# Direct regressions for underwater acoustic source localization in fluctuating oceans

Riwal LEFORT*, Gaultier REAL+, Angélique DRÉMEAU*

*ENSTA Bretagne, 2 rue François Verny, 29806 Brest France
+DGA Naval Systems, Toulon, France
riwal.lefort@ensta-bretagne.fr

**Abstract**

In this paper, we show the potential of machine learning regarding the task of underwater source localization through a fluctuating ocean. Underwater source localization is classically addressed under the angle of inversion techniques. However, because an inversion scheme is necessarily based on the knowledge of the environmental parameters, it may be not well adapted to a random and fluctuating underwater channel. Conversely, machine learning only requires using a training database, the environmental characteristics underlying the regression models. This makes machine learning adapted to fluctuating channels. In this paper, we propose to use non linear regressions for source localization in fluctuating oceans. The kernel regression as well as the local linear regression are compared to typical inversion techniques, namely Matched Field Beamforming and the algorithm MUSIC. Our experiments use both real tank-based and simulated data, introduced in the works of G. Real *et al.* Based on Monte Carlo iterations, we show that the machine learning approaches may outperform the inversion techniques.

*Keywords:* Underwater source localization, fluctuating ocean, Machine learning, Regression

## 1. Introduction

In the underwater domain, the specific sound propagation properties make passive acoustics an interesting tool for underwater source localization. In this context, from the 70's [1, 2] to the present day [3, 4], the inversion strategy has remained a methodological reference. By using inversion, however, we necessarily make the strong assumption that the environmental properties are known, or at least known *a priori*. For instance, we must know the exact seabed depth distribution and the exact time-space distribution of both the temperature and the salinity. Unfortunately, these environmental parameters are in practice very fluctuating both in time and space, leading to strong mismatches between physical models and related real measures [5]. It was otherwise shown that small amplitude environmental fluctuations may induce drastic changes in the propagated acoustic pressure field. The idea behind this phenomenon is that the effect of these small fluctuations of the propagation medium is cumulative (see the so-called $\delta$-correlation approximation in [6]). These strong physical uncertainties make inversion a very tough task, so that researchers have developed some methods to jointly assess the source position and the environmental properties [7].

On the other hand, in a lot of research fields such as computer vision and speech recognition, machine learning has become a methodological reference, especially in the context of big data and deep learning [8, 9]. In addition to enabling real-time processing, the technique has proved to be very successful in comparison to the common baselines. In the opposite direction of inversion methods, machine learning is a "black-box" approach which does not need for any physical prior knowledge. Regarding the task of underwater source localization, it will naturally consider all the environmental parameters as underlying the regression parameters learned during a training step.

Machine learning has already proven its ability to accurately locate sources from sensor measurements. This is especially true in the field of robotics where a humanoid robot assesses a source position from a pair of acoustic sensors [10, 11]. But despite few works forecasting the relevance of machine learning in future developments of underwater passive acoustic systems (*e.g.* [12]), it has still never been used to locate underwater sources. One possible reason may lie in the fact that these methods require to build a training database beforehand, which may be impossible in certain rare, non-reproducible scenarios and, in any case, time consuming. However, we can *a contrario* target many situations where it is possible to acquire such groundtruthed databases. As an example, it is possible to register both time and space position of any oceanic event (*e.g.* seismic prospection, weather events, vessel activity) and to associate this event to the closest array measurement. Such an association between underwater acoustic measurements and the ocean activity has already been carried out in the context of weather forecast [13]. In the context of underwater source localization, we can think of synthetic simulations, miming the real forecasted environmental characteristics, or, *in situ* acquired data, making use of underwater sound synthesizers or taking advantage of sources of opportunity and recording the received acoustic pressure.

In this paper, we make use of two datasets introduced by G. Real *et al.* in [14, 15, 16, 17]. The first one is built from a software that simulates four increasing degrees of fluctuating environments. The other dataset is built from tank experiments where a "random lens" (called RAFAL in this paper, see section 4.1) simulates and reproduces the random effects of a fluctuating propagation channel. Both databases are interesting by their ability to synthesize increasing environmental deteriorations, from an ideal channel without any disturbance to a fully saturated environment [18].

The main contribution of this paper is in using direct regressions for the task of underwater source localization. We experimentally demonstrate that machine learning may outperform the inversion techniques in fluctuating environments. In particular, we investigate two regression models: a kernel regression and a piecewise linear regression, which appeared to be well-suited to our case of interest.

The paper is organized as follows. In section 2, we introduce the main principles of both inversion and machine learning for underwater source localization. In section 3, we present both the methods and the approximation we proposed to improve the computational efficiency. Then, in the experimental section 4, we compare the localization performance of the direct regressions with two of the main inversion references: the Matched Field Beamformer (MFBF) [19] and the MUSIC algorithm [20, 21]. This comparison is based on the measure of the localization error from Monte Carlo iterations. In section 5, we propose a discussion about the limitations of our study and the future perspectives of such machine learning approaches. We finally conclude the paper in section 6.

## 2. Problem statement

We suppose that a source is emitting a monochromatic signal at frequency $f$ from a position $y \in \mathbb{R}^{Q\times 1}$, where $Q$ stands for the number of position coordinates, according to the propagation assumptions (plane, cylindric or spheric waves, $2D$ or $3D$ propagation). This signal is measured by a passive acoustic array composed of $P$ sensors. Let $z \in \mathbb{C}^{P\times 1}$ be the Fourier Transform at frequency $f$ of the complex measured acoustic pressure. For any measurement

$z$, we try to assess its related position $y$. Note that, in practice, underwater source localization considers several snapshots, *i.e.* a set of measurements $\{z_n\}_{n=1}^N$ for a single source position $y$. For a sake of clarity, we only present the methods for a single snapshot, a simple averaging strategy being carried out for several snapshots.

