Shallow water sound source localization using the iterative beamforming method in an image framework

Xun Wang, Shahram Khazaie, Luca Margheri, Pierre Sagaut

Shallow water sound source localization using the iterative beamforming method in an image framework

Xun Wang*, Shahram Khazaie, Luca Margheri, and Pierre Sagaut

Aix Marseille Univ, CNRS, Centrale Marseille, M2P2, 13451 Marseille Cedex 13, France

Abstract

Shallow water is a complicated sound propagation medium due to multiple reflections by water surface and bottom, imprecisely measured sound speed, noisy environment, etc. Therefore, in order to localize a shallow water sound source, classical signal processing techniques must be improved by taking these complexities into account. In this work, the multiple reflections and uncertain reflectivity of water bottom are explicitly modeled. In the proposed model, a measured signal is a mixture of the direct propagation from the source and the multiple reflections. Instead of solving the Helmholtz equation with boundary conditions of reflections, each signal is interpreted as a superposition of signals emitting from the physical source and its image sources in a free space, which results in a fast computation of sound propagation. Then, the source location, along with its amplitude, reflection paths and power loss of bottom reflection, is estimated via the iterative beamforming (IB) method, which alternatively estimates the source contributions and performs beamforming on these estimates until convergence. This approach does not need to compute the sound propagation for all the possible source locations in a large space, which thus leads to a low computational cost. Finally, numerical simulations are introduced to illustrate the advantage of the proposed model and the source estimation method. The sensitivity of the proposed method with respect to model parameter uncertainties is also investigated via a full uncertainty quantification analysis. The localization error of IB is proved to be acceptable in the given error range of sound speed and water depth. Besides, the IB source estimate is more sensitive to the sound speed while the matched-field processing methods have a stronger sensitivity to the water depth: this result can guide the choice of source localization method in the different cases of model parameter uncertainties.

*Electronic address: xun.wang@univ-amu.fr, xunwang00@gmail.com

Keywords: sound source localization, shallow water, iterative beamforming, multiple reflections, uncertainty quantification, acoustic imaging.

1. Introduction

Sound source localization in shallow water is a challenging issue due to the complexity of sound propagation environment and the difficulty of acoustical measurements. Contrary to the ideal assumption in the classical source localization problem, experimental results obtained in shallow water acoustics are always submitted to different kinds of uncertainties [1, 2]. Inhomogeneous and random sound speed profile [3–7], imprecise locations of sensors [8–10], sound reflections by ships, water surface and bottom [11, 12], noisy environment (due to shipping lanes for example) and presence of internal waves [11, 13, 14] are all sources of uncertainties which may dramatically affect the estimation result of sound source. Due to the complicated nature of the underwater environment, until now no signal processing method for source localization can be employed to eliminate all the bad effects from the aforementioned uncertainties. In this study, a model quantifying the multiple reflections between the water surface and bottom is proposed. The reflectivity of water bottom is assumed to be unknown.

By posing the boundary conditions at the water surface and bottom, the depth-dependent function of sound wave in a Pekeris waveguide can be analytically solved [2]. However, in order to compute the Helmholtz equation, numerical methods, e.g., wavenumber integration [15] or normal modes [16], have to be used. These approaches involve a low computational cost but is not negligible when a very large number of sound propagation has to be calculated, which is generally the case of source localization. Alternatively, this paper considers an image source approach [2, 11, 17, 18], in which the reflections between the two boundaries can be described by the images of the physical source. In this case, the boundary conditions are not needed and the sound propagation can be considered in a free-field. Furthermore, the unknown loss of bottom reflection, which depends on the water bottom properties, is also quantified in this image source model via the amplitudes of the image sources. Finally, by identifying the multiple sources (the physical source and its images) in the free-field, the sound source in the shallow water environment can be localized. The reflectivity of water bottom can also be decided via the estimated amplitudes of the sources.

Near-field acoustical holography (NAH), time reversal and beamforming are the most widely used signal processing techniques dedicated to the sound source localization problem. NAH [19] back-propagates the sound field from the measurement plane to a surface near the sound source, which guarantees a high resolution by taking evanescent waves into account. However, this approach only works for the near-field source. The time reversal method [20–24] inverses the measured signal in time and reinject it back into the same medium. This approach is able to refocus the source and return a super-resolution result in a medium with multiple reflections, scattering and refractions [23]. However, time reversal is sensitive to the model uncertainty: if the environmental parameters (e.g., sound speed profile) were imprecisely known, it could not even return a source estimate. Classical beamforming (CB) [25–27] estimates the direction of arrival of plane wave or the location of point source via the delay of signal arrival. However, CB is still limited in the frequency range and minimum resolvable source separation, particularly for the underwater case which has a relatively small size of microphone array and long sound propagation distance. Matched-field processing (MFP) [3–5, 28, 29] is a generalized beamforming method which takes the sound field complexity into account. Conventional and Capon's [30] approaches are the most widely used MFP methods, the former is more stable with respect to the model parameter uncertainties while the latter returns a super-resolution source localization. Furthermore, the MFP approaches can not only localize the sound source but also estimate the parameters related to the sound propagation environment. However, MFP needs to discretize the space of source location and other unknown parameters which often involves a large number of sound field computation. Iterative beamforming [10, 31] is a maximum likelihood (ML) approach for multiple sources, based on the expectation-maximization (EM) algorithm [32, 33]. This method assumes that each acoustical measurement is a sum of latent signals emitted from the multiple sources. Given an initial value of source parameters, IB iteratively computes the latent signals (source contributions) using the current fit of parameters and then performs beamforming on these source contributions to estimate the corresponding source location and amplitude. The theory of the EM algorithm guarantees the increase of likelihood function of model parameter after each iteration and the final result is a maximum likelihood estimate (MLE) of the sound sources [32]. This parametric mechanism of IB is able to avoid the limit of spatial resolution and augment the frequency range of CB. In this paper, the IB approach is used to solve the source localization

problem in the framework of image sources.

Since the underwater environment is complicated and some of the environmental parameters cannot be precisely measured, the sensitivity of the source localization method with respect to these parameters must be investigated. Kriging method [34, 35] is an efficient tool to build response surfaces for surrogate modeling, which is extensively used in uncertainty quantification (UQ) in mechanics [36, 37] and fluid dynamics [38–40]. In this work, a response surface of the estimated sound source as a function of uncertain parameters (water depth and sound velocity) is built using the Kriging method. Then, a quasi-Monte Carlo sampling of the response surface is used to obtain the probability density function (PDF) of the source estimate and to quantify the sensitivities of the source estimation methods with respect to the uncertain environmental parameters.

In this paper, the problem of sound source localization in a shallow water environment is investigated. The sound waves are emitted from a single source and reflected by the water surface and bottom. The sound propagation process is considered via the image source model. The single source localization problem is then transferred to a multiple-source estimation problem in a free-field, which is solved via the IB method. The organization of this paper is as follows. Section 2 introduces the proposed image source model. In Section 3, the source estimation methods are briefly introduced first; the algorithm for estimating a shallow water source using IB is then explicitly given. Section 4 presents experimental results on simulated data. In Section 5, the sensitivity of the source localization methods to the uncertain parameters of sound speed and water depth is investigated. Finally, the conclusions and perspectives are given in Section 6.

2. Sound propagation in a shallow water environment

2.1. Model description

In this section, the sound propagation model in a shallow water environment is presented. It is assumed that the acoustic field is produced by a sound source radiating continuous time signals at $\mathbf{r}_0 = (x_0, y_0, z_0)$. The sound wave propagates through the underwater medium and is measured by an array of sensors whose coordinates are $\mathbf{r}'_m = (x'_m, y'_m, z'_m), m = 1, \cdots, M$. Figure 1 shows the problem in 3D space, in which the cross and circles represent the source and microphone locations respectively. By assuming that the

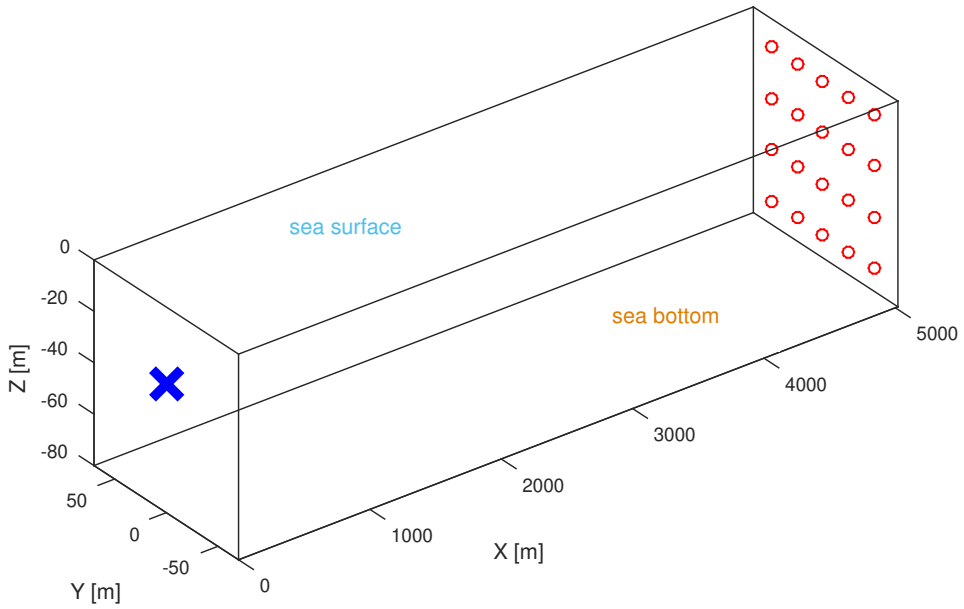


Figure 1: Description of the model in 3D space. The cross and circles represent the source and microphone locations, respectively.

medium is sourceless except at $\mathbf{r}_0$, the sound pressure $\tilde{p}(\mathbf{r}, t)$ as a function of spatial coordinate $\mathbf{r} = (x, y, z)$ and time t is governed by the wave equation

$$\left(\nabla^2 - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right)\tilde{p}(\mathbf{r}, t) = \delta(\mathbf{r} - \mathbf{r}_0)\tilde{W}(t), \tag{1}$$

in which v represents the sound speed, $\tilde{W}(t)$ is a deterministic function of source signal in the time domain, and δ is the Dirac delta function. Taking Fourier transform of the both sides of Eq. (1) with respect to t results in the following Helmholtz equation for the sound field $p(\mathbf{r}, f)$ in the frequency domain:

$$\left(\nabla^2 + k^2\right)p(\mathbf{r}, f) = \delta(\mathbf{r} - \mathbf{r}_0)W(f), \tag{2}$$

where $k = 2\pi f/v$ is the wavenumber, f is the frequency, and $W(f)$ is the Fourier transform of $\tilde{W}(t)$. The solution of the wave equation depends on the boundary condition. The water surface ($z = 0$) is assumed to be totally mirror reflective. The bottom ($z = d < 0$) is smooth and has a frequency-dependent reflectivity (denoted by $\alpha(f) \in [0, 1]$). Moreover, the sound waves propagating through the bottom are assumed to be fully absorbed by the

bottom, i.e., no sound wave reflects back to the considered shallow water region $\{\mathbf{r} : z \in [d, 0]\}$.

