Archives of Acoustics Vol. 50, No. 2, pp. 173–185 (2025), https://doi.org/10.24425/aoa.2025.153663

Research Paper

# Experimental Results of Diver Detection in Harbor Environments Using Single Acoustic Vector Sensor

Saier MAHMOUD*, Louay SALEH, Ibrahim CHOUAIB

*Department of Electronic and Mechanical Systems, Higher Institute for Applied Sciences and Technology*
Damascus, Syria

*Corresponding Author e-mail: saeir.mahmoud@hiast.edu.sy

*Received February 5, 2025; revised March 30, 2025; accepted March 31, 2025; published online May 29, 2025.*

This paper addresses the detection of divers with an open-circuit scuba. An acoustic vector sensor (AVS), which contains four channels, one for the pressure component, and three for orthogonal particle velocity components is proposed to be used. A novel covariance matrix analysis (CMA) method is presented for estimating the signal power using AVS in three-dimensional measurements. This method is based on solving a quartic equation that relates the determinant and trace of the AVS covariance matrix to the reciprocal of the signal-to-noise ratio (SNR) in a three-dimensional isotropic acoustic field with spherical isotropic noise. This method is compared with two traditional methods: the AVS pressure channel power, and the minimum variance distortionless response (MVDR) beamformer, in estimating the acoustic power associated with the diver's breathing. Experimental data from sea trials demonstrate the capability of all three methods to reconstruct the waveform of the acoustic diver signal and highlight the periodic breathing patterns. The diver's breathing rate and corresponding power are estimated using the fast Fourier transform (FFT) of the power signal, therefore serving as a key signature for diver detection. The experiment demonstrates that the CMA method gives better diver detection index compared to the other two methods.

**Keywords:** scuba diver; acoustic vector sensor (AVS); detection; minimum variance distortionless response (MVDR); signal-to-noise ratio (SNR); covariance matrix analysis (CMA); fast Fourier transform (FFT).



## 1. Introduction

An acoustic signal is generated when the diver breathes through an apparatus regulator, with air bubbles discharged from it. This signal is a broad-spectrum (Johansson *et al.*, 2010) ranging from hundreds of hertz to 75 kHz (Tu *et al.*, 2020), quasi-periodic (Gorovoy *et al.*, 2015), with a periodicity ranging from 2.44 s to 7.09 s, and exhibits a repetition pattern corresponding to the diver's breathing rate, which typically falls within the range of 0.14 Hz to 0.41 Hz (Donskoy *et al.*, 2008). This range variability is influenced by several factors such as the diver's age, experience, activity level, and scuba equipment used (Donskoy *et al.*, 2008). Passive sonar systems utilize this periodic signal to detect the presence of a diver. The respiratory cycle of a diver initiates with the inhalation phase, which corresponds to sound frequencies exceeding 2 kHz, followed by the exhalation phase, which corresponds to frequencies below 2 kHz (Tu *et al.*, 2020). Both signals are useful in detection systems (Hari *et al.*, 2015). The specific frequency band of interest for analysis varies among scientific papers, with some focusing on the band with a high-frequency band above 2 kHz (Tu *et al.*, 2020; Jin, Xu, 2020; Lennartsson *et al.* 2009; Li *et al.*, 2015), others focus on the low-frequency band (Korenbaum *et al.*, 2016; Gorovoy *et al.*, 2014). The detectability of a diver can be determined by two distinctive indicators: the power of the frequency band and its repetition rate (Korenbaum *et al.*, 2020), which apply to the two types of scuba, open circuit and closed circuit, while the first emits more acoustic noise (Donskoy *et al.*, 2008), the detection of a diver with a closed circuit is still challenging.

As the acoustic signal propagates away from the source, two distinct fields are produced: pressure and particle motion (PM). Pressure is a scalar quantity

that can be measured using a hydrophone, whereas PM is a vector quantity that is oriented parallel to the direction of wave propagation in the free far field. PM can be measured directly using inertial sensors such as accelerometers or geophones, or indirectly through a configuration of nearby pressure sensors. In last case, the differential measurements approximated the pressure gradient, providing an estimate of PM acceleration (Gray *et al.*, 2016; Nedelec *et al.*, 2021). The direct measurement approach involves challenges related to buoyancy, compliance, suspension, geometry, and flow. In the contrast, the indirect measurement approach faces many issues such as spacing, calibration uncertainty, noise, and flow (Gray *et al.*, 2016). Integrating the PM sensor with an omnidirectional pressure sensor into a single unit results in what is known as an acoustic vector sensor (AVS) (Roh *et al.*, 2022), which has recently gained attention and has increased usability (Yuan *et al.*, 2022; Dong *et al.*, 2024). This sensor has a wide range of applications in both terrestrial and aquatic environments. In terrestrial environment, it is used for localization, tracking, and speech enhancement (Chen *et al.*, 2018; Cao *et al.*, 2017), while in aquatic environment – the focus of this paper – it is employed for detection (Yuan *et al.*, 2022), localization (Chen *et al.*, 2023), and tracking (Nagananda, Anand, 2017).

Detection using a single AVS still requires further research efforts (Yuan *et al.*, 2022). The energy-flux detector is proved to perform as a maximum likelihood ratio detector under isotropic noise conditions (Sun *et al.*, 2003). For horizontal isotropic noise, Yuan *et al.* (2022) introduced a method to estimate the signal power by analyzing the covariance matrix of the 2D-AVS output. Incorporating this estimation into detection has proven to be more effective than traditional energy detectors under nonstationary ambient noise. Furthermore, an adaptive matched filter is proposed with 2D-AVS for passive broadband source detection, demonstrating superior performance against noise and interference (Ma *et al.*, 2019). The signal waveform can be optimally estimated using the minimum variance distortion response (MVDR) beamformer, also known as the Capon beamformer or the minimum power distortion response (MPDR) (Van Trees, 2002; Zhao *et al.*, 2018). This technique is particularly valuable for direction of arrival (DOA) estimation (Zhao *et al.*, 2018), which is a topic of growing interest among researchers. Various algorithms have been explored with a single AVS, including the arctan (Bereketli *et al.*, 2015), intensity-based (Wang *et al.*, 2014; Nehorai, Paldi, 1994), velocity-covariance-based (Nehorai, Paldi, 1994), beamforming (Zhao *et al.*, 2018; Bereketli *et al.*, 2015), maximum likelihood (Levin *et al.*, 2012), multiple signal classification MUSIC (Zhao *et al.*, 2018), and estimation of signal parameters via rotational invariance techniques Esprit (Tichavsky *et al.*, 2001; Paulraj *et al.*, 1985).

Multiple algorithms are available for detection the diver's acoustic signal. While all of these algorithms contribute to estimating the energy of the breathing frequency, they vary in the way they reconstruct the waveform of the diver's signal. The envelope spectrum was used within a 30 kHz–35 kHz bandwidth, achieving a detection range of up to 25 m (Lennartsson *et al.*, 2009). Chung *et al.* (2007) utilized a multiband normalized matched filter but a reference signal is needed. Tu *et al.* (2020) employed the envelope spectrum detection method within 13 kHz–18 kHz bandwidth and extended the range of detection from 20 m to 40 m by using an adaptive noise subtraction approach. All the aforementioned studies utilize data from a single hydrophone. Conversely, other studies have employed two hydrophones to reconstruct the waveform of the diver's signal through cross-correlation. This cross-correlation analysis determines the time delay between the signals received by the hydrophones, which is used in DOA estimation (Korenbaum *et al.*, 2020; Sutin *et al.*, 2013).

