frequency representation starting for a theoretical model of waveguide. Depth estimation is done comparing estimated modes amplitudes with predicted ones extracted from simulations at different depths. The robustness against noise is discussed. Validation on real dataset is given.**

## I. INTRODUCTION

Passive source localization in waveguide has been studied for many decades in underwater acoustics. For shallow water environment, beamforming techniques are inappropriate because they do not consider multipath arrivals and ocean acoustic channel complexity. Matched Field Processing can be used as it takes into account oceanic propagation. This technique consists in building and maximizing an objective function between simulated acoustic field and pressure field recorded on sensors. It can be combined with modal decomposition in shallow water where modes amplitudes (also called modes excitation factors) allow us to estimate the source depth. We are talking then of Matched Mode Processing. These modes amplitudes are usually estimated by spatial integration in depth of the pressure field on a vertical array [5] [6]. An alternative method has been developed using frequency-wavenumber transform and a horizontal array [11].

In this paper, we use both modal decomposition and Matched Mode Processing to estimate the source depth. Main difference with other methods is on the use of time-frequency representations based on one single sensor to estimate the modes excitation factors. To reach depth information, we estimate first the amplitudes on the real data and on simulated data at different depths. For each simulation we calculate an objective function between those two estimations. The estimated depth corresponds to the simulation that maximizes the objective function.

After a short presentation of guided propagation in shallow water environment, we develop a matched mode method of source depth estimation based on time-frequency representations. Choices of the time-frequency representations are discussed. A study of the robustness against noise of this method is presented and we apply it on real data from the North Sea.

## II. NORMAL MODES THEORY IN A WAVEGUIDE

The normal modes theory is appropriate for low frequency waves at a long range in an oceanic waveguide.

We consider a waveguide made of two homogeneous layers of fluids (known as a Pekeris waveguide [12]). The first layer is between perfectly reflecting boundaries at $z = 0$ and the second layer at $z = D$ ($c$ represents the water layer velocity and $\rho$ its density). The study is made for a harmonic point source located at $z = z_s$ but results can be extended for broadband source. Acoustic pressure $P(r, z, t)$ received at point $B(r, z)$ can be expressed by $P(r, z, t) = p(r, z)exp(-i\omega t)$ where $p(r, z)$ verifies the Helmholtz equation:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial p}{\partial r}\right)+\rho\frac{\partial}{\partial z}\left(\frac{1}{\rho}\frac{\partial p}{\partial z}\right)+\frac{\omega^2}{c^2}p=-\frac{\delta(r)\delta(z-z_s)}{2\pi r} \quad (1)$$

with the pulsation $\omega$. According to boundaries conditions and using technique of separation of variable, acoustic pressure field at long range can be expressed as a sum of modes:

$$p(r, z) = C\sum_{m=1}^{+\infty} u_m(z_s)u_m(z)\frac{\exp(ik_{rm}r)}{\sqrt{k_{rm}r}} \quad (2)$$

where $u_m$ is depth dependent as:

$$u_m(z) = \sqrt{\frac{2}{D}}\sin(k_{rm}z_s) \quad (3)$$

The vertical wavenumber component $k_{zm}$ depends of the guide model (perfect, Pekeris...). We see in equation 2 that pressure field depends upon a $u_m(z_s)$ factor: this is the mode amplitude. In the Pekeris waveguide case, mode amplitude depends on the frequency. As broadband signals are consider as integration of harmonic signals, it is not a problem for comparison between real and predicted modes amplitudes.

This short study of propagation in a shallow water waveguide shows us that the source depth $z_s$ only appears in modes amplitudes. As a result we can use these amplitudes to estimate $z_s$. Many methods using this property have been developed; they are called modal decomposition or mode filtering.

## III. Localization matched mode method

### A. Principles

The principle of modal decomposition consists in estimating modes amplitudes to perform localization. Classical amplitude evaluation techniques are based on the modes orthogonality properties along $z$ axis [5] [6]. By projection on modal space, amplitudes values can be deduce. These methods are particularly well adapted for mono-frequency signals (because modes amplitudes vary with frequency) and they use a large vertical array. An alternative method has been developed using horizontal array and frequency-wavenumber transforms [11]. In this specific space and for broadband signals, modes follow strictly separated curves. Helping with modal filtering in this plane, relative modal amplitude can be estimated. We are starting from the same principle but using the time-frequency space where broadband signals are described by modal curves.