### *2.1. Inversion for source localization*

An inversion technique considers the following optimization problem:

$$\hat{y} = \arg\min_y D\left[z, f_\theta(y)\right], \tag{1}$$

where the function $D$ measures how much the current *in situ* observation $z$ fits a given model $f_\theta(y) \in \mathbb{C}^{P\times 1}$.

The model $f_\theta(y)$ is an analytical deterministic expression which predicts the measured acoustic pressure from a source position $y$. The model parameters $\theta$ may refer to any propagation properties such as the temperature, the salinity, the sound speed, the seabed characteristics or the transducer parameters. The analytical expression of $f_\theta(y)$ may derive from a modal form of the sound propagation [22, 23]. Many contributions have focused on choosing an appropriated distance measure $D$. This distance often takes the form of a correlation-based measure [19]. In order to deal with the issue of measuring a distance in a high-dimensional space, other works (see [20, 21]) consider the signal subspace projection by eigen decomposition, the distance $D$ being computed in the mapped space. Sparse-based distance measures have furthermore proven their ability to be more accurate in the presence of multiple sources [24, 25, 3]. An other category of contributions includes the introduction of randomness and uncertainty to model the array noise or a fluctuating environment [26, 7, 27]. In that latter case, the inversion usually consists of assessing both the source position and the environmental properties, by maximizing a likelihood-based criterion:

$$D\left[z, f_\theta(y)\right] = -p(z|y). \tag{2}$$

More recently, the propagation uncertainty has been modeled by using the evidential theory [4]. Note finally that, although the optimization problem (1) is usually solved by grid search, we now find papers dealing with continuous optimizations [25].

### 2.2. Machine learning for source localization

Without a loss of generality, we formalize the machine learning techniques considered in this paper as follows. Let $\mathcal{R}(z)$ (resp. $\mathcal{I}(z)$) denotes the real (resp. imaginary) part of the complex pressure and $x = \{\mathcal{R}(z), \mathcal{I}(z)\} \in \mathbb{R}^{2P\times 1}$ be the vector concatenating them. Then, machine learning directly assesses the related position $y$ from a regression model:

$$\hat{y} = g_\gamma(x), \tag{3}$$

where $\gamma$ denotes the unknown regression parameters.

With all precautions we have given in section 1, we suppose that we are able to build a training database $\{x_n, y_n\}_{n=1}^N$, where $x_n \in \mathbb{R}^{2P\times 1}$ and $y_n \in \mathbb{R}^{Q\times 1}$. Machine learning consists then of optimizing the parameters $\gamma$ from the training data $\{x_n, y_n\}_{n=1}^N$. In comparison to the inversion paradigm, machine learning does not explicitely use the environmental parameters $\theta$. They underlie however the dependencies between each pair of training samples $\{x_n, y_n\}$, $\forall n$. These dependencies are then modeled by the regression function $g_\gamma$ for specific values of the parameters $\gamma$. In other words, while the channel characteristics $\theta$ clearly appear in the inversion expression (1), they disappear in the analytical regression expression (3), in favor of well-managed regression parameters. This makes machine learning highly interesting for random fluctuating environments.

## 3. Non linear regression

Regarding the specific application of passive underwater acoustics, we have experimentally observed that the location $y$ can not be expressed as a linear combination of the components of $x$. This is illustrated in Figure 4 where the error reaches its maximum value in the case of linear regression. Therefore, in this paper, we mainly focus on two non linear regression models, namely the local linear regression and the kernel regression.

### 3.1. Local linear regression

Let us first consider the linear regression model:

$$g_\gamma(x) = Ax, \tag{4}$$

where $\gamma = A \in \mathbb{R}^{Q\times P}$. In the training step, the matrix $A$ is learned from the training database $\{x_n, y_n\}_{n=1}^N$ as follows. Let $X = [x_1, \ldots, x_N]$ (resp.

$Y = [y_1, \ldots, y_N]$) be the matrix of the concatenated $\{x_n\}_{n=1}^N$ (resp. $\{y_n\}_{n=1}^N$), we look for $A$ satisfying $Y = AX$, or equivalently $Y^T = X^T A^T$. This can be achieved by solving, $\forall i \in \{1, \ldots, Q\}$,

$$\hat{a}_i = \arg\min_{a_i} \left\| \tilde{y}_i - X^T a_i \right\|_2^2 + \mu \left\| a_i \right\|_2^2, \tag{5}$$

where $a_i$ is the $i$-th row of $A$ (or equivalently the $i$-th column of $A^T$) and $\tilde{y}_i$ is the $i$-th column of $Y^T$.

Without any *a priori* on the expected values in $\hat{a}_i$, we chose the ridge regularization $\|a_i\|_2^2$ to help improving the conditioning of the problem (see *e.g.* [28]). This choice leads to a convex and differentiable problem, for which simple and efficient resolution algorithms exist, as the well-known gradient algorithm. The value of $\mu$, determining the weight of the regularization term over the data-attached term $\left\| \tilde{y}_i - X^T a_i \right\|_2^2$, is further discussed in section 4.4.

To extend the linear model (4) to a non-linear one, a common strategy consists of fitting a piecewise linear regression [11]. The feature space is first partitioned into $K$ clusters by using any clustering technique. In our case, we use a fast implementation of $K$-means [29]. Let $I_k(x) = 1$ if $x$ belongs to the cluster indexed by $k$, $I_k(x) = 0$ else. The piecewise non linear regression takes thus the form of a sum representing the contribution of each cluster:

$$g_\gamma(x) = \sum_{k=1}^{K} I_k(x) A_k x, \tag{6}$$

where each matrix $A_k \in \mathbb{R}^{Q \times P}$ is learned by using the training samples that belong to the cluster $k$ only, and $\gamma = \{A_k\}_{k=1}^K$. In a formal way, the optimization problem we use to train each matrix $A_k$ is then defined as, $\forall i \in \{1, \ldots, Q\}$,

$$\hat{a}_i^{(k)} = \arg\min_{a_i} \left\| \tilde{y}_i^{(k)} - X_k^T a_i \right\|_2^2 + \mu \left\| a_i \right\|_2^2, \tag{7}$$

where $X_k$ (resp. $Y_k$) is the matrix made up of the $x_n$ (resp. $y_n$) such as $I_k(x_n) = 1$ and $\tilde{y}_i^{(k)}$ is the $i$-th column of $Y_k^T$. The resolution of (7) is then the same as for problem (5).