This paper focuses on utilizing an AVS to capture the acoustic signals emitted by a diver with an open-circuit scuba. Equations that relate the determinant and trace of the AVS covariance matrix to the reciprocal of signal-to-noise ratio (SNR) are extracted in a three-dimensions isotropic acoustic field with spherical isotropic noise. Solving these equations results in power signal estimation. We name this technique as covariance matrix analysis (CMA). Additionally, the MVDR beamformer is used to estimate the power signal by optimizing the azimuth and elevation angle values in order to maximize the MVDR spectrum. The presence of a diver is estimated through comparing the breathing frequency estimated power with a predefined threshold which is estimated empirically using a recorded data for ambient noise. The detection algorithm is evaluated using data obtained from sea trials.

The structure of this paper is arranged as follows. Section 2 describes the AVS with its mathematical model and covariance matrix. The signal power estimation using the AVS $P$-channel is presented in Sec. 3. In Sec. 4, the signal power estimation by maximizing the spectrum of MVDR beamformer is presented. In Sec. 5, the power signal estimation using the CMA method is performed. This method based on analysing the covariance matrix of AVS channels: pressure and velocity, which results in a quartic equation. Solving this equation provides an estimation of the signal power. The effectiveness of this method is evaluated using simulated data. The proposed diver detection approach using AVS is explained in Sec. 6 and evaluated in Sec. 7 using sea trial data. Finally, Sec. 7 concludes the paper by summarizing the key findings and implications derived from the study.

## 2. Acoustic vector sensor model

The AVS sensor has four output channels, one for pressure signal ($P$-channel) and three orthogonal components for particle velocity signal ($\mathbf{V}$-channels). So that, the AVS captures more information about the acoustic field comparing to hydrophones. The relationship between the PM velocity ($\mathbf{v}$), acceleration ($\mathbf{a}$), and the pressure ($p$) given by Euler's equation, is described as

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \mathbf{a} = -\frac{\nabla p}{\rho}, \tag{1}$$

where $t$ is the time, $\nabla$ is the gradient operation, and $\rho$ is the water density.

The PM vector, which has information about the direction of the signal, can be measured in two ways: 1) using a geophone to measure the PM velocity ($\mathbf{v}$), or accelerometer to measure the acceleration ($\mathbf{a}$), this way is named direct measurement type or inertial type; 2) by estimation of the $\nabla p$ using multiple spaced hydrophones, this way is named indirect type or gradient pressure type (Gray *et al.*, 2016; Nedelec *et al.*, 2021).

Under plane wave conditions, the relationship between the pressure signal and the PM velocity signal is expressed as (Abraham, 2019)

$$\mathbf{v} = -\frac{p}{\rho c}\mathbf{u}, \tag{2}$$

where $\rho c$ represents acoustic impedance, while $c$ denotes the speed of sound in water,

$$\mathbf{u} = \begin{bmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & \sin\theta \end{bmatrix}^{\mathrm{T}}$$

is a unit vector oriented from sensor to source, $\phi$ is the azimuth angle, and $\theta$ is the elevation angle (Fig. 1).

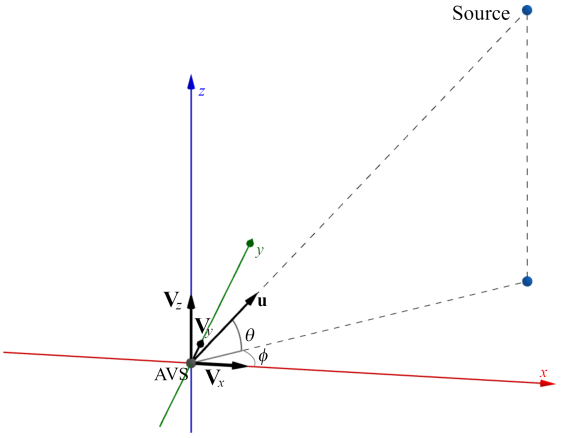


Fig. 1. AVS coordinates.

The AVS output $\mathbf{s}(t)$, after scaling the velocity channels with $\rho c$, is modelled as

$$\begin{aligned}\mathbf{s}(t) &= \begin{bmatrix} p(t) + n_p(t) & \mathbf{v}^{\mathrm{T}}(t) + \mathbf{n}_v^{\mathrm{T}}(t) \end{bmatrix} \\ &= \mathbf{h}(\phi,\theta)p(t) + \mathbf{n}_s(t), \end{aligned} \tag{3}$$

where

$$\mathbf{h}(\phi,\theta) = \begin{bmatrix} 1 & \mathbf{u}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}, \qquad \mathbf{n}_s(t) = \begin{bmatrix} n_p(t) & \mathbf{n}_v^{\mathrm{T}}(t) \end{bmatrix}^{\mathrm{T}}$$

is the ambient-noise, $n_p(t)$ is the pressure noise with power $\sigma_{np}^2$, and $\mathbf{n}_v^{\mathrm{T}}(t)$ is the velocity noise with power $\sigma_{nv}^2$. Under the ambient isotropic noise condition, the relation between $\sigma_{np}^2$ and $\sigma_{nv}^2$ is given as $\sigma_{np}^2 = 3\sigma_{nv}^2$ (Levin *et al.*, 2012).

Assuming $p(t)$ and $\mathbf{n}_s(t)$ are uncorrelated, the covariance matrix of $\mathbf{s}(t)$ is given as

$$\mathbf{R} = \frac{1}{T}\int_T \mathbf{s}(t) * \mathbf{s}^{\mathrm{T}}(t)\,\mathrm{d}t = \mathbf{R}_s + \mathbf{R}_n, \tag{4}$$

where $\mathbf{R}_n$ and $\mathbf{R}_s$ are the covariance matrix of noise and signal. In practice, $\mathbf{R}$ can be estimated from the received signal as following (Yuan *et al.*, 2022; Liu *et al.*, 2019):

$$\widehat{\mathrm{R}} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{s}(n)\mathbf{s}^{\mathrm{T}}(n), \tag{5}$$

where $N$ is the length of snapshot. This matrix is a crucial in array signal processing, as demonstrated in the subsequent sections.

## 3. Signal power estimation using the AVS $P$-channel

The AVS contains a pressure sensor that functions as an omnidirectional hydrophone. The sensor's output, $s_p(t)$, is expressed as

$$s_p(t) = p(t) + n_p(t). \tag{6}$$

The power of measurement signal, assuming the noise $n_p(t)$ is uncorrelated with the signal $p(t)$, is defined as

$$\begin{aligned} y^2 &= \frac{1}{T}\int_T s_p^2(t) = \frac{1}{T}\int_T p^2(t)\,\mathrm{d}t + \sigma_{np}^2 \\ &= \sigma_s^2 + \sigma_{np}^2 = \sigma_s^2\left(1 + \frac{1}{\mathrm{SNR}}\right), \end{aligned} \tag{7}$$

where

$$\sigma_s^2 = \int_T p^2(t)\,\mathrm{d}t$$

is the signal power. For a high SNR, the power of pressure signal approaches to signal power.

In practice, this power is estimated by averaging the instantaneous power over a snapshot of length $N$ as

$$y^2 = \frac{1}{N}\sum_{n=1}^{N} s_p^2(n). \tag{8}$$

This method is used to estimate the power of the acoustic signal emitted by the diver to highlight its periodicity, and is compared to other signal power estimation methods described in the next section.

## 4. Signal power estimation using MVDR beamformer

Beamforming is a technique employed in array signal processing to steer, shape, and concentrate signals received from sensors toward a desired direction. The shape and width of the beam are adjusted by combining multiple signals through weighting and delays. The AVS features an array of four sensors located at the same point, allowing for the application of beamforming without the need for delay adjustments. The first order beamformer is formed as (Levin *et al.*, 2012)

$$y(n) = \mathbf{w}^{\mathrm{T}}\mathbf{s}(n) = \alpha p(n) + (1-\alpha)\mathbf{u}_{\mathrm{str}}^{\mathrm{T}}\mathbf{v}, \tag{9}$$

where

$$\mathbf{w}^{\mathrm{T}} = \begin{bmatrix} \alpha & (1-\alpha)\mathbf{u}_{\mathrm{str}}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$$

is the weights with $\alpha \in [0\ 1]$, and $\mathbf{u}_{\mathrm{str}}$ is the look vector.