### B. Modes excitation factors estimation

Geoacoustic parameters (number of layers, depths of layers, propagation velocities and densities) and the range $R$ allow us to calculate the time-frequency relationship $t_m(\nu_m)$ for each mode.

For the Pekeris model, if $\nu$ is the frequency and $k_{rm}$ the horizontal wavenumber, we obtain the characteristic equation:

$$D\left(\frac{\omega_m^2}{c_1^2}-k_{rm}^2\right)^{1/2}+\arctan\left(\frac{\rho_2\left(\frac{\omega_m^2}{c_1^2}-k_{rm}^2\right)^{1/2}}{\rho_1\left(k_{rm}^2-\frac{\omega_m^2}{c_2^2}\right)^{1/2}}\right)=m\pi \quad (4)$$

with $\omega_m = 2\pi\nu_m$. Group velocity $v_{gm}$ could also be known as a function of $k_{rm}$ [3]. As we have $t_m = R/v_{gm}$, we can deduce the relation between time and frequency $t_m(\nu_m)$ for a given mode $m$. This method can also be done for a multi-layers ($> 2$) case, group velocity can then be calculated numerically [7].

Time-frequency relationship $t_m(\nu_m)$ produces then a curve in the time-frequency plane for each mode $m$ (see figure 1). Modes amplitudes are distributed on these curves. Starting from them, we build binary masks by dilatation process and apply these masks on the time-frequency representation (a mask example is shown in figure 1). We make then a modal filtering by multiplying the time-frequency representation by given masks. If the time-frequency representation gives the energy density, the t-f representation on the mask region is the squared value of mode amplitude.

### C. Depth estimation

As we use Matched Mode Processing technique, we have to compare real modes amplitudes to predicted ones. To obtain these predicted modes amplitudes, we place a point source at each depth in a simulated waveguide. The acoustic field at the sensor is calculated. Simulated fields are obtained using a parabolic equation method made in an environment similar to the real environment (parameters are suppose to be known). Principles of parabolic method are exposed in [4].

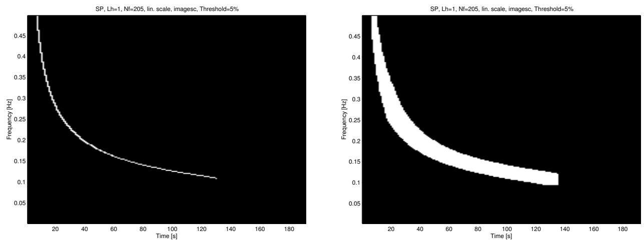

Fig. 1. Theoretical time-frequency curve for mode 3 (left) and mask associated with mode 3 (right)

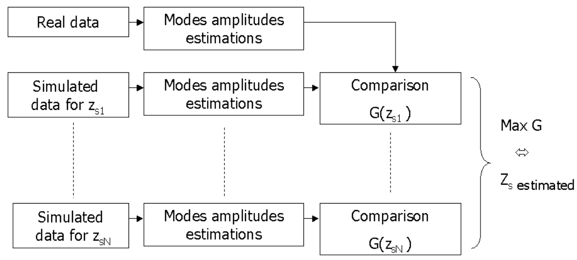


Fig. 2. Depth estimation process

The last step to compare measured and predicted modes amplitudes is to maximize the objective function which is the inverse dB function of the mean square error between the real and predicted modes amplitudes:

$$G_{dB} = 10log_{10}MSE^{-1} = 10log_{10}\left(\frac{nbr}{\sum_{modes}(c_{i_{simu}} - c_{i_{real}})^2}\right) \quad (5)$$

where $nbr$ is the number of modes. Then the estimated source depth is the depth associated to the simulation that maximizes the objective function $G$ (i.e. simulation matching most real data). Figure 2 summarizes the depth estimation process.

## IV. Time-frequency representations

### A. Introduction

Any time-frequency representation is subjected to the theoretical limitation known as the time-frequency uncertainty relationship (Heisenberg-Gabor inequality). This limitation prevents a precise localization simultaneously in time and frequency (a time-frequency compromise is necessary). In our case, this limitation causes an inevitable spreading of spectral elements out of the theoretical curves. Practical curves layouts on time-frequency plane are shown on figure 3 (we can also notice the depth dependence of the modes amplitudes).