### *3.2. Kernel regression*

Kernel regression is one of the first proposed non linear regression techniques [30]. The method aims at approximating the conditional expectation

$E[y|x]$. Introducing the parametric kernel $K_\gamma(x) = \exp\frac{-\|x\|^2}{\gamma}$, this is empirically achieved by the following regression model:

$$g_\gamma(x) = \frac{\sum_{n=1}^{N} K_\gamma(x - x_n)\; y_n}{\sum_{n=1}^{N} K_\gamma(x - x_n)} \simeq E[y|x] \tag{8}$$

This kernel regression does not require a training step, the equation (8) being directly expressed as a function of the training data $\{x_n, y_n\}_{n=1}^N$. However, this method may produce a huge computational cost because computing (8) depends on both the size of the training dataset ($N$) and the size of the measured vector ($2P$). Regarding the problem of source localization in a 3D environment from a large sensor array, we potentially have many training samples living in a high-dimensional space. Consequently, for computational efficiency, we consider a $L$-nearest neighbor-based approximation [31] of the kernel model (8):

$$g_\gamma(x) = \frac{1}{L} \sum_{n \in \mathcal{S}_L(x)} y_n, \tag{9}$$

where the set $\mathcal{S}_L(x)$ contains the index values of the $L$-nearest neighbors of $x$. The algorithm is thus very fast, only consisting of computing the squared Euclidean distance $\|x - x_n\|_2^2$, $\forall n$, and then, of averaging the source position of the $L$ closest samples.

## 4. Experiments

The evaluation databases and the evaluation protocol are respectively presented in section 4.1 and 4.2, while the main results are presented in section 4.3. Finally, in section 4.4, we analyze the parameter sensitivity as well as the way we set the free parameters.

### *4.1. Evaluation databases*

The experiments are based on two databases collected by G. Real *et. al.* [14]. They are composed of experimental signals acquired in a water tank and of the corresponding parabolic equation (PE) simulations. The following paragraphs are dedicated to the description of the tank experiments. The PE code reenacts the experiment using a 3D propagation code adapted from the one developed by X. Cristol *et. al.* [32].

#### 4.1.1. Acquisition protocol

A scaled experimental protocol was developed in order to reproduce faithfully the influence of spatial sound speed fluctuations in an oceanic medium perturbed by phenomena such as linear internal waves. A mobile transducer transmits an ultrasonic wave through a RAndom Faced Acoustic Lens (or RAFAL) presenting a plane "input" face and a randomly rough "output" face. The random roughness of the output profile induces distortions to the propagated acoustic field. The latter is recorded using a mobile hydrophone whose automatic displacements allow to simulate virtual linear arrays. A diagram of this experiment is proposed in Figure 1.

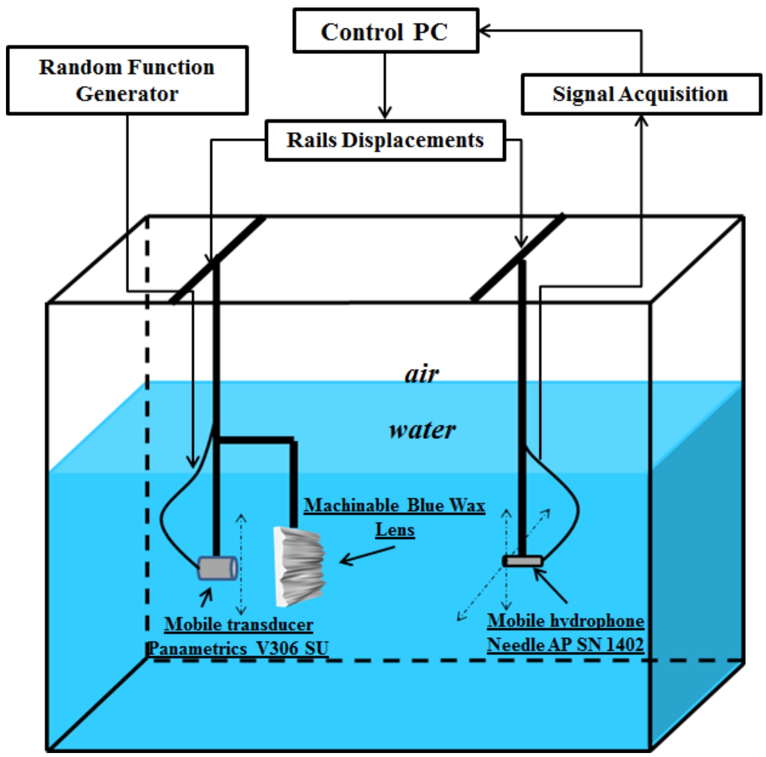



**Figure 1: Tank experiment diagram.**

From the mobile hydrophone, 65-elements virtual arrays were simulated, *e.g.* $P = 65$. The hydrophone displacement was of 0.3 mm in order to satisfy the sampling criterion (displacement $< \lambda/2$, where $\lambda = 0.665$ mm denotes the wavelength of the emitted signal in a fresh water at 20 degrees). The emitted signal is then a monochromatic wave train at a frequency $f = 2.25$ MHz.

The transducer is also fixed on a motorized rail, which allows to acoustically highlight statistically independent areas on the RAFAL. Therefore, multiple realizations of the same process can be obtained, and statistical studies can be carried out.