2.2. Image source model

In this work, the sound wave equation and source localization are considered in an image source framework. The canonical (Pekeris) shallow water acoustic waveguide is considered. The sound speed field is assumed as a constant. The wavelength is much smaller than the water depth and wave propagation range such that the ray theory and the image method can be applied. The sound emitted from the source is reflected by the water surface and bottom. An acoustical measurement obtained by each microphone is a superposition of contributions from the direct propagation and the reflections. Figure 2 demonstrates this multiple reflection problem: the solid lines display the direct propagation (P_0) and the other three paths of reflections (upper reflection P_1, lower reflection P_2 and upper lower reflections P_3).

Instead of solving the wave equation Eq. (2) with the boundary condition, the sound pressure at each point $\mathbf{r}$ is seen as a mixture of contributions propagating from the physical and image sources in a free-field:

$$\begin{aligned} p(\mathbf{r}, f) &= \sum_{s=0}^{\infty} (-1)^{n(s)} \alpha(f)^{n^-(s)} A_0 G(\mathbf{r}|\mathbf{r}_s, f) \\ &= \sum_{s=0}^{\infty} (-1)^{n(s)} \alpha(f)^{n^-(s)} A_0 \frac{\mathrm{e}^{\mathrm{j}k\|\mathbf{r}-\mathbf{r}_s\|}}{4\pi\|\mathbf{r}-\mathbf{r}_s\|}, \end{aligned} \tag{3}$$

in which $\mathbf{r}_s$ ($s \geq 1$) represents the locations of the image sources, A_0 is the amplitude of the physical source $\mathbf{r}_0$, $n(s)$ stands for the number of reflections of s-th mode ($(s+1)$-th term in Eq. (3)), and $n^-(s)$ is the corresponding number of bottom reflections. The coefficient $(-1)^{n(s)}$ corresponds to the sign change after each reflection due to the pressure-release boundary condition and $\alpha^{n^-(s)}$ means that the remaining power after each bottom reflection is $|\alpha A|^2$, where $\alpha \in [0, 1]$ is the reflectivity and A represents the amplitude of incident wave. Note that all the image sources are in the same line $\{\mathbf{r} : x = x_0, y = y_0\}$, as shown in Figure 2. The z-coordinates z_s of the image sources are obtained by successively imaging z_0 (symmetric with respect to the two

boundaries $z = 0$ and $z = d$):

$$z_{4i} = z_0 - 2id, \tag{4a}$$
$$z_{4i+1} = -z_0 - 2id, \tag{4b}$$
$$z_{4i+2} = -z_0 + 2(i+1)d, \tag{4c}$$
$$z_{4i+3} = z_0 + 2(i+1)d, \tag{4d}$$

in which $i \in \{0, 1, 2, \cdots\}$. Figure 2 shows the locations of the first four sources, i.e., $i = 0$ in Eq. (4). It is remarkable that the image source locations $\mathbf{r}_s$, $s \in \{1, 2, 3, \cdots\}$, do not depend on the observation point. Therefore, the sound pressure $p(\mathbf{r}, f)$ measured at any point $\mathbf{r}$ can be seen as being generated by the image sources. Under the image source sorting method in Eq. (4), the corresponding number of reflections and number of bottom reflections are $n(s) = [(s+1)/2]$ and $n^-(s) = [(s+2)/4]$, in which $[\cdot]$ stands for the floor function. Therefore, the amplitudes of the image sources are

$$A_s = (-1)^{n(s)}\alpha^{n^-(s)}A_0 = (-1)^{\left[\frac{s+1}{2}\right]}\alpha^{\left[\frac{s+2}{4}\right]}A_0,\ s = 1, 2, \cdots. \tag{5}$$

By considering the measurement noise and ignoring the terms larger than S (become negligible after several bottom reflections, depending on the desired precision and the reflectivity α), the measured sound pressure from m-th microphone is

$$p(\mathbf{r}'_m) = \sum_{s=0}^{S} A_s \frac{\mathrm{e}^{\mathrm{j}k\|\mathbf{r}'_m - \mathbf{r}_s\|}}{4\pi\|\mathbf{r}'_m - \mathbf{r}_s\|} + e_m, m = 1, \cdots, M, \tag{6}$$

in which the amplitudes A_s is obtained from Eq. (5) for $s \geq 1$ and the measurement error e_m is assumed to follow a 0-mean complex-valued Gaussian distribution [41].

In the framework of image method, the sound propagation can be analytically computed, such that the source localization can be rapidly performed, which is crucial in the source localization problem. In Section 3, acoustic imaging techniques are employed to visualize all the (physical and image) sound sources. Since the presence of the two boundaries are known, the real (physical) sound source can be identified from its images.

3. Sound source localization using iterative beamforming

In this section, a shallow water sound source is localized. This single source localization problem with multiple reflections is transferred to a mul-

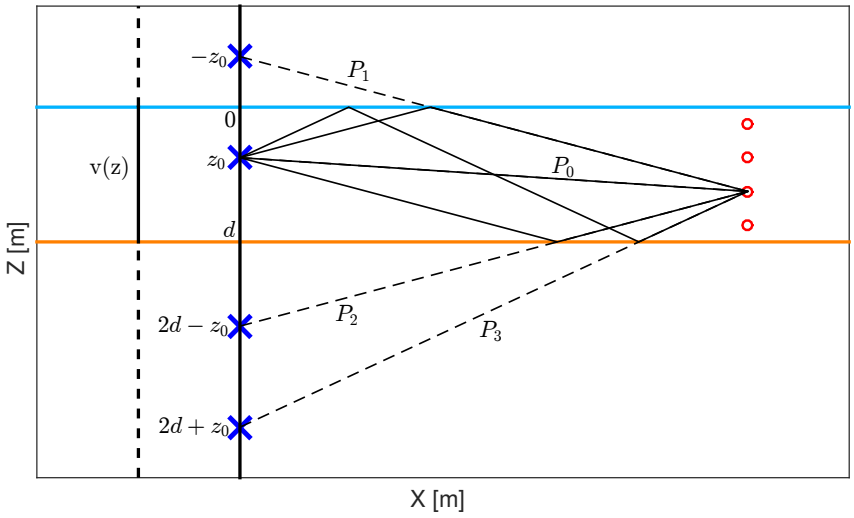


Figure 2: Sound propagation with multiple reflections on the water surface ($z = 0$) and bottom ($z = d$). The crosses and circles represent the (physical and image) sound sources and microphone locations, respectively. The solid and dashed lines stand for the physical and image sound propagation paths.

tiple source localization problem in a free-field using the image source approach, as introduced in Section 2.2. The data used are T snapshots of sound pressure measurements in the frequency domain, denoted as $\mathbf{p} = (\mathbf{p}_1, \cdots, \mathbf{p}_T)$. Each snapshot $\mathbf{p}_t = (p_{1t}, \cdots, p_{Mt})^{\mathrm{T}}$ includes the measurements from M microphones, in which $(\cdot)^{\mathrm{T}}$ stands for the operation of vector transpose.

In Sections 3.1 and 3.2, two sound source localization methods (CB and IB) are introduced. The detailed source localization strategy is given in Section 3.3. Finally, for comparison in terms of computational cost, the MFP methods for the single source based on the image source method are introduced in Section 3.4.

3.1. Classical beamforming

The CB method [25, 26] is based on the assumption of single source. In this case, each snapshot of the measurement $\mathbf{p}_t$ reads

$$\mathbf{p}_t = A_0 \mathbf{G}(\mathbf{r}_0) + \mathbf{e}_t, \tag{7}$$

in which $\mathbf{G}(\mathbf{r}_0) = (G(\mathbf{r}'_1|\mathbf{r}_0), \cdots, G(\mathbf{r}'_M|\mathbf{r}_0))^{\mathrm{T}}$ is the Green's function vector describing sound propagation from the source $\mathbf{r}_0$ to the M microphones and

being computed from the Helmholtz equation (2), A_0 is the amplitude of the sound source, and $\mathbf{e}_t = (e_{1t}, \cdots, e_{Mt})^{\mathrm{T}}$ is the vector of measurement noises which follows an M-dimensional complex Gaussian distribution $\mathbb{N}(\mathbf{0}, \sigma^2\mathbf{I}_M)$ and $\mathbf{I}_M$ is the M-dimensional identity matrix. Note that $\mathbf{p}_t$ is the sound pressure in the frequency domain, which is obtained from a linear combination of measurements in the time domain (discrete Fourier transform). Therefore, the measurements noise vector can be reasonably assumed as Gaussian distributed regardless of the distribution of the time domain noise according to the central limit theorem.