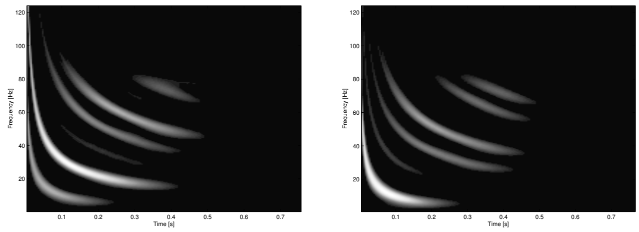

Fig. 3. Time-frequency representation of simulated data for $z_s = D/3$ (left) and $z_s = 2D/3$ (right)

Numerous techniques of time-frequency representations are referenced in signal processing due to the numerous possibilities to circumvent the time-frequency uncertainty. Of course, because we need select the various modes in the time-frequency plane, we would like to have the most precise localization on the time-frequency plane. In addition, a constraint must be respected for the representation choice: as we are working on mode amplitudes, it is necessary for the selected representations to be energetic (i.e. they must preserve the signal energy). Spectrogram, reassigned spectrogram or Lagunas representations respect this condition.

### B. Usual methods

*1) Spectrogram:* The spectrogram [8] is the most classical time-frequency method. It can be seen as the square modulus of the projection of the signal upon a filters bank for each time or upon time-frequency atoms (built on a reference window $h(t)$) which have all the same time duration and frequency bandwidth delayed around a time and central frequency:

$$S_x(\tau,\nu)_h = \left| \int x(t)h^*(t-\tau)e^{-j2\pi\nu t}dt \right|^2 \qquad (6)$$

A spectrogram example on real dataset is shown figure 4.

*2) Reassigned spectrogram:* Reassigned spectrogram [9] is particularly interesting because this technique, starting from the usual spectrogram, reassigns the spectrogram points with more respect of the energy localization. The time-frequency patterns are then more precise. But problems of overlapping due to time-frequency interferences are not resolved. A reassigned spectrogram example on the same real dataset is shown figure 5.

*3) Lagunas representation:* The Lagunas representation [10] is obtained by parametric methods. Starting from the autocorrelation function of the signal, it creates a filters bank respecting conditions:

- Keep the real value for central frequency of the filter;
- Reject the most as possible values for parasite frequencies;
- Keep the signal energy.

We have then:

$$RTF_{Lagunas}(n,\nu) = \frac{\underline{e}_\nu \underline{\underline{C}}^{-1}(n)\underline{e}_\nu^t}{\underline{e}_\nu \underline{\underline{C}}^{-2}(n)\underline{e}_\nu^t} \qquad (7)$$

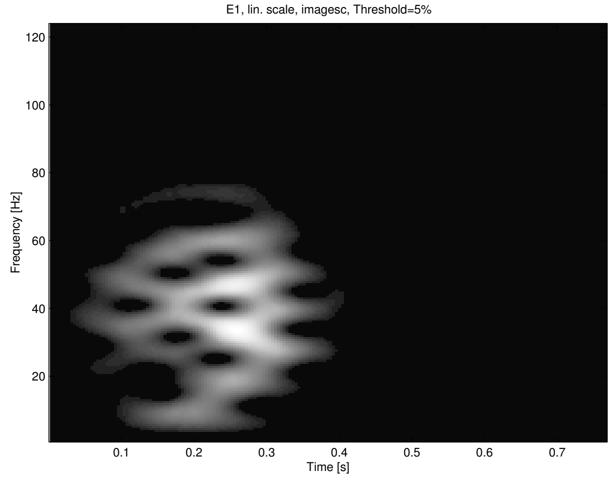


Fig. 4. Spectrogram on real dataset

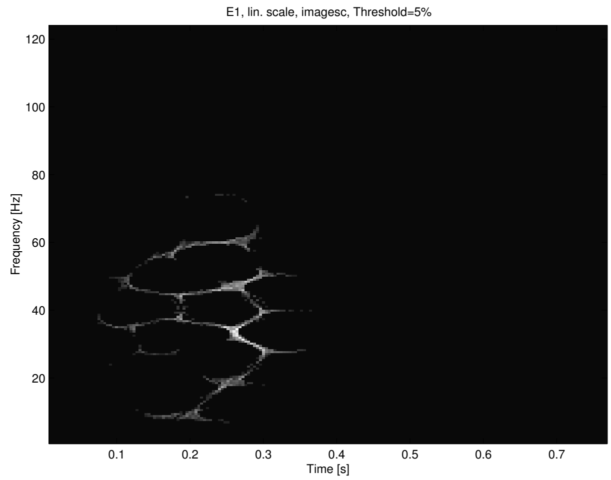


Fig. 5. Reassigned spectrogram on real dataset

with $\underline{\underline{C}}(n)$ the signal autocorrelation matrix and $\underline{e}_\nu = (1, e^{-2\pi j\nu}, ..., e^{-2\pi j\nu P})$ a pure frequency unitary vector. A Lagunas representation example is shown figure 6.