#### *4.1.2. Dimensional analysis*

The induced acoustic distortions are compared to what can be observed in a fluctuating ocean using a dimensional analysis [16]. The evaluation of the strength and diffraction parameters (respectively noted $\Phi$ and $\Lambda$) defined by Flatté [18] allows us to qualitatively relate the acoustic features in our experimental configurations and in an oceanic medium. Calculations (detailed in [17]) provide analytical expressions depending on a set of parameters including signal frequency, propagation distance, RAFAL's output face random roughness amplitude, vertical and horizontal correlation lengths. Equating the henceforth obtained dimensional parameters in this case and in the oceanic case provides a direct correspondence between sets of parameters in both configurations. In the ocean, the parameters involved in the calculation of $\Phi$ and $\Lambda$ are the signal frequency, the sound speed fluctuations amplitude and correlation lengths (horizontal and vertical) and the propagation range. This scaling procedure allows us to, in a controlled and reproducible fashion, acquire acoustic data spanning the various regimes of fluctuations introduced by Flatté [18]:

- The unsaturation (UnS) regime, where phase fluctuations due to medium inhomogeneities.
- The partially saturated (PS) regime, where the appearance of correlated micropaths is likely.
- The fully saturated (FS) regime, where uncorrelated micropaths appear.

In addition, a flat regime (Flat) is added to this study: this is the case where the RAFAL's output face was flat as well (no fluctuations induced). The quantitative accuracy of this scaling process is measured using the mutual coherence function. Both qualitative and quantitative relevance of the presented experimental scheme were validated in [16, 17]. Moreover, the influence of the signal fluctuations on the loss of array gain was exhibited in [15]. These results emphasize the need for innovative signal processing techniques regarding detection and localization of acoustic sources, such as proposed in the present paper.

| | Tank experiments | software |
|---|---|---|
| Flat lens (Flat) | $N = 845$ | $N = 960$ |
| Unsaturated regime (UnS) | $N = 7098$ | $N = 81792$ |
| Partially Saturated regime (PS) | $N = 5577$ | $N = 115200$ |
| Fully Saturated regime (UnS) | $N = 6084$ | $N = 120960$ |

**Table 1: Number of training samples ($N$) for each configuration.**

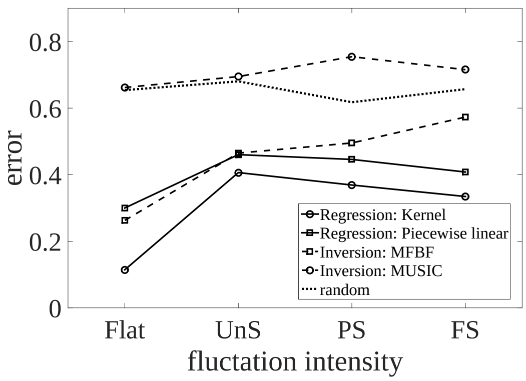
### *4.2. Evaluation protocol*

A total of 80 Monte Carlo iterations is carried out. For each of them, we randomly select a position $y$ and pick 10 corresponding signals measured on the antenna from the dataset. The remaining signals are used as training data to learn the regressions exposed above. The resulting size of the training dataset is given in Table 1 for each fluctuation scenario. To assess the robustness to noise of the different approaches, a zero-mean Gaussian noise of varying variance is added to each of the 10 test snapshots. This protocol allows us to compare the localization performance of both inversion and regression from exactly the same data.

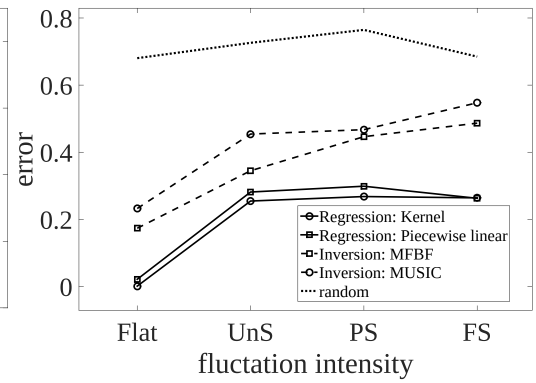
In order to measure the localization performance, at each Monte Carlo iteration, we measure the $L_1$-based distance between the estimated position and the groundtruthed position: $\|y - \hat{y}\|_1$. An alternative solution consists of using a $L_2$-based distance $\|y - \hat{y}\|_2$, but we may be misled by an averaging effect. Note that both the vertical and the horizontal positions are normalized by the domain range $[y_{\min}, y_{\max}]$ that we use for the grid search inversion in equation (1), where $y_{\min}$, $y_{\min} \in \mathbb{R}^{Q\times 1}$. This normalization is necessary to give every space components an equal weight. The error is finally averaged over the 80 Monte-Carlo iterations to obtain a global value.

Four localization methods are compared: the local linear regression (section 3.1), the kernel regression (section 3.2), and two typical inversion strategies, namely the Matched Field Beamforming (MFBF) [19] and the algorithm MUSIC [20, 21]. For those latter, we consider the following replica model, attached to the considered tank experiments [17]:

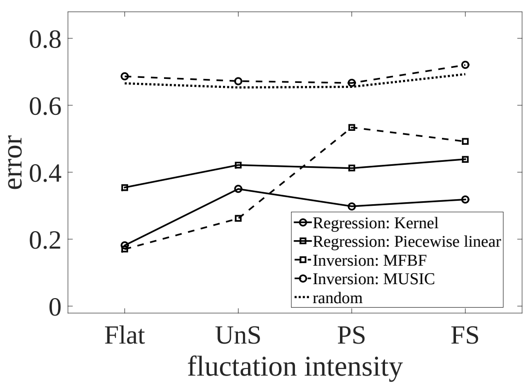
$$a(r, \phi) = S\left(\frac{2\pi}{\lambda}\rho \sin(\phi)\right) e^{-j\frac{2\pi}{\lambda}r}, \tag{10}$$