The CB approach localizes the sound source by "steering" the microphone array. The steering locations, which result in a maximum power, yield the beamforming estimate. "Steering the array" is performed by forming a linear combination of a measurement $\mathbf{p}_t$:

$$B_t = \mathbf{w}^{\mathrm{H}}\mathbf{p}_t = \sum_{m=1}^{M} w_m p_{mt}, \tag{8}$$

in which $(\cdot)^{\mathrm{H}}$ is the operation of conjugate transpose, $\mathbf{w} = (w_1, \cdots, w_M)^{\mathrm{T}}$ is called the steering vector. The CB method is to estimate the source location by maximizing $|B_t|^2$ with respect to $\mathbf{w}$, which finally reduces to maximizing the classical spatial spectrum

$$S(\mathbf{r}_0) = \frac{\mathbf{G}^{\mathrm{H}}(\mathbf{r}_0)\mathbf{p}_t\mathbf{p}_t^{\mathrm{H}}\mathbf{G}(\mathbf{r}_0)}{\mathbf{G}^{\mathrm{H}}(\mathbf{r}_0)\mathbf{G}(\mathbf{r}_0)}. \tag{9}$$

Note that the CB estimate (maximizing Eq. (9)) is a MLE of the source location $\mathbf{r}_0$.

In the case of multiple sources, CB solves the problem in the same way as the single source case. An $\mathbf{r}_0$-$S(\mathbf{r}_0)$ color map for all $\mathbf{r}_0$ in the considered region may be employed: each local maximum in this color map corresponds to a source. However, this is not a parametric approach and thus has a spatial resolution limit, i.e., the sources cannot be separated when they are close to each other. The side lobes of beamformer may also be wrongly identified as a source. In Section 3.2, a generalized beamforming method addressing the case of multiple sources is introduced, which considers the parametric multiple-source model and is a MLE of the multiple sources.

3.2. Iterative beamforming

Multiple sound source identification using the iterative beamforming (IB) approach is introduced in Ref. [10] for deterministic amplitude sources and

in Ref. [31] for random amplitude sources. IB is based on a ML approach to deal with latent variables, known as EM algorithm [32, 33]. In this paper, the shallow water source and its images are viewed as multiple sources, which can thus be estimated using the IB method. Here the principle of this method is briefly reviewed; for an extensive presentation, the readers are invited to refer to Ref. [10]. The detailed algorithm for shallow water source localization is introduced in Section 3.3.

In the case of $S+1$ sources, each snapshot of measurement $\mathbf{p}_t$ has the expression

$$\mathbf{p}_t = \mathbf{G}(\mathbf{r})\mathbf{A} + \mathbf{e}_t. \tag{10}$$

In this equation $\mathbf{r} = (\mathbf{r}_0, \cdots, \mathbf{r}_S)$ represents the locations of the sources, $\mathbf{G}(\mathbf{r}) = (G(\mathbf{r}'_m|\mathbf{r}_s))_{m=1,s=0}^{M,S}$ is the $M \times (S+1)$ Green's function matrix, in which $G(\mathbf{r}'_m|\mathbf{r}_s)$ describes the sound propagation from s-th source to m-th microphone, $\mathbf{A} = (A_0, \cdots, A_S)^{\mathrm{T}}$ is the vector of source amplitudes, and $\mathbf{e}_t = (e_{1t}, \cdots, e_{Mt})^{\mathrm{T}}$, $t = 1, \cdots, T$, are i.i.d. random vectors of measurement noises and follows a complex-valued Gaussian distribution $\mathbb{N}(\mathbf{0}, \sigma^2\mathbf{I}_M)$. The purpose of source identification is to estimate the source positions $\mathbf{r}$ and amplitudes $\mathbf{A}$. The log-likelihood function of the measurements $\mathbf{p}$ (after removing unnecessary terms which are independent of the unknown parameters and have no effect of the MLE) is

$$\log L(\mathbf{r}, \mathbf{A}|\mathbf{p}) = -\sum_{t=1}^{T} \|\mathbf{p}_t - \mathbf{G}(\mathbf{r})\mathbf{A}\|^2. \tag{11}$$

The MLE of $\mathbf{r}$ and $\mathbf{A}$ is obtained by maximizing Eq. (11). In the case of single source, i.e., $\mathbf{r} = \mathbf{r}_0$ and $\mathbf{A} = A_0$, this ML estimate is identical to the CB estimate (maximization of Eq. (9)).

Note that maximizing Eq. (11) is a $(4S+4)$-parameter optimization problem, which is difficult to solve when S is large. However, this maximization problem can be largely simplified by introducing the latent variable, which is defined as an unknown information that could let the parameter estimation straightforward, should it be available. In this multiple-source identification problem, the contributions of the various sources to the measured pressures could be introduced as the latent variables, denoted as $\mathbf{c}_t = (\mathbf{c}_{0t}, \cdots, \mathbf{c}_{St})$, in which

$$\mathbf{c}_{st} = \mathbf{G}_s(\mathbf{r}_s)A_s + \mathbf{e}_{st}, s = 0, \cdots, S. \tag{12}$$

In this equation, $\mathbf{G}_s(\mathbf{r}_s)$ is the $(s+1)$-th column of the matrix $\mathbf{G}(\mathbf{r})$ and $\mathbf{G}_s(\mathbf{r}_s)A_s$ stands for the contribution from s-th source, $\mathbf{e}_{st}$ $(s = 0, \cdots, S,$

$t = 1, \cdots, T$) are independent Gaussian random variables with mean $\mathbf{0}$ and covariance matrix $\Sigma_s = \frac{\sigma^2}{S+1}\mathbf{I}_M$. Therefore, the latent variables $\mathbf{c}_{st}$ and the original measurements $\mathbf{p}_t$ are related via $\mathbf{p}_t = \sum_{s=0}^{S} \mathbf{c}_{st}$ and the log-likelihood function of the latent variables (after removing unnecessary terms which are independent of the unknown parameters and have no effect of the MLE) is

$$\log L(\mathbf{r}, \mathbf{A}|\mathbf{c}) = -\sum_{t=1}^{T}\sum_{s=0}^{S} \|\mathbf{c}_{st} - \mathbf{G}_s(\mathbf{r}_s)A_s\|^2. \tag{13}$$

Since the contributions $\mathbf{c}_{st}$ cannot be measured, it is impossible to directly maximize Eq. (13) with respect to each $\mathbf{r}_s$ and A_s to separately estimate each source from the others using the corresponding source contribution. However, the EM algorithm [32] makes it possible to proceed with Eq. (13) by treating the missing data as random variables. The IB approach starts from an initial parameter $\Phi^0 = (\mathbf{r}^0, \mathbf{A}^0)$; given l-th parameter estimates, the $(l+1)$-th iteration consists in the following steps [10]:

- compute the expected source contribution

$$\hat{\mathbf{c}}_{st}^l = \mathbb{E}(\mathbf{c}_{st}|\mathbf{p}_t, \mathbf{r}^l, \mathbf{A}^l) = \mathbf{G}_s^l A_s^l + \frac{1}{S+1}\left(\mathbf{p}_t - \sum_{s=0}^{S} \mathbf{G}_s^l A_s^l\right); \tag{14}$$

- decide the source location by performing a beamforming projection using the estimated source contribution:

$$\mathbf{r}_s^{l+1} = \arg\max_{\mathbf{r}_s} \frac{\mathbf{G}_s^{\mathrm{H}}(\mathbf{r}_s)\hat{\mathbf{c}}_s^l(\hat{\mathbf{c}}_s^l)^{\mathrm{H}}\mathbf{G}_s(\mathbf{r}_s)}{\mathbf{G}_s^{\mathrm{H}}(\mathbf{r}_s)\mathbf{G}_s(\mathbf{r}_s)}, \tag{15}$$

  in which $\hat{\mathbf{c}}_s^l = \frac{1}{T}\sum_{t=1}^{T} \hat{\mathbf{c}}_{st}^l$;

- estimate the source contribution via

$$A_s^{l+1} = \frac{\mathbf{G}_s^{\mathrm{H}}(\mathbf{r}_s^{l+1})\hat{\mathbf{c}}_s^l}{\mathbf{G}_s^{\mathrm{H}}(\mathbf{r}_s^{l+1})\mathbf{G}_s(\mathbf{r}_s^{l+1})}. \tag{16}$$

It is important to remark that the IB estimate is a MLE of the multiple-source parameters. The principle of EM algorithm guarantees the increase of

likelihood function after each iteration [32] and its convergence to a stationary point [33]. In order to obtain the estimate at the global maximum, a multiple initialization strategy could be employed: the above algorithm is performed with different initial parameters, and finally retain the estimate with highest likelihood.

3.3. Shallow water source localization using iterative beamforming

In this section, the sound source localization in a shallow water environment is considered. As introduced in Section 2.2, the problem of single source with multiple reflections is considered in a model of multiple sources in a free-field. In this case, the locations of the sound sources are denoted by $\mathbf{r} = (\mathbf{r}_0, \cdots, \mathbf{r}_S)$, in which $\mathbf{r}_0 = (x_0, y_0, z_0)$ is the coordinate of the physical source. The image source coordinates are $\mathbf{r}_s = (x_s, y_s, z_s)$, $s \geq 1$, in which $x_s = x_0$, $y_s = y_0$, and z_s is obtained from Eq. (4). The amplitudes of the sources are represented by $\mathbf{A} = (A_0, \cdots, A_S)^{\mathrm{T}}$, in which A_0 is the amplitude of the physical source, and the amplitudes A_s ($s \geq 1$) of the image sources are obtained via Eq. (5).