### C. Adapted methods

A time-frequency representation adapted to the pressure signal propagation in a perfect waveguide has been developed in [1]. The principle is to project the signal on the theoretical modal curves. To return on time-frequency domain, we allow to the mode number $m$ to be real and continuous. The representation is then a projection of the signal upon time-frequency "atoms". Atoms are built on a reference window, based on the layout of the modal curves and which cover all the time-frequency evolution domain of the signal. We have then:

$$TFR_h(\tau,\nu) = |\Psi_h(\tau,m)|^2_{m=m'} \qquad (8)$$

with $\Psi_h(\tau,m) = \int x(t)h(t-\tau)exp(j2\pi \int \nu_m(\theta)\ d\theta)\ dt$ the signal projection on for a given mode number $m$ and $m' = 2\nu D(C_1^2\tau^2 - R^2)^{1/2}(C_1^2\tau)^{-1} + 1/2$ the mode number value

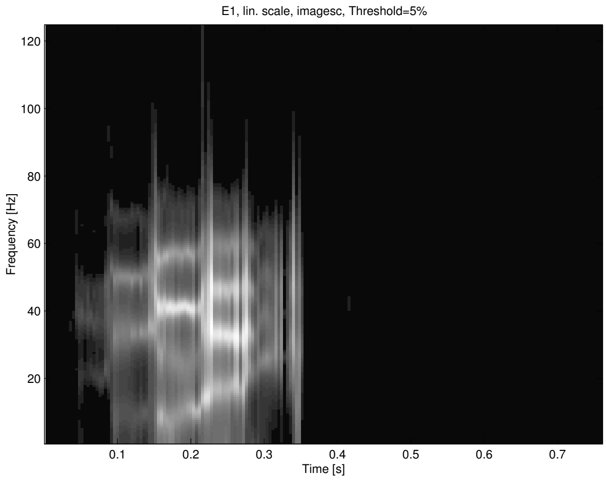


Fig. 6. Lagunas representation on real dataset

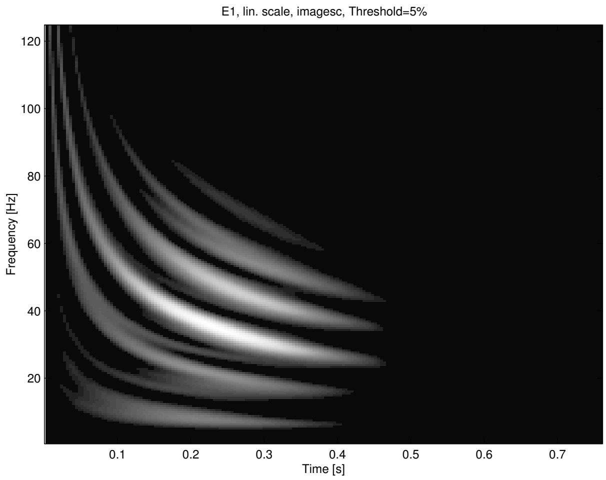


Fig. 7. Adapted method for perfect waveguide on real dataset

on a unspecified point of coordinate $(\tau, \nu)$, not necessary on a modal curve ($m'$ is then not necessary an integer so as to cover all the t-f space). This representation is shown on figure 7.

Starting from this method, we complete it in determining the atoms equations to allow this representation to be energetic: the closure condition on the window projection defined in [2] $\int\int h_{\tau,\nu}(t)h^*_{\tau,\nu}(t') \; d\tau \; d\nu = \delta(t-t')$ must be respected. Starting from the same principle we also create a representation adapted to an approximated Pekeris model. This model gives good results as shows the figure 8. In the same way, we determine atoms equations for the representation to be energetic.

## V. Applications

Techniques and methods described above are now used to estimate the source depth in two different environments: a noisy simulated environment and a real one.