The parameter $\alpha$ significantly influences the beam pattern, functioning as a monopole for $\alpha = 1$, and a dipole for $\alpha = 0$, while the optimum value of

$$\alpha = \alpha_0 = \frac{\sigma_{nv}^2}{\sigma_{np}^2 + \sigma_{nv}^2}$$

corresponds to the DOA estimator with maximum reliability (Levin *et al.*, 2012).

One of the most widely employed beamformers is the MVDR. The objective is to determine the weights $\mathbf{w}^{\mathrm{T}}$ that produces the signal $p(n)$ at the beamformer's output without distortion while minimizing noise power. This problem is mathematically formulated as follows:

$$\begin{aligned} \mathbf{w}_{\mathrm{MVDR}} &= \arg\min_{\mathbf{w}} y^2(n) = \arg\min_{\mathbf{w}} \mathbf{w}^{\mathrm{T}}\mathbf{R}\mathbf{w} \\ \text{subject to} \quad & \mathbf{w}^{\mathrm{T}}\mathbf{h}(\theta,\phi) = 1. \end{aligned} \tag{10}$$

Using the Lagrange multiplier (Van Trees, 2002), the solution is given as

$$\mathbf{w}_{\mathrm{MVDR}} = \frac{\mathbf{R}^{-1}\mathbf{h}}{\mathbf{h}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{h}}, \tag{11}$$

the covariance matrix $\mathbf{R}$ should represent the noise covariance matrix in MVDR, and the array measurement covariance matrix in MPDR. When the steering direction aligns with the signal direction, MVDR and MPDR are identical (Van Trees, 2002).

If the vector $\mathbf{h}$ is known, and the covariance matrix $\mathbf{R}$ is estimated using Eq. (5), the weight vector $\mathbf{w}_{\mathrm{MVDR}}$ can be estimated, allowing the estimation of $\widehat{p}(n) = y(n) = \mathbf{w}_{\mathrm{MVDR}}^{\mathrm{T}}\mathbf{s}(n)$ in the output of beamformer. And the associated power is:

$$\begin{aligned} y^2(n) &= \mathbf{w}_{\mathrm{MVDR}}^{\mathrm{T}}\mathbf{R}\mathbf{w}_{\mathrm{MVDR}} \\ &= \frac{1}{\mathbf{h}^{\mathrm{T}}(\theta,\phi)\mathbf{R}^{-1}\mathbf{h}(\theta,\phi)} = f(\theta,\phi). \end{aligned} \tag{12}$$

In case where the direction angles are unknown, the MVDR spectrum $f(\theta,\phi)$, which is a measure of power radiating from direction $(\theta,\varnothing)$, is computed by changing $(\theta,\varnothing)$ over all possible values. The solution for the DOA problem is by searching $(\theta,\varnothing)$ for maxima of $f(\theta,\phi)$. Therefore, the MVDR can be used in addition to DOA estimation, to estimate the max power radiating from this estimated direction. The estimated power can be applied to reconstruct the waveform of AVS signals. Figure 2 presents an example of the spectrum of the MVDR for a single source emitting a linear frequency signal within the band [600–1000] Hz, with −3 dB power, from a position $(\theta,\varnothing) = (45°, 45°)$, and with an SNR = 5 dB for additive Gaussian noise.

For the experimental data, only the estimated signal power is utilized, as this paper focuses on diver detection rather than the diver's DOA.

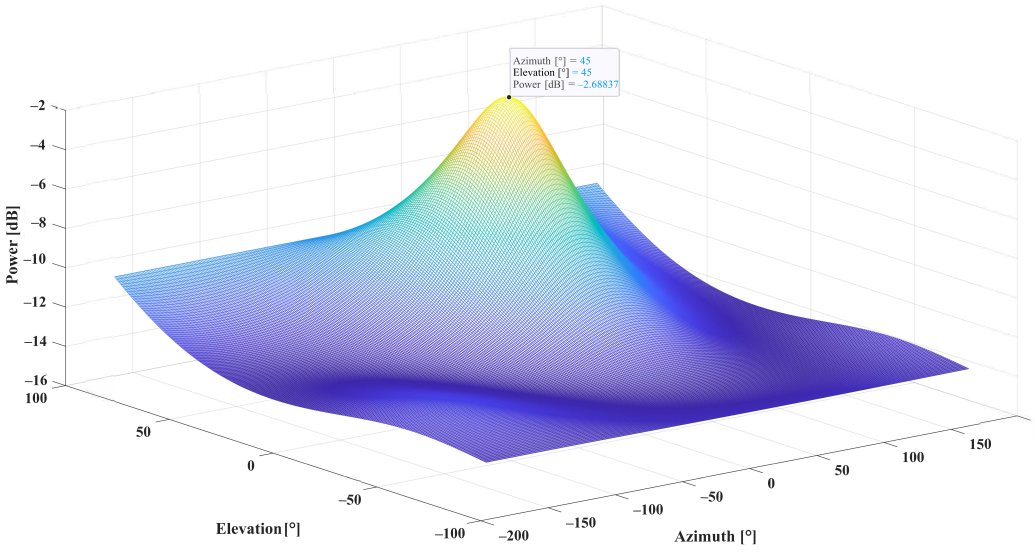


Fig. 2. MDVR spectrum for source radiates signal from $(\theta,\varnothing) = (45°, 45°)$ angles and SNR = 5 dB.

## 5. Signal power estimation using the AVS channels through CMA

For 2D-AVS, the signal power was estimated by solving a cubic equation that relates signal and noise power, to the determinant and trace of the AVS covariance matrix. The cubic equation is derived under the assumption of a horizontal isotropic acoustic field (Yuan *et al.*, 2022). In this section, the solution for 3D-AVS with isotropic noise in both horizontal and vertical dimensions is presented.

The covariance matrix of the AVS output is the sum of signal covariance $\mathbf{R}_s$ and the noise covariance $\mathbf{R}_n$, as shown in the Eq. (4). The matrix $\mathbf{R}_s$ is expressed as follows:

$$\mathbf{R}_s = \mathbf{h}(\theta,\phi)\mathbf{h}^{\mathrm{T}}(\theta,\phi)\sigma_s^2 = \begin{bmatrix} 1 & \cos\theta\cos\phi & \cos\theta\sin\phi & \sin\theta \\ \cos\theta\cos\phi & \cos^2\theta\cos^2\phi & \cos^2\theta\cos\phi\sin\phi & \cos\theta\sin\theta\cos\phi \\ \cos\theta\sin\phi & \cos^2\theta\cos\phi\sin\phi & \cos^2\theta\sin^2\phi & \cos\theta\sin\theta\sin\phi \\ \sin\theta & \cos\theta\sin\theta\cos\phi & \cos\theta\sin\theta\sin\phi & \sin^2\theta \end{bmatrix}\sigma_s^2. \tag{13}$$

This matrix has four eigenvalues $\lambda_1 = 2\sigma_s^2$ and $\lambda_2 = \lambda_3 = \lambda_4 = 0$.