Given an initial parameter of the physical source parameters, denoted as $\mathbf{r}_0^0 = (x_0^0, y_0^0, z_0^0)$ and A_0^0, the corresponding initial location of each image sources is $\mathbf{r}_s^0 = (x_s^0, y_s^0, z_s^0)$, in which $x_s^0 = x_0^0$, $y_s^0 = y_0^0$, and z_s^0 is obtained by inserting z_0^0 back into Eq. (4). The initial amplitudes of the image sources are obtained by replacing A_0 in Eq. (5) by A_0^0 and giving an initial value of bottom reflectivity α^0. Then, the source localization algorithm presented in Section 3.2 is proceeded with the initial parameters $\mathbf{r}^0 = (\mathbf{r}_0^0, \cdots, \mathbf{r}_S^0)$ and $\mathbf{A}^0 = (A_0^0, \cdots, A_S^0)^{\mathrm{T}}$. The returned source estimates are represented by $\mathbf{r}^* = (\mathbf{r}_0^*, \cdots, \mathbf{r}_S^*)$ and $\mathbf{A}^* = (A_0^*, \cdots, A_S^*)$. Then, each estimated image source $\mathbf{r}_s^* = (x_s^*, y_s^*, z_s^*)$, $s \geq 1$, is reflected back to the region $z = [d, 0]$ to obtain the corresponding physical source estimate, denoted by $\mathbf{r}_s^{**} = (x_s^*, y_s^*, z_s^{**})$, in which

$$z_s^{**} = z_s^* + 2\left[\frac{s}{4}\right] d, \text{ if } \left\{\frac{s}{4}\right\} = 0 \tag{17a}$$

$$z_s^{**} = -z_s^* - 2\left[\frac{s}{4}\right] d, \text{ if } \left\{\frac{s}{4}\right\} = \frac{1}{4} \tag{17b}$$

$$z_s^{**} = -z_s^* + 2\left(\left[\frac{s}{4}\right] + 1\right) d, \text{ if } \left\{\frac{s}{4}\right\} = \frac{1}{2} \tag{17c}$$

$$z_s^{**} = z_s^* - 2\left(\left[\frac{s}{4}\right] + 1\right) d, \text{ if } \left\{\frac{s}{4}\right\} = \frac{3}{4}. \tag{17d}$$

Here, $\{x\} = x - [x]$ is the sawtooth function. Then, the estimate of the physical source location can be obtained by

$$\hat{\mathbf{r}}_0 = \frac{1}{S+1}\left(\mathbf{r}_0^* + \sum_{s=1}^{S} \mathbf{r}_s^{**}\right). \tag{18}$$

The reflectivity could be also estimated by

$$\hat{\alpha} = \frac{1}{S-1}\sum_{s=2}^{S}\left(\frac{|A_s^*|}{|A_0^*|}\right)^{1/\left[\frac{s+2}{4}\right]}. \tag{19}$$

Finally, the amplitude of the physical source can be estimated by an average of the image sources considering the power losses of bottom reflections:

$$\hat{A}_0 = \frac{1}{S+1}\sum_{s=0}^{S}(-1)^{\left[\frac{S+1}{2}\right]}\hat{\alpha}^{-\left[\frac{S+2}{4}\right]}A_s^*. \tag{20}$$

The IB strategy for localizing the shallow water source is summarized in Algorithm 1.

Algorithm 1 Sound source localization in shallow water using IB

1. For $l = 0$, pick starting values for the source parameters ($\mathbf{r}_0^0$ and A_0^0) and generate the initial parameters of image sources ($\mathbf{r}_s^0$ and A_s^0, $s = 0, \cdots, S$) via Eqs. (4) and (5).
2. For $l \geq 1$:

   repeat

   estimate the source contributions $\hat{\mathbf{c}}_{st}^l$ via Eq. (14) for $s = 0, \cdots, S$;
   estimate the source locations $\mathbf{r}_s^{l+1}$ via Eq. (15) for $s = 0, \cdots, S$;
   estimate the source amplitudes A_s^{l+1} via Eq. (16) for $s = 0, \cdots, S$.

   until the relative increase of the measured data log-likelihood Eq. (11) is less than a given threshold κ.

   return location and amplitude estimates of the physical and image sources ($\mathbf{r}_s^*$ and A_s^*, $s = 0, \cdots, S$).
3. The location and amplitude of the physical source and the reflectivity of water bottom are estimated via Eqs. (18), (20) and (19), respectively.

3.4. Matched-field processing and computational costs

In the underwater source localization problem, MFP [28] is probably the most widely used method, which generalizes the beamforming approach to estimate both source location and other unknown environmental parameters. In this section, conventional and Capon's MFP methods are briefly introduced, their computational costs are quantified and compared with the IB method.

In this work, the MFP methods are based on the image source model. The Green's function at m-th microphone $\mathbf{r}'_m$ is

$$G(\mathbf{r}'_m|\mathbf{r}_0,\alpha)=\sum_{s=0}^{S}(-1)^{n(s)}\alpha(f)^{n^-(s)}\frac{\mathrm{e}^{\mathrm{j}k\|\mathbf{r}'_m-\mathbf{r}_s\|}}{4\pi\|\mathbf{r}'_m-\mathbf{r}_s\|}. \tag{21}$$

The conventional and Capon's MFP estimates for the source location $\mathbf{r}_0$ and the reflectivity α are respectively:

$$\{\hat{\mathbf{r}}_0,\hat{\alpha}\}=\arg\max_{\mathbf{r}_0,\alpha}\frac{\mathbf{p}_t^H\mathbf{G}(\mathbf{r}_0,\alpha)\mathbf{G}^H(\mathbf{r}_0,\alpha)\mathbf{p}_t}{|\mathbf{G}(\mathbf{r}_0,\alpha)|^2} \tag{22}$$

and

$$\{\hat{\mathbf{r}}_0,\hat{\alpha}\}=\arg\max_{\mathbf{r}_0,\alpha}\frac{1}{\mathbf{G}^H(\mathbf{r}_0,\alpha)\hat{\mathbf{K}}^{-1}\mathbf{G}(\mathbf{r}_0,\alpha)}, \tag{23}$$

in which $\mathbf{G}(\mathbf{r}_0,\alpha)=(G(\mathbf{r}'_1|\mathbf{r}_0,\alpha),\cdots,G(\mathbf{r}'_M|\mathbf{r}_0,\alpha))$.

Then, the computational costs of the MFP approaches are estimated. The 3D space and $\alpha\in[0,1]$ are discretized; N_x, N_y, N_z and N_α denote the number of discrete points in x-, y-, z-axis and $\alpha\in[0,1]$, respectively. In order to obtain the MFP estimate, the sound propagation has to be calculated $SN_xN_yN_zN_\alpha$ times, such that the computational cost is very high when the possible region of source is large and high estimation precision is requested.

By contrast, the computational cost of IB is relatively low, which mainly comes from the optimization step. For this approach, the sound propagation process needs to be calculated $\sum_{l=1}^{L}\sum_{s=1}^{S}N_{sl}$ times, in which N_{sl} is the number of calculation of Green's function in the optimization step for s-th source and l-th iteration, depending on the optimization method. Therefore, the computational cost ratio between IB and MFP is approximately $N_xN_yN_zN_\alpha/(L\bar{N}_{sl})$, where $\bar{N}_{sl}$ is the average of N_{sl} and in general is less than 100. In most cases, IB requires much less computational cost than the MFP approaches. Besides, IB is not limited to the estimation precision while high estimation accuracy of MFP results in a high computation cost.

4. Numerical experiments

In this section, numerical experiments are introduced to illustrate the proposed model and source estimation method. The experimental setup is displayed in Figure 1. The Pekeris shallow water waveguide is assumed with a sound speed $v = 1500$ m/s. The water depth is 100 m, i.e., $d = -100$ m. The sound source is located at $\mathbf{r}_0 = (0, 0, -32)$ m and its amplitude is $A_0 = 1$ at the frequency $f = 1500$ Hz. Note that in order to apply the image model, the wavelength must be much smaller than the water depth. Here, the wavelength is $\lambda = 1$ m, which is sufficiently small to guarantee the model accuracy. Besides, the source localization accuracy depends on the frequency. Refs. [10, 31] have discussed the influence of frequency on IB and MFP, therefore it is not the main concern of the present paper. In this experiment, five towed hydrophone arrays are used, each array has 7 sensors. More specifically, $M = 35$ sensors are placed on an x-plane $x'_m = 5000$ m; their y- and z-coordinates are $y'_m \in \{-100, -50, 0, 50, 100\}$ m and $z'_m \in \{-95, -80, \cdots, -5\}$ m, respectively. The water surface is totally mirror reflective; the water bottom is smooth and has a reflectivity $\alpha = 0.1$. Generally speaking, an interference having an incident angle smaller than 75° can be ignored due to more bottom reflections (depending also on the material of water bottom and the microphone-source distance) [1]; in the data generation of this experiment, 34 modes (i.e., $S = 33$) are considered such that minimum incident angle is around 75°. Thus, the sound waves propagating from the source are within a cone whose apex angle is 30°. The sound propagation and measurement process are simulated via Eq. (3). The number of snapshots is $T = 30$. The signal-to-noise ratio (SNR) is assumed to be 10^4, the standard deviation of measurement noise is thus $\sigma = 10^{-4}\frac{1}{M}\|\mathbf{GA}\|$.