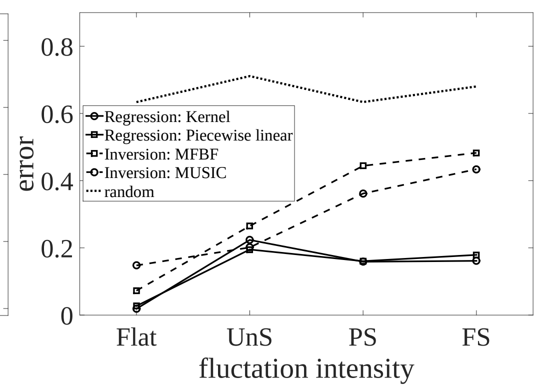

**(a)** Tank experiments, SNR=-10dB




**(b)** Tank experiments, SNR=10dB




**(c)** Software-based, SNR=-10dB



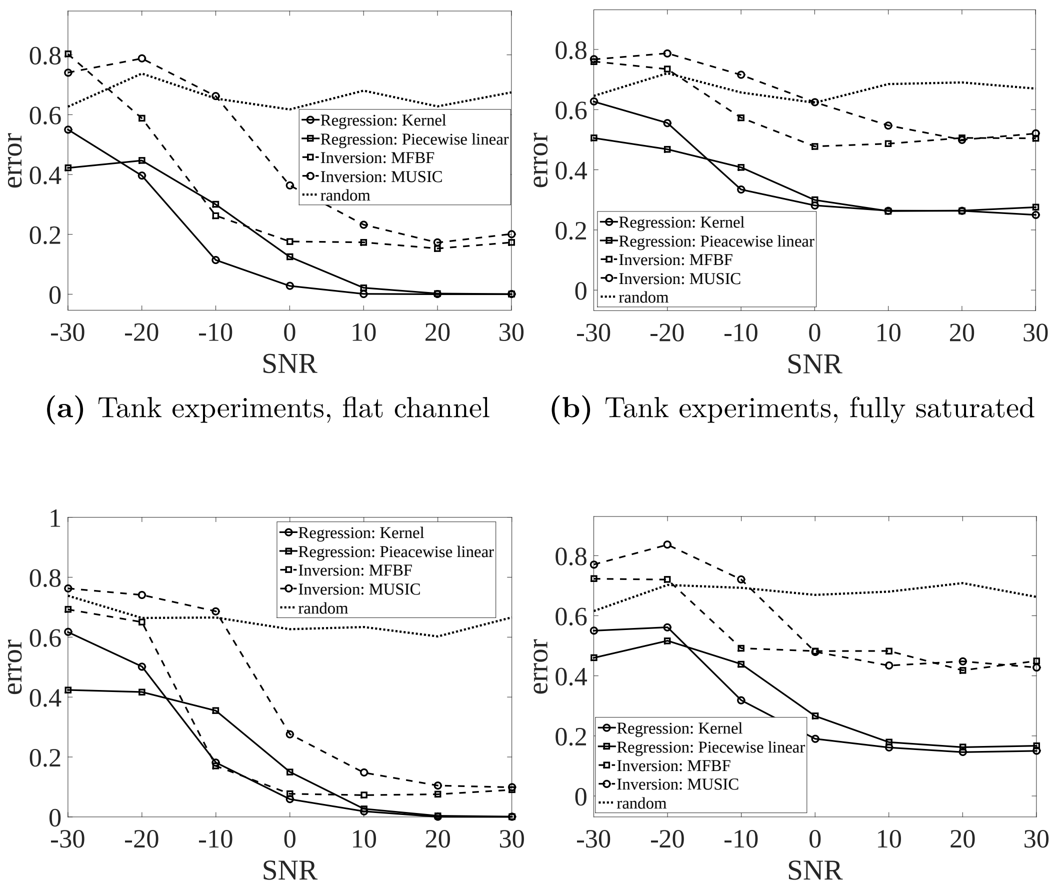

**(d)** Software-based, SNR=10dB

**Figure 2: Source localization error as a function of the channel perturbation regime.**

where $r$ and $\phi$ are respectively the propagation distance and the source elevation angle and constitute the position coordinates of interest, $\rho = 6.5$ mm the transducer radius and $S(.)$ stands for the so-called Sombrero function as defined in [33].

### *4.3. Main results*

In Figures 2 and 3, the regression methods are represented by continuous lines while the inversion ones are represented by dashed lines. In order to realize how much these methods perform, we also report the localization results




**(a)** Tank experiments, flat channel **(b)** Tank experiments, fully saturated

**(c)** Software-based, flat channel **(d)** Software-based, fully saturated

**Figure 3: Source localization error as a function of the Signal to Noise Ratio (SNR) in decibel.**

from a random source placement. In the figures, this method is qualified as "random" in the figures and represented by dot lines.

In Figure 2, we report the localization errors as a function of the channel perturbation regime, from a Flat regime without perturbation to a fully saturated regime (FS). As expected, the improvement led by the regression methods is more visible when the fluctuations become larger (the gap between regression and inversion methods increases). We notice that in average the regression outperforms the inversion-based method. But, the most interesting observation is that a machine learning strategy is more recommended for fluctuating regimes where the environmental characteristic $\theta$ are unknown and the mismatch between $f_\theta(y)$ and $z$ reaches its maximum. Indeed, for the two regression techniques, the localization performance remains quite stable from the unsaturated regime (UnS) to the fully saturated regime (FS). In comparison, the localization error obtained by the inversion methods increases when the channel fluctuation increases. This trend is perfectly illustrated in Figure 2d: while inversion and regression provide quite similar performance for both the Flat and the unsaturated regimes (UnS), regression outperforms inversion for both the partially saturated (PS) and fully saturated regimes (FS).

The above description remains valid regarding Figure 3, where we have reported the source localization error as a function of the SNR in decibel. As expected, the higher the SNR, the less the error. We observe the general trend that machine learning outperforms inversion, not only regarding the way the channel is fluctuating, but also regarding the robustness to the noise. This is especially true for highly saturated regimes (Figure 3b and Figure 3d).