The matrix $\mathbf{R}_n$, under isotropic noise, is expressed as follows:

$$\mathbf{R}_n = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/3 \end{bmatrix}\sigma_{np}^2. \tag{14}$$

As a result, the matrix $\mathbf{R}$ can be expressed as follows:

$$\mathbf{R} = \begin{bmatrix} 1+\alpha & \cos\theta\cos\phi & \cos\theta\sin\phi & \sin\theta \\ \cos\theta\cos\phi & \cos^2\theta\cos^2\phi+\frac{\alpha}{3} & \cos^2\theta\cos\phi\sin\phi & \cos\theta\sin\theta\cos\phi \\ \cos\theta\sin\phi & \cos^2\theta\cos\phi\sin\phi & \cos^2\theta\sin^2\phi+\frac{\alpha}{3} & \cos\theta\sin\theta\sin\phi \\ \sin\theta & \cos\theta\sin\theta\cos\phi & \cos\theta\sin\theta\sin\phi & \sin^2\theta+\frac{\alpha}{3} \end{bmatrix}\sigma_s^2, \tag{15}$$

where

$$\alpha = \frac{\sigma_{np}^2}{\sigma_s^2} = \mathrm{SNR}^{-1}.$$

The eigenvalues of $\mathbf{R}$ are

$$\lambda_{1,2} = \frac{1}{3}\left(2\alpha \pm \sqrt{\alpha^2+9}+3\right)\sigma_s^2 \quad \text{and} \quad \lambda_3 = \lambda_4 = \frac{1}{3}\alpha\sigma_s^2.$$

When there is only a signal $\alpha \to 0$, the eigenvalues are $\lambda_1 = 2\sigma_s^2$, and $\lambda_2 = \lambda_3 = \lambda_4 = 0$. These correspond to the eigenvalues of $\mathbf{R}_s$. Conversely, when $\alpha \to \infty$, indicating the absence of a signal and the presence of noise only, the eigenvalues are given as $\lambda_1 = \sigma_{np}^2$, and $\lambda_2 = \lambda_3 = \lambda_4 = \frac{\sigma_{np}^2}{3}$ that are associated with the eigenvalues of $\mathbf{R}_n$.

The trace of $\mathbf{R}$ is expressed as

$$\mathrm{trace}(\mathbf{R}) = \sum_{i=1}^{4}\lambda_i = 2(\alpha+1)\sigma_s^2. \tag{16}$$

The determinant of $\mathbf{R}$ is expressed as

$$\det(\mathbf{R}) = \prod_{i=1}^{4}\lambda_i = \frac{1}{27}\alpha^3(\alpha+4)\sigma_s^8. \tag{17}$$

Substituting $\sigma_s^2$ into Eqs. (16) and (17) yields a quartic equation in $\alpha$, which can be presented as

$$k\alpha^4 + 4k\alpha^3 + 6\alpha^2 + 4\alpha + 1 = 0 \tag{18}$$

where

$$k = 1 - \frac{\mathrm{trace}(\mathbf{R})^4}{432\det(\mathbf{R})}.$$

This equation has a real positive root corresponding with $k < 0$ as shown in Appendix. By solving Eqs. (16)–(18), the values of $\alpha$, $\sigma_s^2$, and $\sigma_{np}^2$ can be determined. In practice, by estimating the covariance matrix $\widehat{\mathbf{R}}$ as indicated in Eq. (5), these parameters can be determined using the following equations:

$$\widehat{k} = 1 - \frac{\mathrm{trace}(\mathbf{R})^4}{432\det(\widehat{\mathbf{R}})},$$

$$\widehat{k}\widehat{\alpha}^4 + 4\widehat{k}\widehat{\alpha}^3 + 6\widehat{\alpha}^2 + 4\widehat{\alpha} + 1 = 0, \tag{19}$$

$$\widehat{\sigma}_s^2 = \frac{\mathrm{trace}(\mathbf{R})}{2(\widehat{\alpha}+1)}, \qquad \widehat{\sigma}_{np}^2 = \alpha\widehat{\sigma}_s^2.$$

Figure 2 shows the implementation of this method for a linear frequency signal within the band [600–1000] Hz, with additive Gaussian noise and using 1000 Mont Carlo runs. Figure 3a presents the SNR esti-

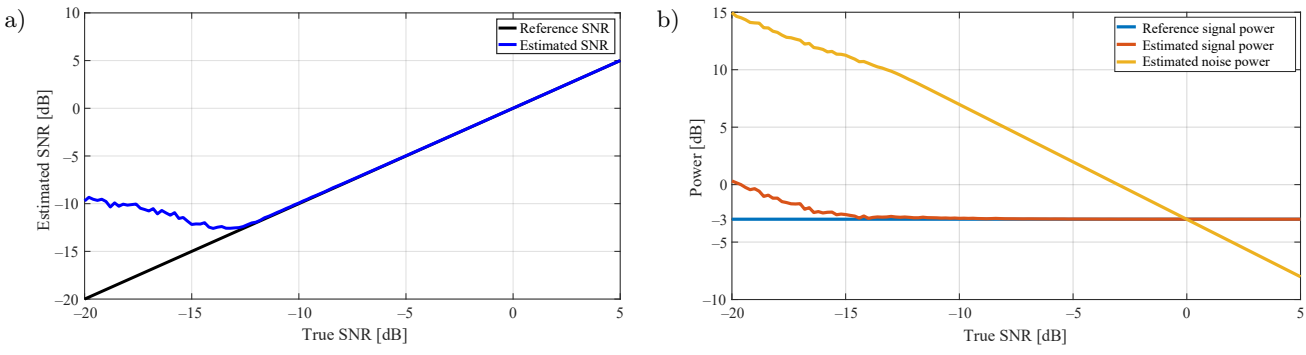


Fig. 3. Estimation using CMA: a) SNR; b) signal and noise power.

mation as the actual SNR varies within the range of [−20 5] dB. The accurate estimation achieved for SNR > −12 dB. While Fig. 3b presents the estimation of noise power and signal power as the SNR changes within the same range. And the true power signal estimation (the true value is −3 dB) is for SNR > −12 dB.

This simulation data demonstrates the technique's capability in estimation the signal power under low SNR. So, this method with the above two method is used to detect the acoustic signal power generated by the diver.

## 6. Steps for diver detection using AVS

Our method for detecting the diver presence involves analyzing the energy associated with their breathing rate. The process is shown in Fig. 4.

The process has the following steps:

1) capture the acoustic signal from the diver using an AVS;
2) apply a bandpass filter to increase the SNR;
3) reshape the signal to resemble a periodic waveform, employing various techniques including estimated signal power using the AVS pressure channel ($P$-channel), CMA, and MVDR beamformer;
4) compute the power of the diver's acoustic signal power within the frequency band of 0.14 Hz–0.41 Hz, utilizing the fast Fourier transform (FFT);
5) compare the computed breath power energy against a predefined threshold.

The proposed methods, AVS pressure channel, CMA, and MVDR, are evaluated and compared in diver detection under sea trials, as discussed in the following section.

## 7. Experimental results

A passive sonar system (Fig. 5a) was placed 5 m deep on the floor of the marine basin in Tartous harbour. Figure 5 illustrates both the system and the experimental location. The system includes two hydrophones and an AVS. The AVS is of the type VHS-90, with a sensitivity of −180 dB across four channels. The VHS-90 sensor contains three pairs of accelerometers arranged along three orthogonal directions, and six hydrophones connected in parallel to a single output in order to obtain an omnidirectional response. All sensors are encapsulated and coated uniformly with polyurethane material to satisfy the waterproof and sound-permeable requirements.

An expert young diver, with an open-circuit scuba, navigated around the sensor without following a consistent path due to the highly murky water conditions. The movement is roughly drawn in Fig. 6. The diver started his trajectory at point A (18, 5) directed to point B (0, 10). He rested at B for 0.35 min, then pro-

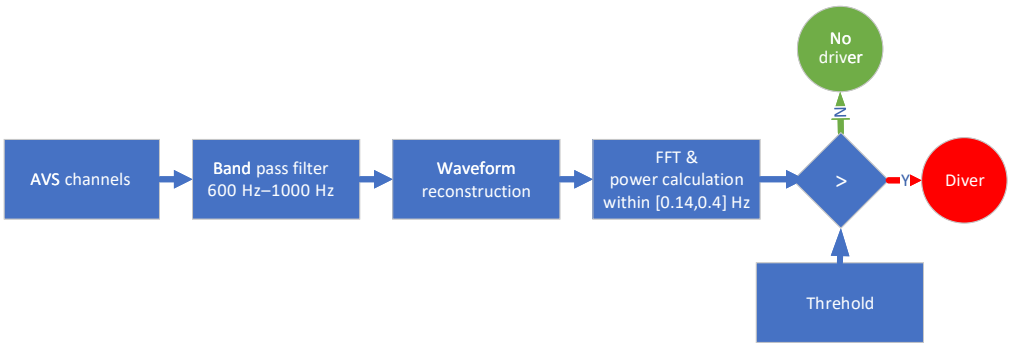


Fig. 4. Diver detection algorithm.