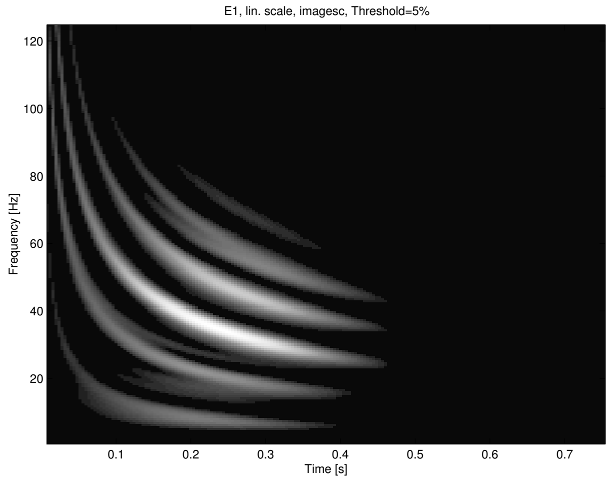


Fig. 8. Adapted method for approximated Pekeris waveguide on real dataset

### A. Masks dilatation

Dilatation of masks is necessary for two main reasons:
- The chosen model does not ever fit perfectly the real waveguide because of constant parameters hypothesis, uncertainties on geoacoustic parameters and eventually approximations (if we take a perfect waveguide model for instance).
- The time-frequency uncertainty, even when it is circumvented, has the consequence to displace time-frequency pattern around his exact place.

Other considerations must be respected:
- Masks must not overlap each other because we cannot decide that a time-frequency point makes part of two different modes.
- If we consider the noise as white Gaussian distribution, it is equidistributed on the time-frequency plane. Masks of different modes must approximately cover the same surface to avoid that one contains more noise (to have the same noise energy on the mode amplitude evaluation).

### B. Sensitivity to noise

To study the robustness against noise, we made manys simulation with a parabolic model in a Pekeris waveguide. The model configuration is presented figure 10, it is the real dataset environment studied on next part. We simulate the propagation of 25 sources located at different depths in the waveguide. A with Gaussian noise is added for each simulation. Source depth is then estimated with the method described in section III. We make this study with a Signal Noise Ratio of 3dB. Results are presented in table I.

Source depth estimations are satisfactory. A small error is frequently present but estimation misses never completely the true depth.

For the SNR problem, size of masks is an important consideration. Bigger the masks size is more the noise disturbs the results. In addition, we need a minimum masks size for reasons exposed above. Results presented are made with the spectrogram but we have the same proportions with the

| | Number of estimations | % of estimations |
|---|---|---|
| Error$< 1m$ | 4 | 16 |
| $1m$ <Error$< 3m$ | 15 | 60 |
| $3m$ <Error$< 5m$ | 5 | 20 |
| $5m$ <Error$< 7m$ | 1 | 5 |
| Error$> 7m$ | 0 | 0 |

TABLE I

ERRORS REPARTITION FOR A $SNR = 3dB$

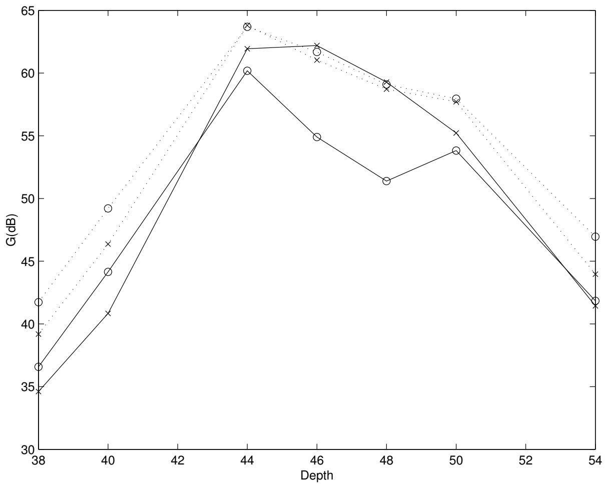


Fig. 9. Evolution of $G(dB)$ with depth. $SNR = 3dB$. Dotted line for spectrogram. Solid line for adapted method for approximated Pekeris. 'o' for a classical dilatation (dilatation index=7). 'x' for a smaller dilatation index (=2).

adapted method for approximated Pekeris waveguide (using the same masks size). The adapted method does not seem to present any new interest as it does not improve results. But if we take masks smaller, we have best results with the adapted method. Figure 9 presents the $G(dB)$ function layout for a real source depth $z_s = 46m$. We find a $2m$ error estimation with both methods for a classical dilatation (dilatation index=7) whereas true depth is found only by adapted method with smaller dilatation (dilatation index=2).