First, CB is employed to estimate the sound source. By inserting the mean of T measured pressures in the frequency domain into Eq. (9) and plotting the $\mathbf{r}_0$-$S(\mathbf{r}_0)$ color map for each $\mathbf{r}_0$ on the source plane $x = 0$, the CB sound field (sound pressure level whose reference pressure is the maximum value in the plane) of source plane can be constructed, as is shown in Figure 3 (a). It is clear that CB has a problem of spatial resolution, i.e., the physical and the first image source cannot be separated. Actually, the resolution of CB can be theoretically estimated [27] by:

$$\text{Res} = \frac{|x'_m - x_0|\lambda}{D}, \tag{24}$$

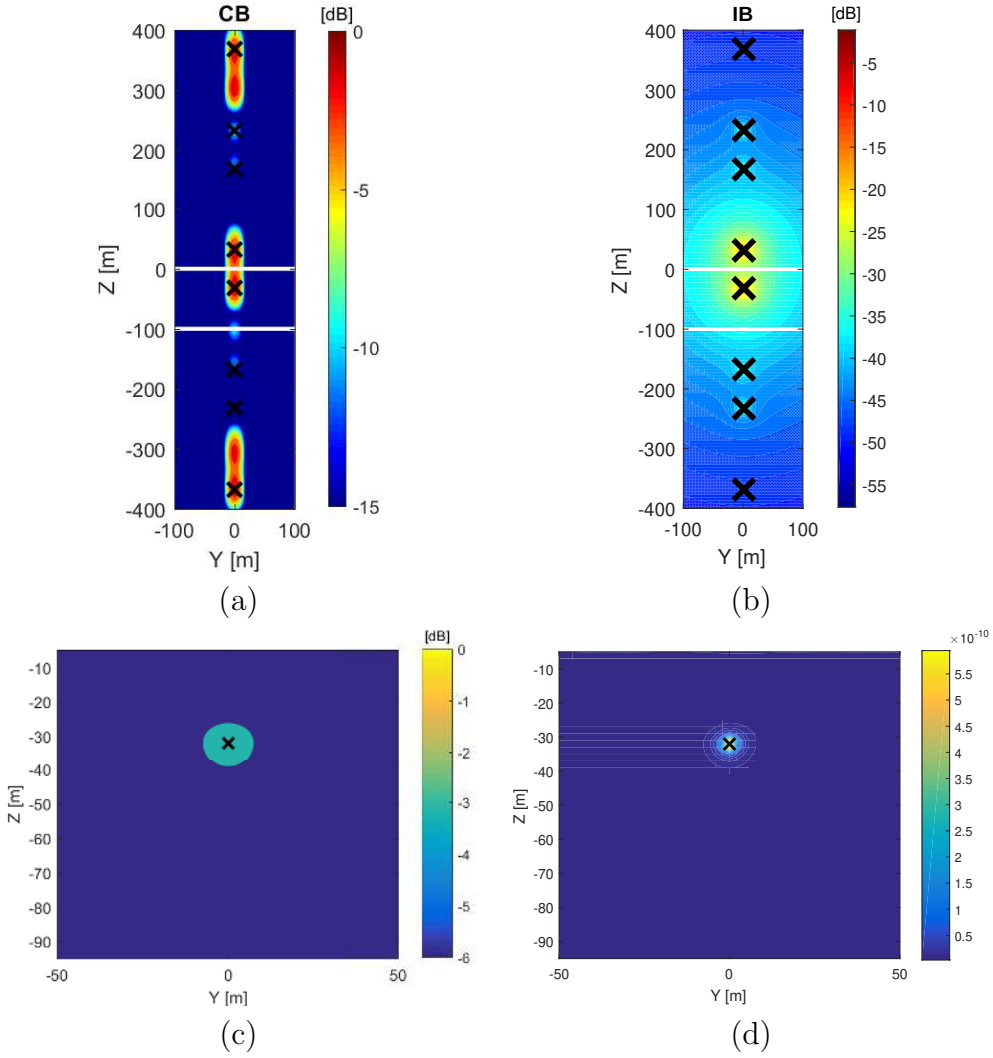


Figure 3: Sound source localization using (a) classical beamforming, (b) iterative beamforling, (c) conventional MFP, and (d) Capon's MFP. The distance between the source and the sensor plane is 5 km . The frequency of the sound source is 1500 Hz. In subfigures (a) and (b), the crosses stand for the physical and image source locations, and the two lines represent the water surface and bottom. (For interpretation of the references to color in this figure, the reader is referred to the web version of this paper.)

in which D is the size of the hydrophone array and λ is the wavelength. In this experiment, the size of the microphone array is 200 m in the y-direction and 90 m in the z-direction, thus the theoretical spatial resolution is 25 m $\times$

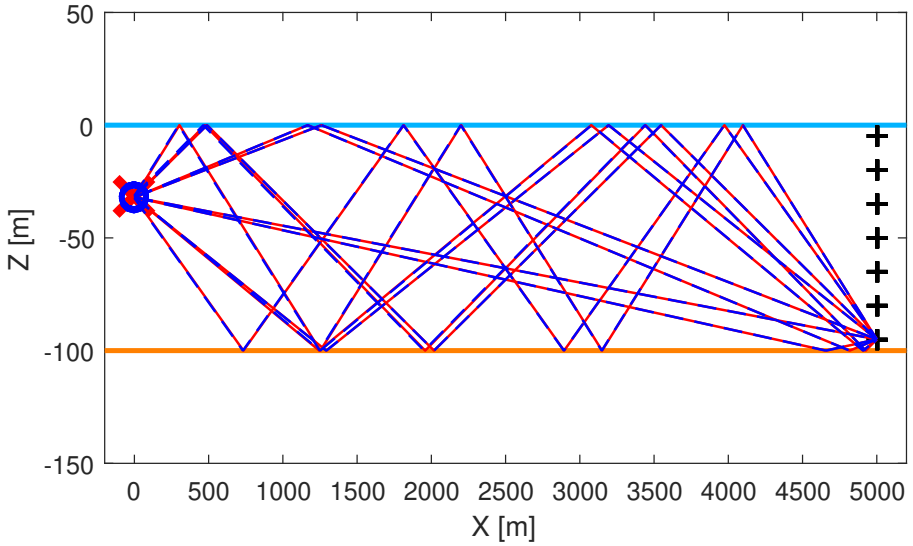


Figure 4: Actual (solid lines) and estimated (dash lines) paths of sound propagation from the source to each microphone. The cross and the circles stand for the actual source and the estimates of the image sources reflected back to the physical region.

56 m, which is consistent with the numerical result.

However, the above limitation can be overcome by IB [31]. Since the IB approach is parametric (the estimates of point sources are explicitly obtained), the reconstructed sound field is not limited by the minimum resolvable source separation of CB. Here, in order to reduce the computational cost, the number of modes assumed in the IB estimation process is 10, i.e., only the sources with powers larger than $|\alpha^3 A_0|^2$ are considered. Ten initial source location of $\mathbf{r}_0$ is randomly obtained, following a 3D uniform distribution with support $\{(x, y, z) : x \in [-50, 50], y \in [-50, 50], z \in [-60, 0]\}$. The corresponding initial source amplitudes are randomly generated from a uniform random variable in $[0.8, 1.2]$. Then, the IB estimates are obtained using the method introduced in Section 3.3: the estimate with highest likelihood is retained as the final source estimate. Figure 3 (b) shows the reconstructed sound pressure level on the source plane: for each discrete point $\mathbf{r}$ on the source plane $z = 0$ ($\mathbf{r} \neq \mathbf{r}_s^*$), the reconstructed sound pressure is computed by

$$\mathbf{p}^*(\mathbf{r}) = \sum_{s=0}^{S} A_s^* |G(\mathbf{r}|\mathbf{r}_s^*)|. \tag{25}$$

This figure shows that IB avoids the limitation of spatial resolution of CB and can clearly separate all the sources. Besides, the source amplitude and the bottom reflectivity are estimated via Eqs. (20) and (19), being $\hat{A}_0 = 0.99$ (the actual value is 1) and $\hat{\alpha} = 0.107$ (the actual value is 0.1). Figure 4 displays the source estimates projected back to the physical region $z \in [d, 0]$ on the X-Z plane via Eq. (17), denoted by the circles. These estimated physical sources are almost overlapped with the actual one, which illustrates the precision of the IB estimates. The sound propagation paths from the real source (solid lines) and the estimated sources (dashed lines, each of which is emitted from the corresponding projected image source) to each hydrophone are also shown in Figure 4. It is clear that the sound propagation paths can also be accurately identified by IB. Furthermore, the IB algorithm converges fast such than the computational cost is low. Actually, the number L of iteration steps is always less than 100 (and maximum 1000) while the threshold κ of relative likelihood function increment is 10^{-4}.

Finally, the conventional and Capon's MFP methods are used. The 3D coordinate of source and the reflectivity of water bottom are discretized by $\{(x, y, z) : x = -50 : 0.1 : 50, y = -50 : 0.1 : 50, z = -5 : -0.1 : -95\}$ m and $\alpha = 0 : 0.01 : 0.5$, respectively. The source estimates are obtained from Eqs. (22) and (23) and the reconstructed sound fields in the source plane $x = 0$ are shown in Figure 3 (c) and (d). Besides, the estimated reflectivity obtained from the both approaches are respectively $\hat{\alpha} = 0.1$ and $\hat{\alpha} = 0.09$. It is clear that the both MFP methods are also accurate in terms of source localization, although conventional MFP has a relatively large focal spot size. However, it is remarkable that the MFP methods require a much higher computational cost: in this example, the sound propagation process has to be computed 4.6×10^{11} times. By contrast, the IB approach needs to calculate the sound propagation maximum 10^6 times (with 10 initial values). Therefore, the computation speed of IB is approximately 4.6×10^5 times faster than the MFP approaches.

5. Sensitivity to environmental parameter uncertainties

In the above experiments, the sound speed and water depth are assumed to be precisely known in the source estimation process. In real applications, however, these environmental parameters often cannot be accurately measured. For example, Ref. [3] proposes that the error ranges of shallow water sound speed and water depth are ± 2.5 m/s and ± 2.5 m. In this section, an

UQ analysis is presented. The source estimation method is proceeded with inaccurate information regarding the sound speed and water depth. Kriging is used to build a response surface (i.e., a meta-model of the true simulation output) of the predicted acoustic source localization. A quasi-Monte Carlo sampling of the response surface is then used to obtain the PDF of the source estimate, the sensitivities of the environmental parameters are also discussed.