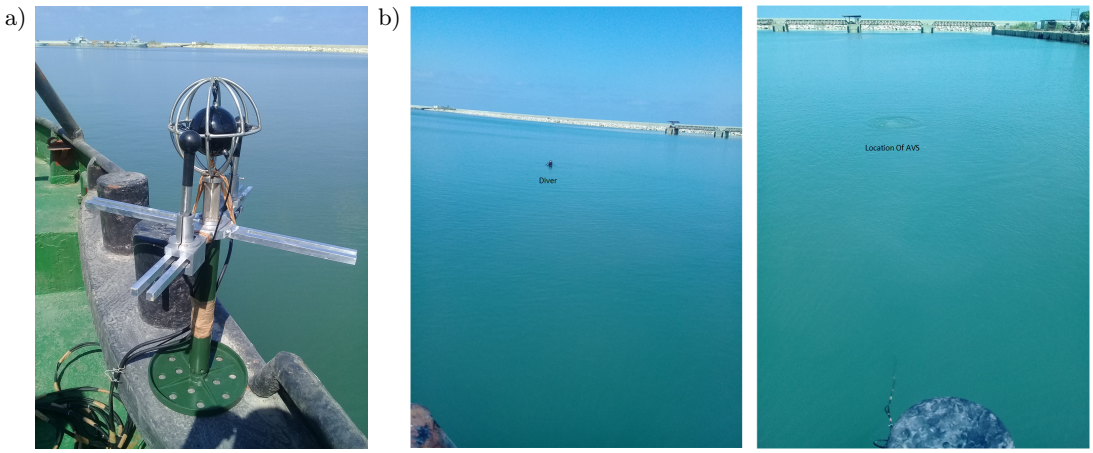


Fig. 5. Our system and experimental location: a) passive sonar system; b) experimental location.

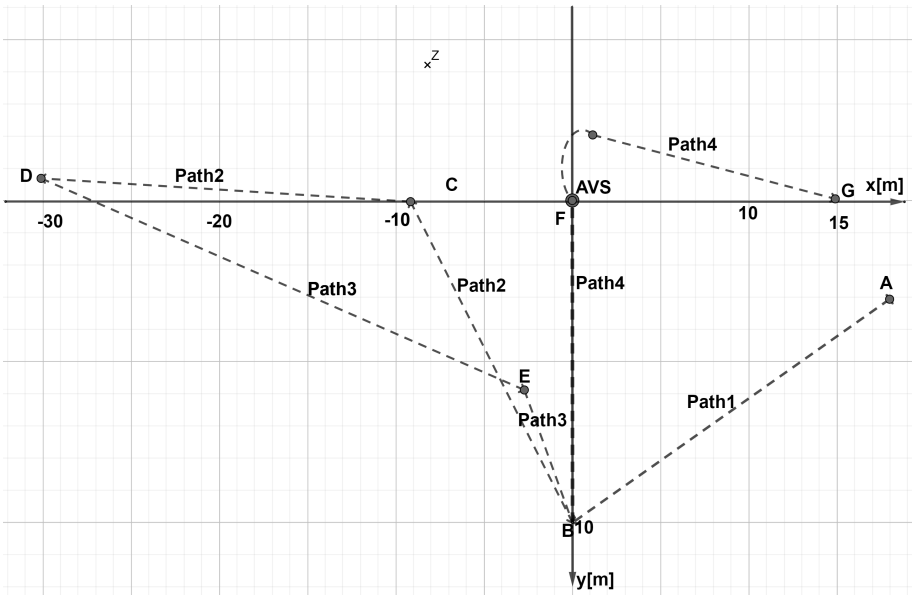


Fig. 6. Approximation of movement of diver around AVS.

ceeded to point C, then D, which is approximately 30 m away from the AVS. The diver has then returned toward point E, before heading back to point B. The diver rested at B 0.5 min, then went toward the AVS, he circled around the sensor and heading back to point G.

The data was captured using an A/D converter at a sampling rate of 44 100 Hz with 24-bit resolution. The MATLAB program was used for processing.

The recorded data was filtered by the 100th-order bandpass FIR filter within the 600 Hz–1000 Hz range in the time domain, corresponding to the exhalation and air bubble signals. Figure 7 illustrates the AVS's channels over time and their corresponding power (the power is calculated using a sliding window with 200 ms and 50 % overlapping). The figure shows a repeated pattern with period 6.3 s. Each pulse corresponds with one breathing of the diver.

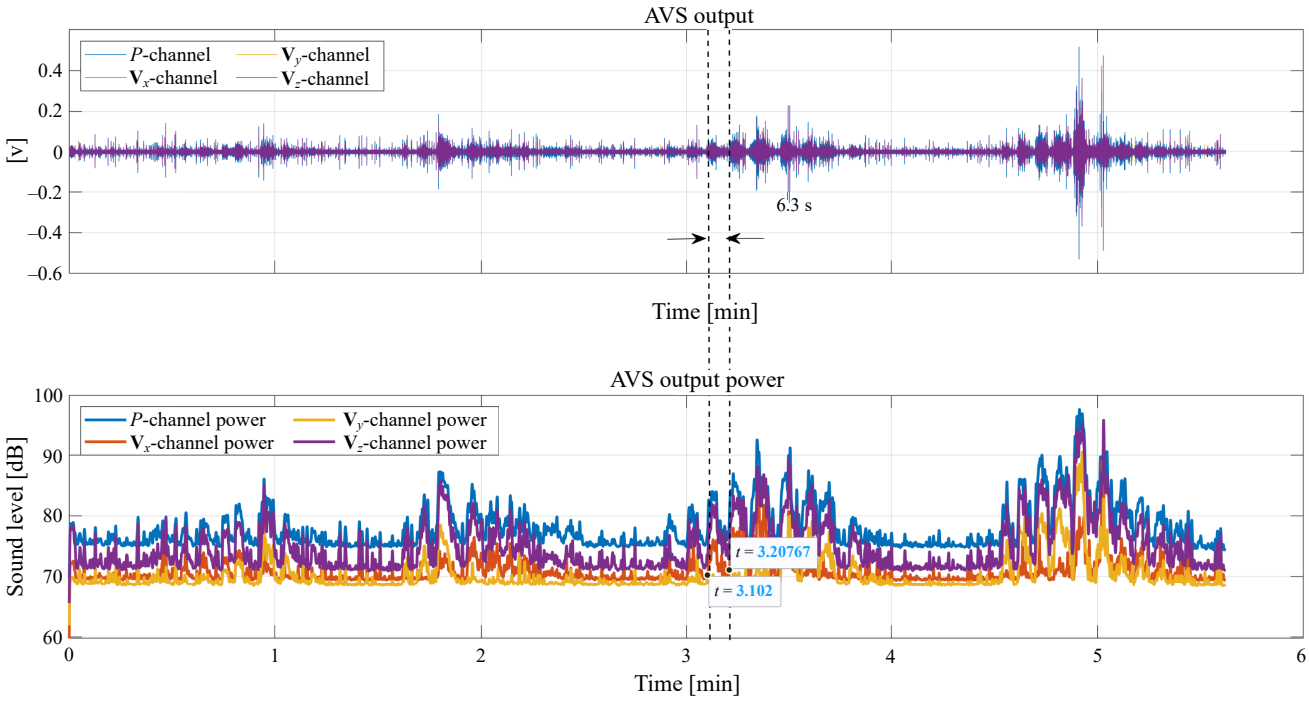


Fig. 7. AVS output signals and corresponding power levels.