It can be explained considering that time-frequency uncertainty is circumvented along the mask layout with adapted methods. Time-frequency patterns are then less spread perpendicularly with the theoretical curves.

To conclude with noise problem, results can be improved using adapted method, as it is possible to use smaller masks in a more relevant way.

### C. Real data

We now estimate source depth on real data from North Sea experiment using various methods of time-frequency representations presented in previous section. The source is an air gun (5-80 Hz) located in the water layer (first layer). The pressure field is recorded by a single sensor laid on bottom at

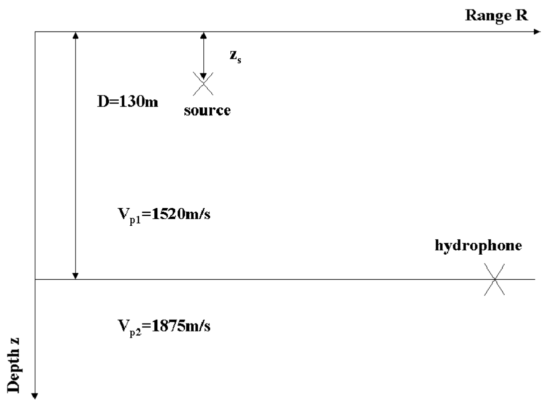


Fig. 10. Model configuration

$R = 3500m$. Time-frequency representations of this dataset are shown figures 4, 5, 6, 7 and 8. The second layer can be considered as a half space, experimental geometry is thus a Pekeris waveguide. The model configuration is presented figure 10.

We have the following geoacoustic parameters: in the water layer $V_{p1} = 1520m/s$, $\rho_1 = 1$ ; in the half space $V_{p2} = 1875m/s$ and $\rho_2 = 3$. Then using the t-f representations and our method described in section IV, modes amplitudes $c_{i_{real}}$ are calculated for this real data. A set of simulations at different source depths is realized. For each simulation, modes amplitudes $c_{i_{simu}}$ are calculated and compared to $c_{i_{real}}$ using function $G$. The source depth estimation is given by the depth associated to the simulation which maximizes $G$. Same masks are used on all the time-frequency representations. $G$ function is shown on figure 11 (main lobe is zoomed on figure 12).

We notice first that all methods give the same global layout. We have a main lobe around $z = 20$ and two sidelobes around $z = 64$ and $z = 122$. In all cases (except the Lagunas one), the estimated $z_s$ is $z_{s_{est}} = 20m$ with a bigger probability to be near $18m$ than $22m$. We have no exact information about the source depth except it is between $10m$ and $20m$. Moreover, ref. [11] finds for the same dataset $z_{s_{est}} = 17$. Our results are thus consistent with what we know and can be considered as a good estimation. The Lagunas layout is particular because higher than others. But it is the worst: as we see on figure 12 maximum value is for $z_{s_{est}} = 22m$ and we know that the true value is down $20m$, the main lobe is larger than for other methods and the difference between minimum and maximum of $G(dB)$ is smaller than for other cases.

Studies on Matched Mode Processing [5] [6] have also shown that signs and number of modes considered are important for the depth estimation. Here, we ignore the sign and work on a small number of modes ($nbr = 7$). Deducing the modes coefficients signs starting from the t-f phase plane will be the subject of a future study.

In a noisy environment, the adapted method gives better results as we have shown on previous part. But when it is not a noisy case as for this dataset and despite the

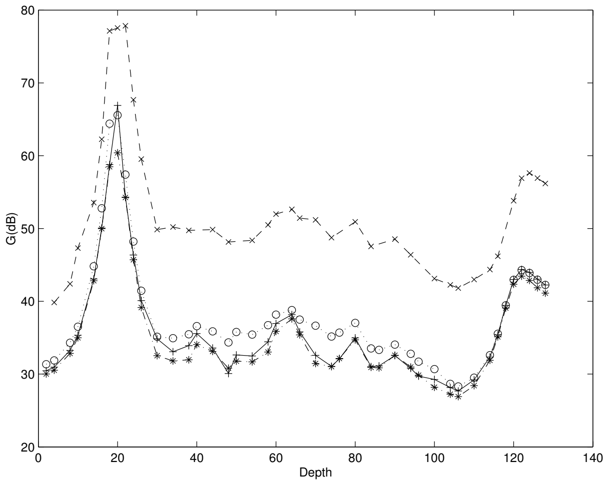


Fig. 11. Evolution of $G(dB)$ with depth. Solid '+' line for Adapted method, dotted 'o' line for spectrogram, dashdot '*' for reassigned spectrogram, dashed 'x' for Lagunas method