5.1. Building a Kriging response surface

The Kriging method proposes a weighted linear estimator, i.e., given n values of the generic quantity of interest (QoI) $X(\omega)$ sampled over the uncertain space Ω, the estimate of the QoI $\hat{X}(\omega)$ at an unmeasured location ω is:

$$\hat{X}(\omega) = \sum_{s=1}^{n} \lambda_s(\omega) X(\omega_s), \tag{26}$$

where $\lambda_s(\omega)$ is the Kriging weight assigned to sample $X(\omega_s)$ and ω_s is the sample location. It is important to notice that the Kriging weights are functions of location at which an estimation is needed, since the s-th weight depends on the distance between the unmeasured sample and the s-th sample location. The QoI $X(\omega)$ is arbitrarily decomposed into a residual $R(\omega)$ and a trend component $\beta(\omega)$:

$$X(\omega) = R(\omega) + \beta(\omega). \tag{27}$$

The residual component can be treated as a stationary random field with 0-mean and covariance $\mathrm{Cov}(R(\omega), R(\omega + h)) = C(h)$, where $C(h)$ is called the covariance function. The fundamental step of Kriging is to correctly select the covariance function defining the underlying Gaussian process. Here, the spline covariance function is used to build the response surface, i.e.,

$$C(h) = \begin{cases} 1 - 6(h\theta)^2 + 6(h\theta)^3, & h < \frac{1}{2\theta} \\ 2(1 - h\theta)^3, & \frac{1}{2\theta} \leq h < \frac{1}{\theta} \\ 0, & h \geq \frac{1}{\theta} \end{cases}, \tag{28}$$

in which θ can be obtained via MLE or empirical tuning.

The Kriging weights $\lambda_s(\omega)$ are determined by minimizing the variance of the estimator $\mathrm{Var}(\hat{X}(\omega) - X(\omega))$ under the unbiasedness constraint $\mathbb{E}[\hat{X}(\omega) - X(\omega)] = 0$. As a matter of fact, Kriging estimator is a best linear unbiased predictor (BLUP). A number of variants of Kriging exists in literature,

among which the Ordinary Kriging approach is preferred, given that there is no a-priori assumption on trend and given the simplicity and robustness of Ordinary Kriging compared to other variants. In this approach, the trend part of the random field $X(\omega)$ is assumed to be constant but with unknown $\beta(\omega) = \mu_0$. Unbiasedness condition requires that the sum of Kriging weights is equal to 1. Then, by minimizing the error variance and adding a Lagrange parameter to respect the unbiasedness constraint, the well known Ordinary Kriging system is obtained:

$$\begin{cases} \sum_{j=1}^{n} \lambda_j(\omega) C(\omega_i - \omega_j) + \mu_0 = C(\omega_i - \omega) \quad i = 1, \cdots, n \\ \sum_{j=1}^{n} \lambda_j(\omega) = 1 \end{cases} \tag{29}$$

To assess the reliability of the meta-model, a leave-one-out cross validation strategy [42] is adopted. The cross-validation values are expressed here as a global relative L_2-error:

$$CV_s = \frac{\sqrt{\int_\Omega (\hat{X} - \hat{X}^{-s})^2 \mathrm{d}\omega}}{\sqrt{\int_\Omega \hat{X}^2 \mathrm{d}\omega}} \tag{30}$$

where $\hat{X}^{-s}$ is the Kriging response surface built without the s-th sample. A threshold ε on the computed cross-validation value CV_s for each sample is imposed to assess convergence.

5.2. Sensitivity estimators from the Kriging meta-model

The main statistics and the Sobol' variance-based sensitivity indices [43] are obtained through quasi Monte Carlo (Sobol sequences) estimators [44]. The mean estimate $\mu(X)$ is calculated as

$$\mu(X) = \frac{1}{q_{MC}} \sum_{i=1}^{q_{MC}} \hat{X}_i, \tag{31}$$

where q_{MC} is the number of quasi Monte Carlo samples $\hat{X}_i$ of the Kriging meta-model of $X(\Omega)$. Similarly, the variance $\sigma^2(X)$ for large q_{MC} can be obtained by

$$\sigma^2(X) = \frac{1}{q_{MC}} \sum_{i=1}^{q_{MC}} (\hat{X}_i - \mu(X))^2. \tag{32}$$

Finally, for the first order sensitivity indices, a common approach is to generate two independent quasi Monte Carlo sample sets of Ω. The independent sets Ω_1 and Ω_2 are obtained from a $q_{MC} \times 2D$ matrix, which is a quasi Monte Carlo sample series in dimension $2D$, D is the dimension of the original problem. The Sobol' indices $S_j = \sigma_j^2/\sigma^2$ for j-th dimension is derived from the estimate of the variance and the partial variance σ_j^2:

$$\sigma_j^2 = \frac{1}{qMC} \sum_{i=1}^{qMC} \hat{X}_i(\Omega_2)(\hat{X}_i(\Omega_1^j) - \hat{X}_i(\Omega_1)), \tag{33}$$

where Ω_1^j is the first samples set Ω_1 where the j-th column has been replaced by the corresponding column of the second sample set Ω_2. The residual variance can be estimated by the difference between the total variance and the sum of first order partial variances.

5.3. Uncertainty quantification of the source localization to the environmental parameters

In this section, the UQ analysis of source localization with respect to uncertain environmental parameters is presented. As previously discussed, the shallow water sound speed and the water depth are assumed to be uncertain, the ranges of which are set to be $v = 1500 \pm 2.5$ m/s and $d = -100 \pm 2.5$ m, respectively. Here, both parameters are assumed as independent random variables. The UQ analysis of source localization using IB, conventional and Capon's MFP is done. The sound source is located at $(0, 0, -32)$ m and the frequency is 1500 Hz. The sensor locations are the same as in Section 4. In order to decrease the computational costs, the x-coordinate of sound source is assumed to be known for all the the three methods, thus only the y- and z-coordinates have to be estimated. Note that in this case the computational cost of MFP decrease (the region of possible source location becomes two-dimensional), but is still much heavier than IB: the computation speed of IB is approximately 460 times faster than the MFP approaches. Besides, in order to eliminate the influence of random initialization of IB, the starting values for the source parameters are obtained from a grid $\{(y, z) : y = -100, -80, \cdots, 100, z = 0, -10, \cdots, -70\}$.

A regular grid with 25 samples of uncertain parameters ($v = 1500, 1500 \pm 1.25, 1500 \pm 2.5$ m/s, $d = 100, 100 \pm 1.25, 1000 \pm 2.5$ m) is used to assess the reliability of the Kriging-based response surface. The number of samples of the UQ grid depends in general on the dimensionality of the UQ problem, on

the physical ranges of each UQ parameter and on the expected complexity of the response. In this case, a grid with 25 samples proved to be enough to get a reasonable convergence of the Kriging response surface. Further tests with additional samples are proved to be inefficient in terms of response surface reliability.

Figure 5 shows the computation results of IB with four different assumed parameters: (a) $v = 1500$ m/s, $d = -98.75$ m; (b) $v = 1500$ m/s, $d = -97.5$ m; (c) $v = 1501.25$ m/s, $d = -100$ m; (b) $v = 1502.5$ m/s, $d = -100$ m. In the latter two cases, the image sources with low strengths cannot be accurately estimated. However, the first two sources (the physical source and its first image which have the highest amplitudes) can be identified in all the cases. Since the main purpose of this work is to localize the physical source, these results are sufficient: the estimated source between the two boundaries can be identified as the estimate of the physical source. Here, the source location estimate of IB is the first mode of the multiple-source estimate. In the following, the estimation accuracy of all the three methods is quantified.

A response surface is built for the z-coordinate estimate of source through ordinary Kriging with spline covariance function and $\theta = 0.1$ (tuned to well fit the meta-model). Actually, the experimental results of IB show that the estimation error of y-coordinate is always much smaller than z-coordinate and relatively negligible (smaller than 0.1 m), thus the z-coordinate estimation error $|z_0 - z_0^*|$ can be seen as the source estimation error $\|\mathbf{r}_0 - \mathbf{r}_0^*\|$. Although the main interest of the UQ analysis is to quantify the error on the source localization, using it to build the response surface is not so efficient as it may appear, since the absolute value operator results in a response surface with sharp points, which can not be correctly fitted by a Kriging meta-model with a small sample size. Therefore, the response surface of the z-coordinate estimate is constructed and the error is then derived. The interested reader can see Ref. [40] for a discussion on the importance of the choice of the QoI in building the response surface and assessing its reliability. Furthermore, the leave-one-out cross validation is used to assess the reliability of the Kriging-based response surface, where the cross-validation value is $CV_s < 1\%$ for all s.

At first both uncertain parameters are assumed to be uniformly distributed; then the results are compared to the truncated Gaussian distributions, which let the range of each parameter mentioned above be equal to six standard deviation centered at the mean value. The PDFs, along with the mean value $\mu(z)$ and the confidence interval $\mu(z) \pm \sigma(z)$ ($\sigma(z)$ is the

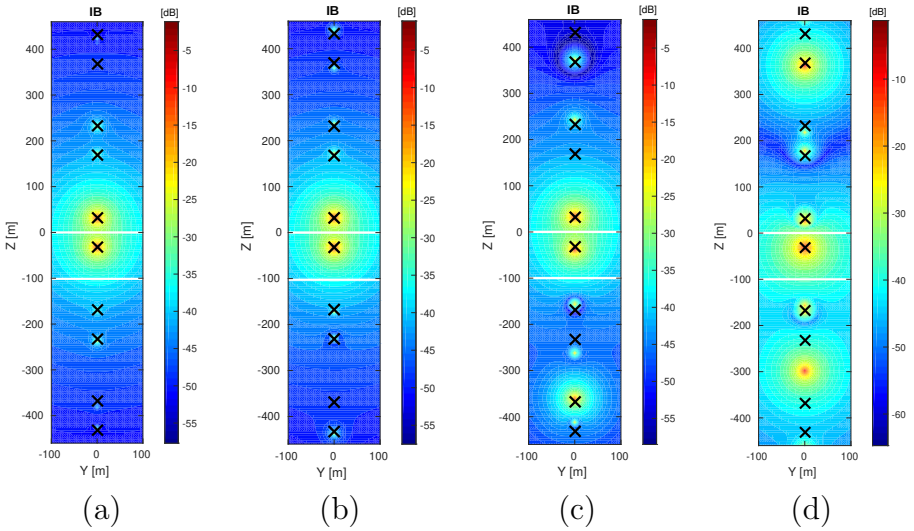


Figure 5: Reconstructed sound pressure levels using IB with imprecise information of water depth and velocity: (a) $d = -98.75$ m, $v = 1500$ m/s; (b) $d = -97.5$ m, $v = 1500$ m/s; (c) $d = -100$ m, $v = 1501.25$ m/s; (d) $d = -100$ m, $v = 1502.5$ m/s. The black crosses stand for the locations of (physical and image) sound sources. The white lines indicate the water surface and bottom. The source-microphone range is 5000 m and the frequency is $f = 1500$ Hz. (For interpretation of the references to color in this figure, the reader is referred to the web version of this paper.)