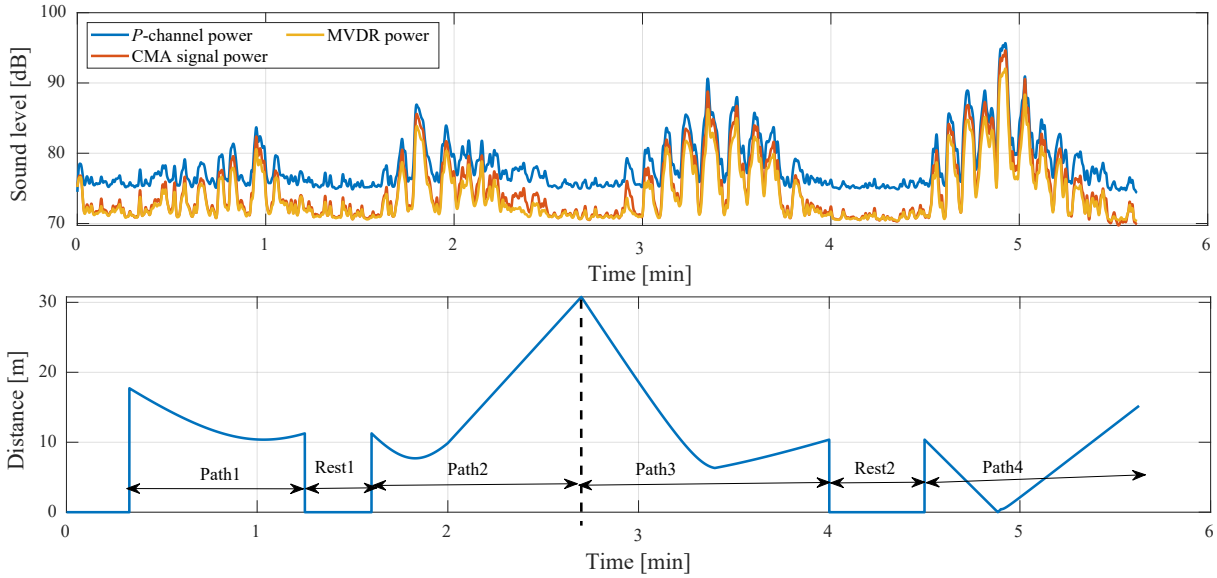


Fig. 8. Signal power estimation and distance between diver and AVS.

The three estimation methods ($P$-channel power, MVDR, and CMA) were implemented, and the results are presented in Fig. 8, which also shows the distance between the diver and the AVS over time. As illustrated in this figure, the signal power decreases with increasing the distance from diver to AVS, falling below the noise power by the end of Path3. This power decreasing is resulted from attenuation caused by the underwater environment. Furthermore, this figure illustrates that all three methods effectively estimate the waveform of the diver's signal, which exhibits periodic characteristics resulting from repetition in the breathing process. However, the interference between pulses, caused by scattering air bubbles, is evident especially when the diver is close to sensor (Path4), where the acoustic signal from the previous breath persists as the current breath begins.

Utilizing a $P$-channel for detection is feasible; however, this approach relies only on the pressure sensor that captures the noise from all directions. In contrast, the CMA method depends on the correlation between the pressure and velocity measurements, similar to the MVDR method. However, the MVDR method involves greater computational complexity resulted from searching for maximum power.

Figure 8 illustrates that the CMA and MDVR methods highlight the breathing pulses more effectively than the $P$-channel method. This is because they reduce noise power when the diver is silent or when SNR is low (e.g., Rest2). However, when the SNR is high (Path4), the estimated powers of all three methods converge to approximate the signal power.

After estimating the power signal which showed the periodicity in the diver's signal, it is necessary to determine the power within the diver breath rate range ([0.14–0.42] Hz), as this serves as an indicator of the diver's presence. This can be done by calculation the FFT of estimated power signal over a window of appropriate length that must contain multiple breathing cycles. Figure 9 illustrates FFT using a 13-second window when the diver is not silent. A local maximum at 0.15 Hz is found – indicating a period with 6.7 s – for all the three estimation methods, with advantages in value to CMA first, MVDR, then $P$-channel. The diver index $D_{\text{index}}$ is calculated by summing the squares of FFT values between 0.14 Hz–0.42 Hz as

$$D_{\text{index}} = \frac{1}{M} \sum_{f=0.14}^{f=0.42} \left| \text{FFT}\{y^2(n)\}(f) \right|^2 ,$$

where $M$ is the number of frequencies within the specified range. The results from repeating this process across the entire signals are illustrated in Fig. 10, indicating an increasing in diver index when the diver is actively breathing. By comparing the diver index to the predefined threshold, a decision of diver detection can be made. The CMA method shows a higher diver index value than the MVDR and $P$-channel methods, giving it an advantage in diver detection. For assertion, the metrics: accuracy, recall, precision, and $F1$-score are calculated and compared over these methods.

The predefined threshold $T_h$ can be estimated by applying the three methods to recorded ambient noise and selecting the maximum value as $D^N_{\text{index}}$. We then set the threshold $T_h = 1.3 D^N_{\text{index}}$ as indicated by Tu (2020). The probability of detection $P_d$ can be expressed as

$$P_d = \begin{cases} 1 & D_{\text{index}} \geqq 2T_h, \\ (D_{\text{index}} - T_h)/T_h & T_h < D_{\text{index}} < 2T_h, \\ 0 & D_{\text{index}} \leqq T_h. \end{cases} \tag{20}$$

This formula was applied to recorded data, and the results are presented in Fig. 10, which shows false alarms