From both Figure 2 and Figure 3, we observe that the kernel regression slightly outperforms the local linear regression. This is mainly due to the fact that the kernel regression is a continuous model. Conversely, the local linear regression is based on a vector quantization of the feature space by using a $K$-means clustering. The localization performance of the local linear regression thus depends on the space partition we get. The ideal local regression would consider a supervised learning of this partition. In other words, we should solve an optimization problem that learns the best clustering realization for each targeted database. Placing the clustering problem into a Bayesian framework, we could also consider using an Expectation-Maximization (EM) algorithm (as *e.g.* in [34]) to weight the contributions of the entire dataset rather than an “in-out” strategy.

We explain the poor results obtained by MUSIC by the weak number of snapshots we simulated. Actually, we use only 10 snapshots, which is not enough to correctly assess the covariance matrix on which the eigen decomposition is based.

In Table 2, we analyze the standard deviation of the error we obtain from the 80 Monte Carlo iterations. The standard deviation is reported as a function of the saturation regime (Flat, UnS, PS, FS), and for different values of the SNR. For the task of source localization or source classification, we often observe that the better a method performs, the less the standard deviation. Following this trend, the standard deviation due to the regression is less important than the one obtained from the inversion techniques. This

(a) Tank-based experiments:

| Fluctuating regime | Flat | | UnS | | PS | | FS | |
|---|---|---|---|---|---|---|---|---|
| SNR (dB) | -10 | +10 | -10 | +10 | -10 | +10 | -10 | +10 |
| kernel regression | 0.11 | 0.00 | 0.26 | 0.21 | 0.22 | 0.19 | 0.22 | 0.16 |
| MFBF | 0.17 | 0.11 | 0.31 | 0.28 | 0.30 | 0.26 | 0.30 | 0.28 |

(b) Software-based experiments:

| Fluctuating regime | Flat | | UnS | | PS | | FS | |
|---|---|---|---|---|---|---|---|---|
| SNR (dB) | -10 | +10 | -10 | +10 | -10 | +10 | -10 | +10 |
| kernel regression | 0.16 | 0.02 | 0.29 | 0.22 | 0.23 | 0.11 | 0.21 | 0.12 |
| MFBF | 0.11 | 0.04 | 0.26 | 0.17 | 0.26 | 0.24 | 0.28 | 0.22 |

**Table 2: The standard deviation of the localization error from the Monte Carlo iterations is reported as a function of both the fluctuating regime (Flat, Unsaturated (UnS), Partially Saturated (PS) and Fully Saturated (FS) and the signal to noise ration (SNR). The standard deviation is reported for both (a) the tank-based experiments and (b) the simulated experiments.**

is even truer when the channel perturbation increases.

### *4.4. Parameter sensitivity*

For the sake of simplicity and to reduce the computational time, the sensitivity of the free parameters (namely $\mu$ and $K$ in (6)-(7), $L$ in (9)) is analyzed on a single scenario. We consider the specific case of UnSaturated regime (UnS) and a SNR that equals 30 dB. In addition, we consider a single random split to design training and test data, and there are only 10 iterations to generate the random additive noise.

The sensitivity of the local linear regression is reported in Figure 4. The localization error is evaluated as a function of both the number of nearest neighbors $L$ and the regularization parameter $\mu$. As expected, the higher $K$, the higher it outperforms. This illustrates that a pure linear regression ($K = 1$) does not satisfy our non-linear problem. Regarding the regularization parameter $\mu$, we are encouraged to use low values. Indeed, for value such that $\mu \leqslant 10^{-1}$, the localization performance remains stable. Note that this experiment points out the interest of using a ridge constraint, the optimal values being different from $\mu = 0$ for $K \leqslant 512$ in equation (6).

The sensitivity of the kernel parameter $L$ is studied in Figure 5. From this result, we notice that, for this specific scenario, the localization performances are quite stable in the range $L \in [4, 128]$.

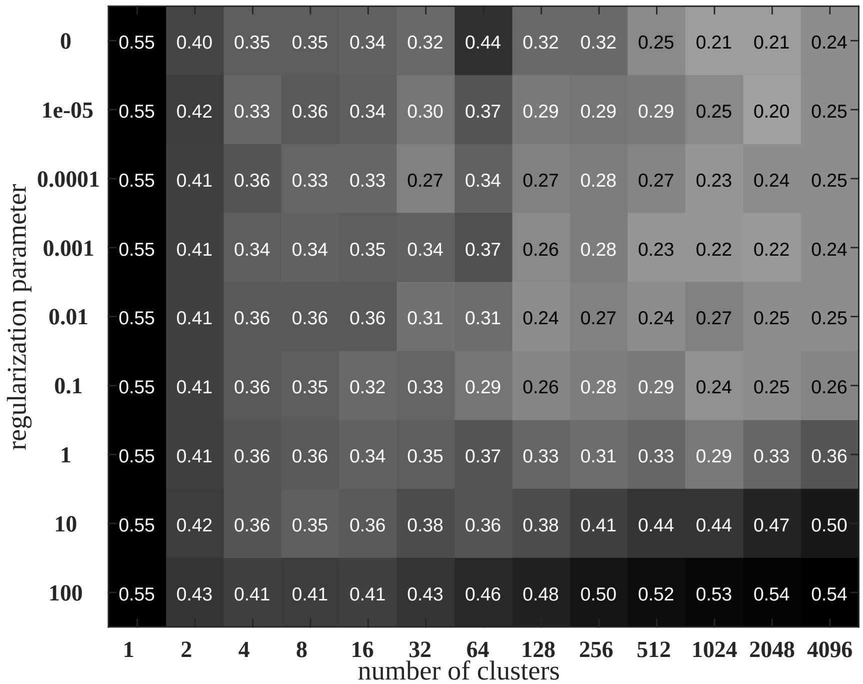



**Figure 4: Parameter sensitivity of the local linear regression. The average $L_1$ error is reported as a function of both the regularization parameter $\mu$ and the number of clusters $K$.**

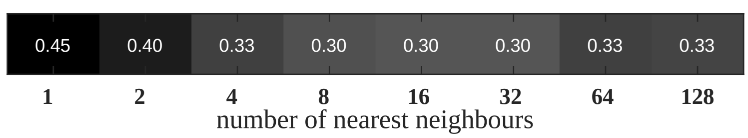



**Figure 5: Parameter sensitivity of the kernel regression. The average $L_1$ error is reported as a function of the number of nearest neighbors $L$.**

The baseline MUSIC relies on a separation between the noise and the signal. This classification is based on a projection onto a basis defined by the eigen vectors that correspond to the lowest eigen values. We must set the number of lowest eigen values, say $\delta$, *i.e.* the size of the projection space. In Figure 6, we report the localization performance as a function of the projection space dimension $\delta$. Based on this analysis, we encourage to consider a space size in the range $\delta \in [5, 20]$.