standard deviation), of the estimated z-coordinate using the three source localization methods with uniform or Gaussian assumption of the uncertain environmental parameters are shown in Figure 6. For the IB estimation, in the case of uniform distribution, the mean value of the z-coordinate estimates is $\mu(z_0^*) = -31.88$ m; by contrast, in the Gaussian case, the mean value is $\mu(z_0^*) = -31.91$ m. In both cases the PDF is spread around the reference value $z_0 = -32$ m and the confidence interval contains z_0, meaning that the IB method is able to retrieve the exact solution in most of the combinations of the uncertain environmental parameters. Similarly, the conventional MFP estimate has mean value $\mu(z_0^*) = -32.18$ m for uniform distribution and $\mu(z_0^*) = -32.11$ m for Gaussian distribution, which imply that the exact estimate can also be retrieved by the mean value. The mean value error of conventional MFP is more or less same as IB but the variance is smaller, which can be observed from Figure 6. The Capon's MFP estimate results in $\mu(z_0^*) = -33.92$ m for uniform distribution and $\mu(z_0^*) = -33.61$ m for Gaussian distribution, and the confidence interval is far away from z_0, which leads

to a much larger localization error than the two methods. Table 1 displays α-th percentile ($\alpha = 5\%, 10\%, 25\%, 50\%, 75\%, 90\%, 95\%$) of the distribution of the z coordinate estimate obtained from the three methods under the uniform or Gaussian assumption for the uncertain environmental parameters. It is interesting to notice that for IB the error is slightly larger far from the nominal value for the uniform assumption, so that in the case of Gaussian approximation for the environmental parameters the percentiles corresponding to a given error of 1 m increase by about 15%: the z-coordinate estimate at -33 m corresponds to the percentile being 10% for uniform assumption and 25% for the Gaussian assumption.

To further investigate the error distribution against the uncertain environmental parameter and to compare it with the MFP approaches, the response surface of $|\mathbf{r}_0^* - \mathbf{r}_0|$ is analysed for all the three source estimation methods. Maps of the source estimation error $|\mathbf{r}_0^* - \mathbf{r}_0|$ for IB, conventional and Capon's MFP are shown in Figure 7. As can be noticed, the conventional MFP approach returns the most robust result while Capon's MFP is most sensitive with respect to uncertain parameters. The IB estimates have an intermediate error in terms of dispersion of the results, but it is also reliable enough since the absolute error never exceeds 2 m in the given range of uncertain parameters. Here, we may conclude that IB is a good choice which returns an acceptable source localization error in presence of environmental parameter uncertainties with a relatively low computational cost.

Finally, Sobol' index calculation for IB shows a weak sensitivity of source estimation to the shallow water depth ($S_d \approx 3\%$), a strong sensitivity to the sound speed ($S_v \approx 92\%$) and a low effect of coupling between the two uncertain environmental parameters ($S_{d-v} = \sigma_{d-v}^2/\sigma^2 \approx 5\%$ where σ_{d-v}^2 is the residual variance). On the contrary, the MFP methods are more sensitive to the water depth and to the coupling of the two environmental parameters: the Sobol' indices are $S_d \approx 67\%$, $S_v \approx 1\%$ for conventional MFP and $S_d \approx 62\%$, $S_v \approx 2\%$ for Capon's MFP, respectively. The same conclusions can be observed in Figure 7. It is remarkable that this result is able to guide the choice of source localization method: when the uncertainty level of the water depth is higher than the sound speed IB can more accurately estimate the source, while in the opposite case the MFP methods are preferable in the sense of estimation robustness.

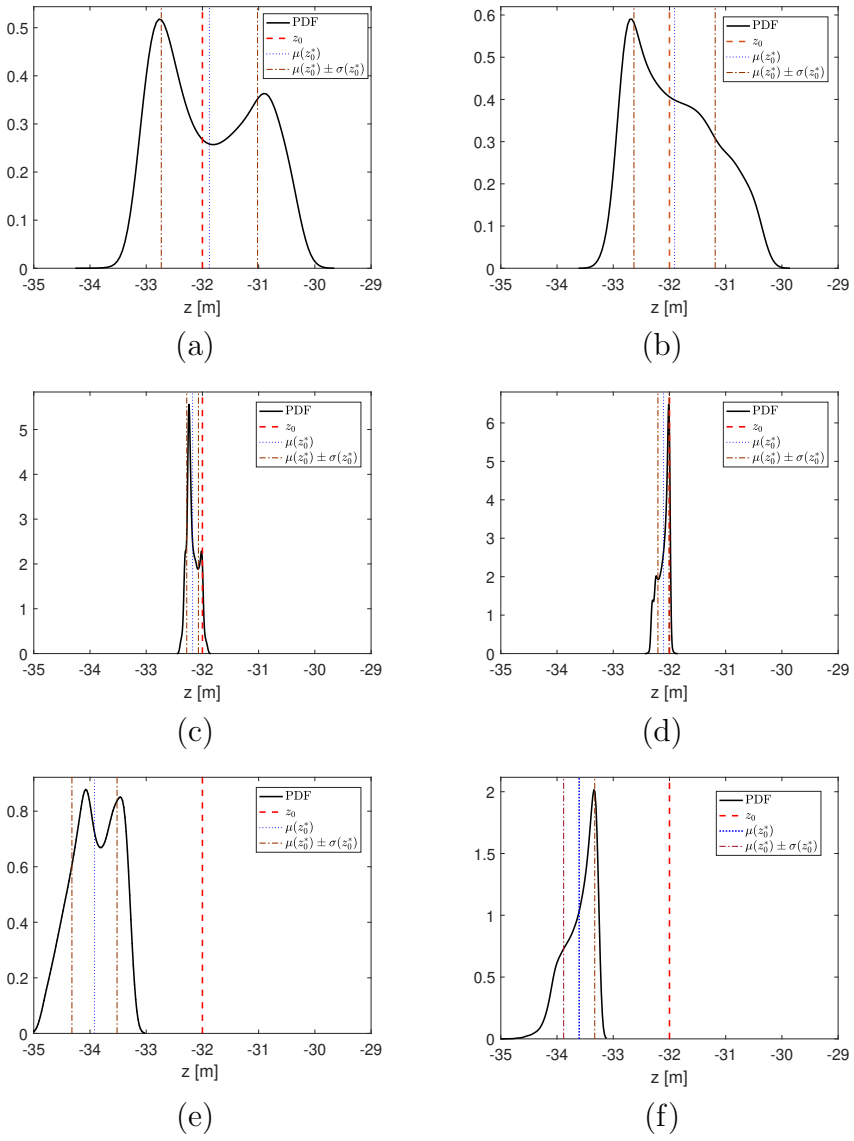


Figure 6: Probability density functions of the z-coordinate estimate using the IB (a,b), conventional MFP (c,d) and Capon's MFP (e,f) methods under the assumption of a uniform distribution (a,c,e) and a Gaussian distribution (b,d,f) for the uncertain environmental parameters.

6. Conclusions

In this paper, sound source localization in a shallow water environment is addressed. The multiple reflections on the water surface and bottom during

α	IB		Conventional MFP		Capon's MFP	
	Uniform	Gaussian	Uniform	Gaussian	Uniform	Gaussian
5%	-33.14	-32.97	-32.33	-32.30	-34.65	-34.14
10%	-33.00	-32.86	-32.31	-32.27	-34.50	-34.03
25%	-32.72	-32.59	-32.26	-32.19	-34.25	-33.82
50%	-32.03	-32.02	-32.20	-32.09	-33.93	-33.55
75%	-31.14	-31.38	-32.10	-32.03	-33.59	-33.38
90%	-30.72	-30.88	-32.02	-32.00	-33.42	-33.31
95%	-30.54	-30.65	-31.99	-31.99	-33.34	-33.29

Table 1: α-th percentile ($\alpha = 5\%, 10\%, 25\%, 50\%, 75\%, 90\%, 95\%$) of the distribution of the z-coordinate estimate obtained from the IB, conventional MFP and Capon's MFP methods under the uniform and Gaussian assumption for the uncertain environmental parameters.

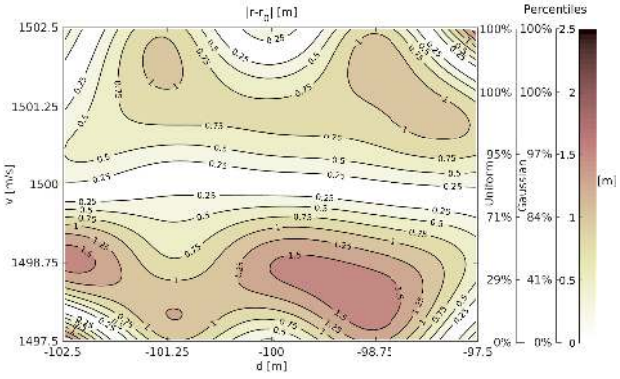
the sound propagation are considered. Instead of solving the wave equation with boundary conditions of reflection, an image source model including multiple sound sources in a free-field is proposed. Therefore, the multiple-reflection problem is transferred to a multiple source problem which results in an analytical solution of sound propagation. Then, the multiple sources are estimated using the iterative beamforming approach, which is a generalization of classical beamforming to deal with the case of multiple sources. The proposed source localization technique is a parametric method based on the maximum likelihood, which avoids the limits of spatial aliasing and resolution of classical beamforming.