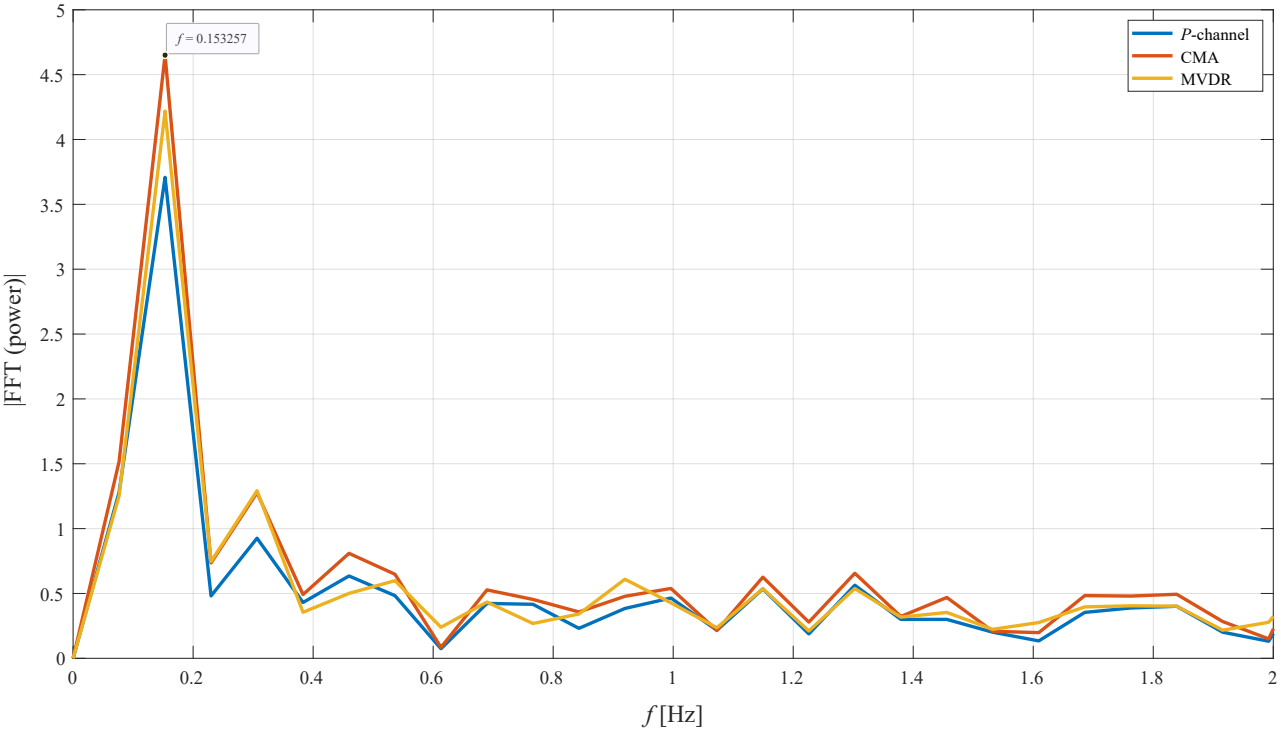


Fig. 9. FFT of window with 13 s length where the diver is present.

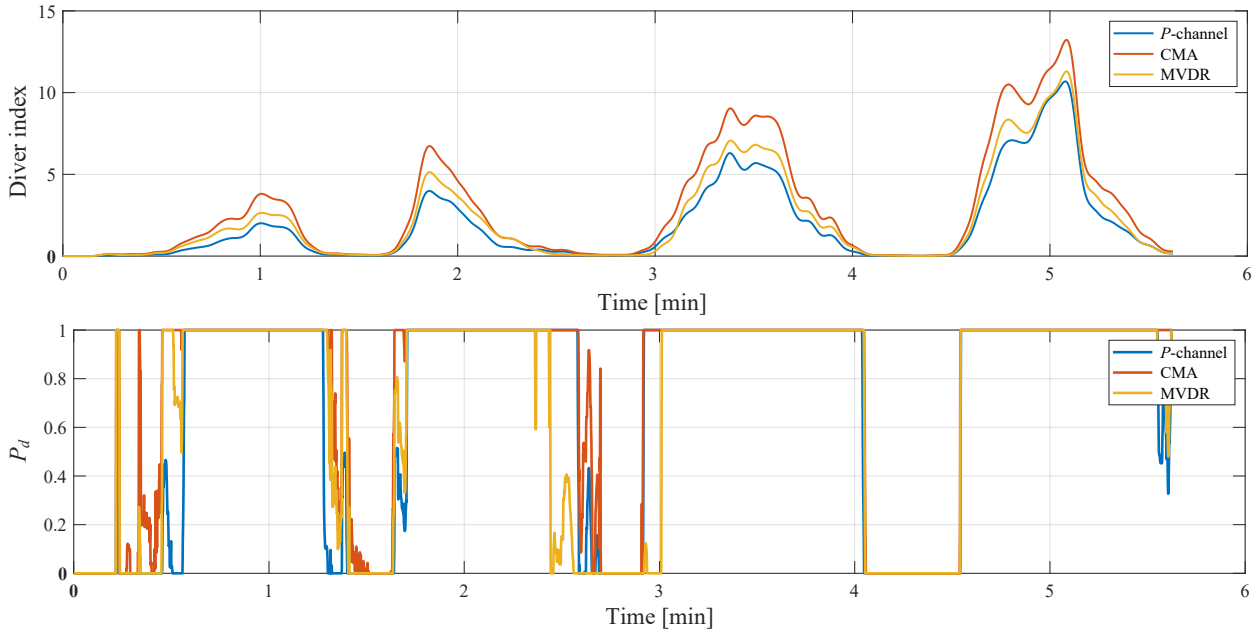


Fig. 10. Calculation of: a) the diver index for journey; b) probability of detection.

in the beginning. Additionally, a low value of $P_d$ is observed when the diver is either silent or far away from the sensor. To determine the detection distance for the three methods, the detection probability for Path3 was calculated and shown in Fig. 11. This figure shows that the distance of detection is approximately 29 m for both the $P$-channel and CMA methods with a preference for CMA, while this distance is about 26 m for the MVDR method.

To compare between the three detection methods, several commonly metrics were calculated as shown in Fig. 12. These metrics were derived based on the confusion matrix provided in Table 1.

The confusion matrix was estimated based on the defined threshold $T_h = 1.3D^N_{\text{index}}$. The matrix initially indicates comparable performance among the three methods, with the CMA method showing an advantage in correct classifications (true positives), while

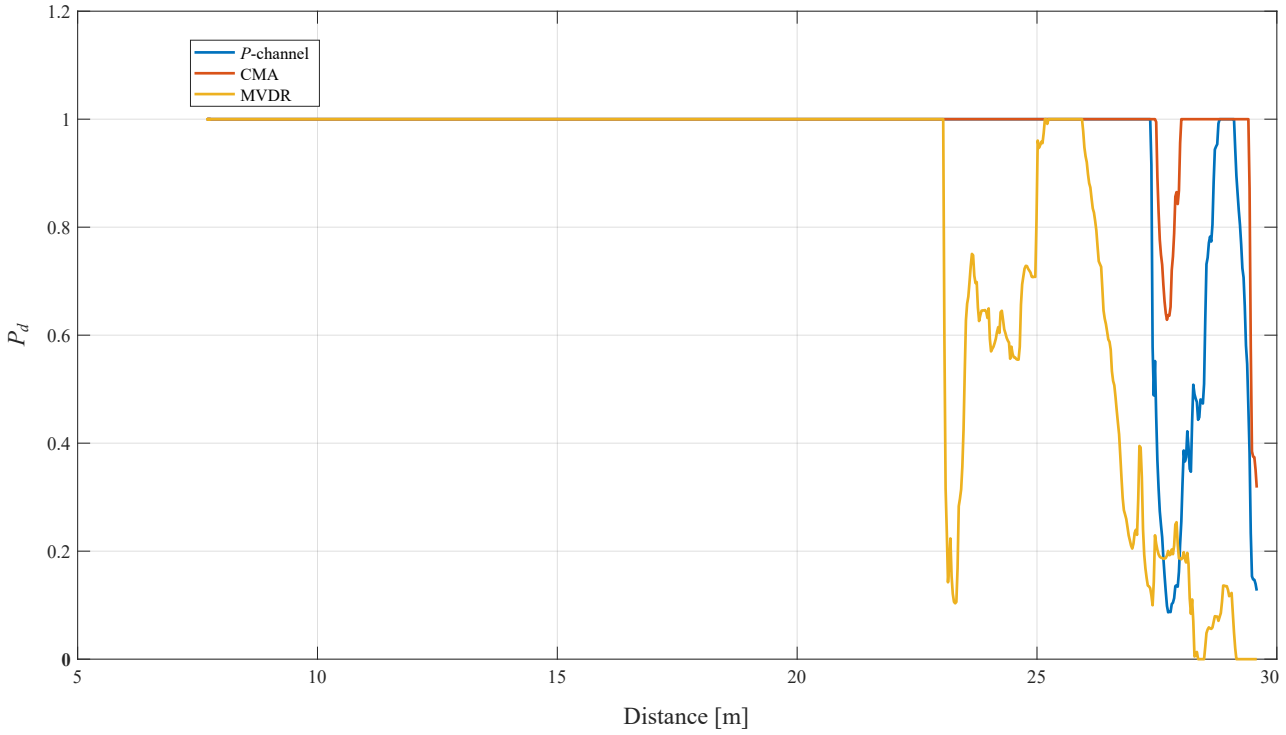


Fig. 11. Detection probability.