We use this sensitivity analysis to set the free parameters of each localization method. In machine learning, these parameters are usually set by

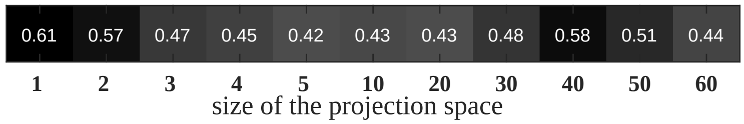



**Figure 6: Parameter sensitivity of MUSIC algorithm. The average $L_1$ error is reported as a function of the size $\delta$ of the projection size.**

cross-validation. In this paper, instead, for the sake of simplicity and to reduce the computational time, we set these parameters on the single previous scenario that we use to analyze the parameter sensitivity. From this specific scenario, Figure 4 shows that localization performances are quite stable for $K \geqslant 512$ and $\mu \leqslant 10^{-1}$, the regularization parameter for the local linear regression is thus set to $\mu = 10^{-4}$ and the number of clusters to $K = 4096$. In the same way, Figure 5 shows that the error is quite satisfactory in the range $L \in [4, 128]$, the number of nearest neighbors for kernel regression is thus set to $L = 8$. From Figure 6, we conclude that the size of the projection space should be in the range $\delta \in [5, 20]$, we set it to $\delta = 10$.

## 5. Discussion and perspectives

The quantitative analysis of section 4 illustrates how much machine learning may be efficient with regards to source localization in fluctuating environments. However, their robustness to the uncertainties of the propagation medium has a counterparty: their precision and performance are directly linked to the representativeness of available training data. Machine learning could therefore be of interest for example within acoustic observatories, where data can be collected during long periods, making use of sources of opportunity.

Conversely, machine learning may not be a relevant approach in complex configurations when only few training data is acquired. In such a scenario, we encourage to fuse the knowledge we have from the acoustic propagation and the one that training data can provide. An inversion scheme with a physical acoustic model can actually benefit from the real few measures of the fluctuations. In this context, a fusing model, that integrates the decisions from both acoustic replica model and machine learning-based model, would be appropriated. More generally, in the case where there is not enough *in situ* training data, the training database we handle to train regression parameters can be extended by using synthetic samples from the acoustic replica model.

The clear need for an exhaustive training database is not the only one drawback we can identify by using a machine learning approach. Indeed, in underwater acoustic, detecting several signals at the same time in not straightforward. The method we have proposed here only support a single source. A conventional beamformer, or any matched field technique, bases its multi-source localizer from a threshold which is applied to the spectrum output. Following this idea, we can propose an inversion scheme by using a regression. This specific regression would predict the antenna measure from the source position. An other solution consists of registering a training database that considers a set of records in the presence of several sources.

Finally, we would emphasize that detecting a source position from under-acoustic measurements is not the only one task the underwater acoustician is interested in. Because inversion requires a replica model of the acoustic measure that depends on several environmental parameters, it would be interesting to assess these parameters from machine learning. For instance, the celerity profiles and the seabed properties may be assessed by machine learning. In the same way, the task of detecting the source presence/absence can also be dealt by machine learning. Especially given that the experimental training data acquisition seams easier in this case, indeed, we do not need to know the exact source position. Note that, unlike this paper which considers a monochromatic signal, in order to consider such new applications, we would have to use other acoustic signatures in order to model the hidden involved parameters. Time-frequency parameters would be on top of interest in such a case.

## 6. Conclusion

In this paper, we have addressed the task of source localization in fluctuating underwater environments from a machine learning point of view. In particular, two regression methods are confronted to two classical inversion approaches, namely a Matched Field Beamforming and the MUSIC algorithm. The data considered to train and test the regression approaches have been collected in tank conditions [15]. They constitute ideal study subjects for machine learning approaches: they reproduce fluctuating environments in closed and well-mastered settings. In a more general view, they give insight into the performance that should be achieved by machine learning methods within localization of underwater sources.

The quantitative analysis we carried out illustrates the potential of machine learning regarding fluctuating environments. More precisely, our experiments show that the source localization error is decreased by using machine learning. In this regard, their good behavior tends to underline their interest in more general settings. In particular, they do not rely on an explicit propagation model and reveal thus suitable to situations where no or too few *a priori* information is available on the environmental characteristics.

## Acknowledgment

This work has been supported by the DGA/MRIS.

## References

[1] H. Bucker, "Use of calculated sound fields and matched-field detection to locate sound sources in shallow water", The Journal of the Acoustical Society of America, volume 59(2), pages 368-373, 1976.

[2] H. L. Wilson and F.D. Tappert, "Acoustic propagation in random oceans using the radiation transport equation", The Journal of the Acoustical Society of America, volume 66(1), pages 256-274, 1979.

[3] P. Gerstoft, A. Xenaki and C.F. Mecklenbräuker, "Multiple and single snapshot compressive beamforming", The Journal of the Acoustical Society of America, volume 138(4), pages 2003-2014, 2015.

[4] X. Wang, B. Quost, J.-D. Chazot and J. Antoni, "Estimation of multiple sound sources with data and model uncertainties using the EM and evidential EM algorithms", Mechanical Systems and Signal Processing, volume 66-67, pages 159-177, 2016.

[5] C. Soares, M. Siderius and S.M. Jesus, "Source localization in a time-varying ocean waveguide", The Journal of the Acoustical Society of America, volume 112(5), pages 1879-1889, 2002.