The results obtained via numerical simulations show the interest of the proposed model and the source estimation method. The experiment justifies that classical beamforming is limited in the underwater source localization problem due to a long microphone-source distance and a relatively small size of microphones array. However, iterative beamforming is able to accurately estimate the sound source and return a super-resolution reconstruction of sound field. Furthermore, the proposed method has a much lower computational cost than the matched-field processing methods. In this paper, the uncertainty quantification study also justifies that the proposed iterative beamforming method is stable with respect to the model uncertainties: when the sound speed and water depth are imprecisely known (the error range is ± 2.5 m for water depth and ± 2.5 m/s for sound speed), the sound source can still be accurately localized.

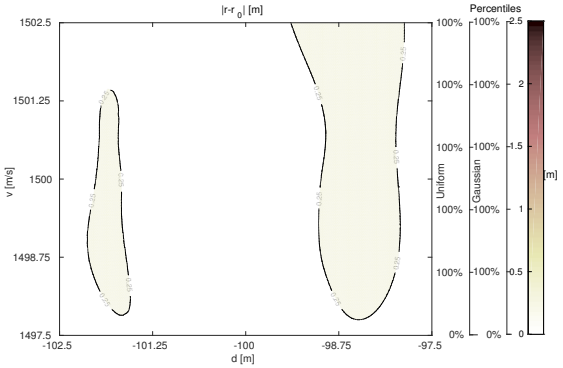
In this paper, only the case of homogeneous sound speed field is considered. However, the proposed image model and the source localization method can be naturally generalized to the inhomogeneous case. Furthermore, this work analyzes the sensitivities of the proposed method with respect to uncertain sound speed and water depth. Other uncertainties, for example the roughness of water bottom, can also be considered. It would be also interesting to model the uncertain parameters as random variables, which may further improve the accuracy of source localization.

Acknowledgments

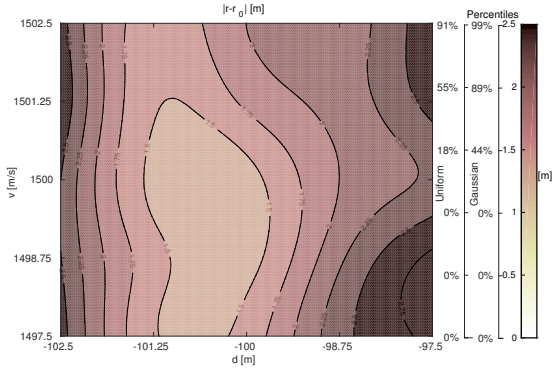
This work has been carried out thanks to the support of the A*MIDEX grant (reference number: ANR-11-IDEX-0001-02) funded by the French Government "Investissements d'Avenir" program.



(a)



(b)



(c)

Figure 7: Estimation error of sound source localization using (a) IB, (b) conventional MFP, and (c) Capon's MFP with imprecise information of water depth ($d \in [-102.5, -97.5]$) and sound velocity ($v \in [1497.5, 1502.5]$). Corresponding percentiles with respect to the uniform and Gaussian input distributions are given on the right side of the error map. (For interpretation of the references to color in this figure, the reader is referred to the web version of this paper.)

References

[1] W. A. Kuperman and J. F. Lynch. Shallow-water acoustics. *Physics Today*, pages 55–61, October 2004.

[2] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt. *Computational Ocean Acoustics*. Springer, second edition, 2011.

[3] J. Tabrikian and J. L. Krolik. Robust maximum-likelihood source localization in an uncertain shallow-water waveguide. *Journal of the Acoustical Society of America*, 101(1):241–249, 1997.

[4] J. Tabrikian and J. L. Krolik. Barankin bounds for source localization in an uncertain ocean environment. *IEEE Transactions on Signal Processing*, 47(11):2917–2927, 1999.

[5] S. E. Dosso. Environmental uncertainty in ocean acoustic source localization. *Inverse Problems*, 19:419–431, 2003.

[6] X. Wang, S. Khazaie, and P. Sagaut. Sound source localization in a randomly inhomogeneous medium using matched statistical moment method. *Journal of the Acoustical Society of America*, 138(6):3896–3906, 2015.

[7] S. Khazaie, X. Wang, and P. Sagaut. Localization of random acoustic sources in an inhomogeneous medium. *Journal of Sound and Vibration*, 384:75–93, 2016.

[8] A. Jakoby, J. Goldberg, and H. Messer. Source localization in shallow water in the presence of sensor location uncertainty. *IEEE Journal of Oceanic Engineering*, 25(3):331–336, 2000.

[9] P. Castellini and M. Martarelli. Acoustic beamforming: analysis of uncertainty and metrological performances. *Mechanical Systems and Signal Processing*, 22:672–692, 2008.

[10] 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*, 66–67:159–177, 2016.

[11] S. Kim, G. F. Edelmann, W. A. Kuperman, W. S. Hodgkiss, H. C. Song, and T. Akal. Spatial resolution of time-reversal array in shallow water. *Journal of the Acoustical Society of America*, 110(2):820–829, 2001.

[12] R. Zhang, W. Li, X. Qiu, and G. Jin. Reverberation loss in shallow water. *Journal of Sound and Vibration*, 186(2):279–290, 1995.

[13] K. Yoo and T.C. Yang. Broadband source localization in shallow water in the presence of internal waves. *Journal of the Acoustical Society of America*, 106(6):3255–3269, 1999.

[14] D. E. Weston and H. W. Andrews. Acoustic fluctuations due to shallow-water internal waves. *Journal of Sound and Vibration*, 31(3):357–367, 1973.

[15] F. R. DiNapoli and R. L. Deavenport. Theoretical and numerical Green's function field solution in a plane multilayered medium. *Journal of the Acoustical Society of America*, 67(3):92–105, 1980.

[16] C. L. Pekeris. Theory of propagation of explosive sound in shallow water. *Geological Society of America Memoirs*, 27:1–116, 1948.

[17] P. Roux and M. Fink. Time reversal in a waveguide: Study of the temporal and spatial focusing. *Journal of the Acoustical Society of America*, 107(5):2418–2429, 2000.

[18] J. Cuenca, F. Gautier, and L. Simon. The image source method for calculating the vibrations of simply supported convex polygonal plates. *Journal of Sound and Vibration*, 322(4):1048–1069, 2009.

[19] E. G. Williams. *Fourier Acoustic: Sound Radiation and Nearfield Acoustical Holography*. Academic Press, 1999.

[20] M. Fink. Time-reversed acoustics. *Scientific American*, pages 91–97, November 1999.

[21] M. Fink and C. Prada. Acoustic time-reversal mirrors. *Inverse Problems*, 17(1):1–38, 2001.

[22] G. Papanicolaou, K. Solna, and L. Ryzhik. Statistical stability in time reversal. *SIAM Journal on Applied Mathematics*, 64(4):1133–1155, 2004.

[23] P. Blomgren, G. Papanicolaou, and H. Zhao. Super-resolution in time-reversal acoustics. *Journal of the Acoustical Society of America*, 111(1):230–248, 2002.

[24] P. Druault, R. Marchiano, and P. Sagaut. Localization of aeroacoustic sound sources in viscous flows by a time reversal method. *Journal of Sound and Vibration*, 332(15):3655–3669, 2013.

[25] B. D. Van Veen and K. M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE Acoustic, Speech and Signal Processing Magazine*, 5:4–24, 1988.

[26] H. Krim and M. Viberg. Two decades of array signal processing research. *IEEE Signal Processing Magazine*, 13(4):67–94, 1996.

[27] J. J. Christensen and J. Hald. Technical review: beamforming. *Bruel & Kjaer*, 2004.

[28] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky. An overview of matched field methods in ocean acoustics. *IEEE Journal of Oceanic Engineering*, 18(4):401–424, 1993.

[29] D. Kim, W. Seong, Y. Choo, and J. Lee. Localization of incipient tip vortex cavitation using ray based matched field inversion method. *Journal of Sound and Vibration*, 354:34–46, 2015.

[30] J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418, 1969.

[31] X. Wang, B. Quost, J.-D. Chazot, and J. Antoni. Iterative beamforming for identification of multiple broadband sound sources. *Journal of Sound and Vibration*, 365:260–275, 2016.

[32] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B (Methodology)*, 39(1):1–38, 1977.

[33] J. C. F. Wu. On the convergence properties of the EM algorithm. *Annals of Statistics*, 11:95–103, 1983.

[34] D. G. Krige. A statistical approach to some basic mine valuations problems on the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 52:119–139, 1951.

[35] G. Matheron. Principles of geostatistics. *Economic Geology*, 58:1246–1266, 1963.

[36] B. Echard, N. Gayton, and M. Lemaire. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. *Structural Safety*, 33(2):145–154, 2011.

[37] N. Gayton, J.M. Bourinet, and M. Lemaire. CQ2RS: a new statistical approach to the response surface method for reliability analysis. *Structural Safety*, 25(1):99–121, 2003.

[38] J.C. Jouhaud, P. Sagaut, and B.Labeyrie. A Kriging approach for CFD/wind-tunnel data comparison. *Journal of Fluids Engineering*, 128, 2006.

[39] T. Bracconier, M. Ferrier, J.C. Jouhaud, M. Montagnac, and P. Sagaut. Towards an adaptive POD/SVD surrogate model for aeronautic design. *Computer and Fluids*, 40:195–209, 2011.

[40] L. Margheri and P. Sagaut. A hybrid anchored-ANOVA–POD/Kriging method for uncertainty quantification in unsteady high-fidelity CFD simulations. *Journal of Computational Physics*, 324:137–173, 2016.

[41] B. Picinbono. Second-order complex random vectors and normal distributions. *IEEE Transactions on Signal Processing*, 44(10):2637–2640, 1996.

[42] S. Geisser. The predictive sample reuse method with applications. *Journal of the American Statistical Association*, 70:320–328, 1975.

[43] I. Sobol. Sensitivity estimates for non-linear mathematical models. *Mathematical Modeling and Computational Experiment*, 1(4):407–414, 1993.

[44] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola. Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259–270, 2010.