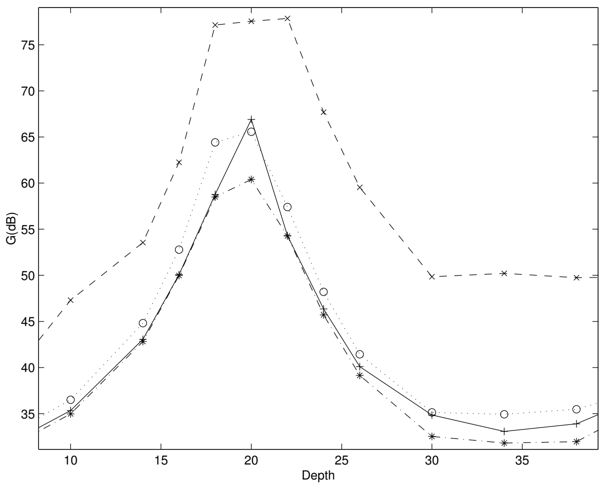


Fig. 12. Zoom on main lobe of the evolution of $G(dB)$ with depth. Solid '+' line for Adapted method, dotted 'o' line for spectrogram, dashdot '*' for reassigned spectrogram, dashed 'x' for Lagunas method

different pattern layouts (as we can show on various t-f figures), results are sensitivity the same whatever the time-frequency representation choice. Actually, a precise localization of modes is not necessary a paramount condition. This observation leads us to test a simple use of global t-f comparison (without the mode concept), i.e. of Matched Field Processing on the time-frequency plane. This will be the subject of future studies.

## VI. Conclusion

The depth of an acoustic source in the ocean can be determined by modes coefficients estimation. We propose an estimation based on time-frequency representations (needing only one sensor) and Matched Mode Processing. This method is providing good and accuracy results as we observe on real data. Results are comparable to those obtained on the same dataset starting from frequency-wavenumber transform. Time-frequency methods give all good results even in a noisy environment. However, for this last case, adapted method based on Pekeris waveguide model gives better results.

## References


[1] G. Faye Bourdeaux-Bartels C-S Chen and J.H. Miller. Time-frequency representations for broadbandacoustic signals in shallow water. In *Oceans 2003*

[2] P. Flandrin. *Temps-fréquence*. HERMES, 1993.

[3] L. Tolstoy and C.S. Clay *Ocean Acoustics : Theory and Experiment in Underwater Sound*. American Institute of Physics, 1987.

[4] F.B. Jensen W.A. Kupperman M.B. Porter and H. Schmidt *Computational ocean acoustics*. AIP Press, New-York, 1994.

[5] E.C. Shang Source depth estimation in waveguides *J.Acoust.Soc.Am.*, 77:1413–1418, 1985.

[6] T.C. Yang A method of range and depth estimation by modal decomposition. *J.Acoust.Soc.Am.*, 82:1736–1745, 1987.

[7] D.M.F. Chapman and D.D. Ellis The group velocity of normal mdoes. *J.Acoust.Soc.Am.*, 74(3):973–979, 1983.

[8] J.B. Allen and L.R. Rabiner A unified approach to Short Time Fourier Analysis and Synthesis *Proc. IEEE*, 65(11):1558–1564, 1977.

[9] F. Auger and P. Flandrin Improving the readability of Time-Frequency Time-Scale Representation by the Reassignment method *IEEE Transaction on Signal Processing*, 43(5):1068–1089, 1995.

[10] M.A. Lagunas M.E. Santamaria A. Gasull and A. Moreno Maximum likehood filters in spectral estimation problems *Signal Processing*, 10:19–34, 1986.

[11] B. Nicolas J. Mars and J-L. Lacoume. Source depth estimation using modal decomposition and frequency-wavenumber transform. In *12th European Signal Processing Conference, Vienna*, 2004.

[12] C.L. Pekeris. Theory of propagation of explosive sound in shallow water. In *Geol.Soc.Am.Mem*, 1947.