[6] V. Tatarskii, "The effects of the turbulent atmosphere on wave propagation", Book, Jerusalem: Israel Program for Scientific Translations, 1971.

[7] Y. Jin and B. Friedlander, “Detection of distributed sources using sensor arrays”, IEEE Transaction on Signal Processing, volume 52(6), pages 1537-1548, 2004.

[8] G. Hinton, L. Deng, D. Yu, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T.S.G. Dahl and B. Kingsbury, “Deep Neural Networks for Acoustic Modeling in Speech Recognition”, volume 29(6), pages 82-97, 2012.

[9] M. Oquab, L. Bottou, I. Laptev and J. Sivic, “Learning and Transferring Mid-Level Image Representations using Convolutional Neural Networks”, International Conference on Computer Vision and Pattern Recognition, 2014.

[10] M.S. Datum, F. Palmieri and A. Moiseff, “An artificial neural network for sound localization using binaural cues”, The Journal of the Acoustical Society of America, volume 100(1), pages 372–383, 1996.

[11] A. Deleforge, R. Horaud, Y.Y. Schechner, L. Girin, “Co-Localization of Audio Sources in Images Using Binaural Features and Locally-Linear Regression”, IEEE Transactions on Audio, Speech and Language Processing, volume 23(4), pages 718-731, 2015.

[12] B. Clark, “The emerging era in undersea warfare”, report from the Center for Strategic and Budgetary Assessments, 2015. http://csbaonline.org/publications/2015/01/undersea-warfare/

[13] S. Pensieri, R. Bozzano, J.A. Nystuen, E.N. Anagnostou, M.N. Anagnostou and R. Bechini, “Underwater Acoustic Measurements to Estimate Wind and Rainfall in the Mediterranean Sea”, Advances in Meteorology, volume 15, pages 1-18, 2015.

[14] G. Real, J.-P. Sessarego, X. Cristol and D. Fattaccioli, “Decoherence effects in underwater acoustics: scaled experiments”, Underwater Acoustics Conference and Exhibition, 2014. https://www.researchgate.net/profile/Gaultier_Real/publications

[15] G. Real, X. Cristol, D. Habault and D. Fattaccioli, “Influence of de-coherence effects on sonar array gain: scaled

experiment, simulations and simplified theory comparison", Underwater Acoustics Conference & Exhibition, 2015. https://www.researchgate.net/profile/Gaultier_Real/publications

[16] G. Real, X. Cristol, D. Habault, J.-P. Sessarego and D. Fattaccioli, "RAFAL: RAndom Faced Acoustic Lens used to model internal waves effects on underwater acoustic propagation", Underwater Acoustics Conference & Exhibition, 2015. https://www.researchgate.net/profile/Gaultier_Real/publications

[17] G. Real, "An ultrasonic testbench for reproducing the degradation of sonar performance in a fluctuating ocean", PhD Thesis, University of Aix-Marseille, France, 2015. https://hal.inria.fr/tel-01239901/document

[18] S.M. Flatté, R. Dashen, W.H. Munk, K.M. Watson, F.Zachariasen, "Sound Transmission through a Fluctuating Ocean", Part of Cambridge Monographs on Mechanics, 2010.

[19] A. Baggeroer, W. Kuperman and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem", The Journal of the Acoustical Society of America, volume 83(2), pages 571-587, 1988.

[20] G. Bienvenu and L. Kopp, "Optimality of high resolution array processing using the eigensystem approach", IEEE Transaction on Acoustics, Speech and Signal Processing, volume 31(5), pages 1235-1248, 1983.

[21] R. Schmidt, "Multiple emitter location and signal parameter estimation", IEEE Transaction on Antennas and Propagation, volume 34(3), pages 276-280, 1986.

[22] G.R. Wilson, R.A. Koch and P.J. Vidmar, "Matched mode localization", The Journal of the Acoustical Society of America, volume 84, pages 310-320, 1988.

[23] T.C. Yang, "Effectiveness of mode filtering: A comparison of matched-field and matched-mode processing", The Journal of the Acoustical Society of America, volume 87, pages 2072-2084, 1990.

[24] A. Xenaki, P. Gerstoft, and K. Mosegaard, “Compressive beamforming”, The Journal of the Acoustical Society of America, volume 136(1), pages 260-271, 2014.

[25] A. Xenaki and P. Gerstoft, “Grid-free compressive beamforming”, The Journal of the Acoustical Society of America, volume 137(4), pages 1923-1935, 2015.

[26] S.E. Dosso, “Environmental uncertainty in ocean acoustic source localization”, Inverse Problems, volume 19(2), pages 419-431, 2003.

[27] S.E. Dosso, “Bayesian multiple-source localization in an uncertain ocean environment”, The Journal of the Acoustical Society of America, volume 129(6), pages 3577-3589, 2011.

[28] A. E. Hoerl, “Application of ridge analysis to regression problems", Chemical Engineering Progress, volume 1958, pages 54-59, 1962.

[29] A. Vedaldi and B. Fulkerson, “An Open and Portable Library of Computer Vision Algorithms”, http://www.vlfeat.org, 2008.

[30] E. A. Nadaraya, E. A., “On Estimating Regression”, Theory of Probability and its Applications, volume 9(1), pages 141-142, 1964.

[31] O. Kramer, “Unsupervised nearest neighbor regression for dimensionality reduction”, Soft Computing, volume 19(6), pages 1647-1661, 2015.

[32] X. Cristol, D. Fattaccioli and A.-S. Couvrat, “Alternative criteria for sonar array-gain limits from linear internal waves”, European Conference of Underwater Acoustics, 2012.

[33] J. Gaskill, “Linear systems, Fourier transforms and optics”, Wiley New York, 1978.

[34] A. Drémeau and C. Herzet, “An EM-algorithm approach for the design of orthonormal bases adapted to sparse representations”, IEEE Int’l Conference on Acoustics, Speech and Signal Processing (ICASSP), 2010.