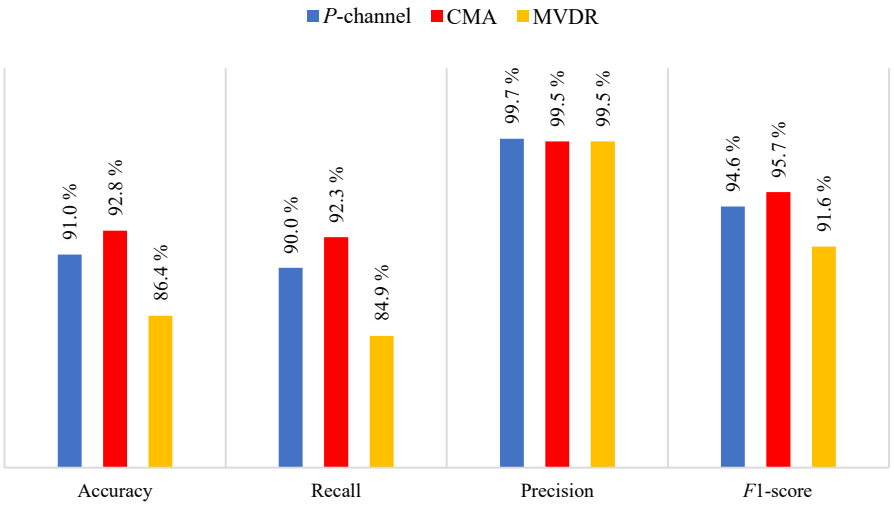


Fig. 12. Comparison of the three methods in detection performance under $T_h = 1.3D^N_{\text{index}}$.

Table 1. Confusion matrix under $T_h = 1.3D^N_{\text{index}}$.

| | | P-channel | | MVDR | | CMA | |
|---|---|---|---|---|---|---|---|
| | | Actual values | | | | | |
| | | Positive | Negative | Positive | Negative | Positive | Negative |
| Predicted values | Positive | 4102 | 11 | 3868 | 20 | 4203 | 21 |
| | Negative | 454 | 628 | 688 | 619 | 353 | 618 |

the $P$-channel method exhibits a lower value of false positives. It is important to note that increasing the threshold value affects the confusion matrix, generally resulting in degraded performance. However, beyond a certain threshold value, the performance of the MVDR method becomes superior to that of the $P$-channel method.

The results in Fig. 12 illustrate that: the accuracy of detection for the three methods was 86 % to 93 %, the recall was 85 % to 92 %, the precision was 99 % to

100 %, and the $F1$-score was 92 % to 96 %. These values demonstrate the effectiveness of the three methods in detecting the presence and absence of divers underwater, with advantage to the CMA method.

The experimental data demonstrated the possibility of using the three methods ($P$-channel, CMA, and MVDR power) in diver detection with a preference for the CMA method.

Therefore, detection is impacted by the method used to represent the periodicity resulting from the breathing process, which affects the diver index value. Furthermore, the long detection window (13 s) is a limitation of using the diver's breath rate as a key signature.

## 8. Conclusion

This study investigated the feasibility of detecting a diver with an open-circuit scuba by analyzing the acoustic signals emitted during respiration, which are captured by AVS. The AVS focuses on a low-frequency range of [600–1000] Hz, corresponding to the exhalation phase. The detection method involves estimating the signal power associated with the diver's breathing. In this paper, a novel method was presented to estimate the signal power by analyzing the covariance matrix of the 3D-AVS channels. This analysis leaded to the derivation of a quartic equation that relates the determinant and trace of the AVS covariance matrix to the reciprocal of the SNR. Solving this equation allows for the estimation of both the signal power and the SNR. Simulated data demonstrated the effectiveness of this method in estimating signal power under low SNR (−12 dB). Additionally, the paper presented the MVDR beamformer, which showed potential for estimating signal power along with azimuth and elevation angles using simulated data. The estimated power via CMA, MVDR, and $P$-channel methods was also compared using trial data, showing a periodicity corresponding to the diver's exhalation. The trial data demonstrated that the CMA method provides a stronger diver detection index compared to MVDR and the $P$-channel. However, a limitation of these methods is the long detection duration. This issue could potentially be addressed using AI algorithms to identify other features in acoustic diver signals. Investigating this approach will be a focus of our future work.

## Appendix. Roots of quartic equation

The existence of one real positive number for the quartic equation (Eq. (18)) is demonstrated, and it can be rewritten as

$$a\alpha^4 + b\alpha^3 + c\alpha^2 + d\alpha + e = 0,$$

$$a = kb = 4kc = 6, \qquad d = 4, \qquad e = 1.$$

By applying the change of variable $\alpha = \beta - \frac{b}{4a} = \beta - 1$, the depressed quartic equation has been obtained, which has the following form:

$$\beta^4 + p\beta^2 + q\beta + r = 0,$$

where

$$p = \frac{8ac - 3b^2}{8a^2} = 6\left(\frac{1}{k} - 1\right) = 6g,$$

$$q = \frac{b^3 - 4abc + 8a^2 d}{8a^3} = -8\left(\frac{1}{k} - 1\right) = -8g,$$

$$r = \frac{16ab^2 c - 64a^2 bd - 3b^4 + 256a^3 e}{256a^4} = 3\left(\frac{1}{k} - 1\right) = 3g,$$

$$g = \left(\frac{1}{k} - 1\right).$$

The presence of real roots can be determined by evaluating the signs or values of two terms (Prodanov, 2021) in the following form:

$$\delta(p,q,r) = 256r^3 - 128p^2r^2 + 144pq^2r + 16p^4r - 27q^4 - 4p^3q^2 = 6912g^3(g+1)^2,$$

$$L(p,q,r) = 8pr - 9q^2 - 2p^3 = -432g^2(g+1).$$

There are several cases depending on the value of $g$, as follows:

1) $g = 0 (k = 1) \rightarrow \delta = 0$, $L = 0$, $p = 0$: rewrite Eq. (18) as $(\alpha + 1)^4 = 0$, the root is $\alpha_i = -1$, this value is invalid because it is a negative number;
2) $g = -1\,(|k| \rightarrow \infty) \rightarrow \delta = 0$, $L = 0$, $p < 0$: rewrite Eq. (18) as $\alpha^3(\alpha + 4) = 0$, there are two real roots: $\alpha_1 = 0$, and $\alpha_2 = -4$. The root $\alpha_1$ indicates that the power of noise equals to zero, and the root $\alpha_2$ is invalid;
3) $g < 0 (k < 0 \cup k > 1) \rightarrow \delta < 0$: Eq. (18) has two distinct real roots;
4) $g > 0 (0 < k < 1) \rightarrow \delta > 0$, $L < 0$: Eq. (18) has not any real roots.

Thus, the acceptable case is 3, where there are two distinct real roots. To determine the sign of these roots, Descartes' rule of signs is applied, which is based on analyzing the sign changes in the coefficients of the polynomial, as follows:

- $k > 1$: the polynomial $f(\alpha) = k\alpha^4 + 4k\alpha^3 + 6\alpha^2 + 4\alpha + 1$ has no changes in the sign of coefficients. So, the two distinct roots are negative. And this case is invalid;
- $k < 0$: the polynomial $f(\alpha) = k\alpha^4 + 4k\alpha^3 + 6\alpha^2 + 4\alpha + 1$ has one sign change, indicating that the Eq. (14) has one positive real root. Additionally, the polynomial $f(-\alpha) = k\alpha^4 - 4k\alpha^3 + 6\alpha^2 - 4\alpha + 1$ shows three sign changes, meaning it has either three or one negative roots. Since Eq. (18) has two distinct roots, this results one invalid negative real root and one valid positive real root.

As a result, when $k < 0$, Eq. (18) has one positive root, which represents the valid solution for the reciprocal of SNR.

## References


1. Abraham D.A. (2019), *Underwater Acoustic Signal Processing: Modeling, Detection, and Estimation*, Springer, USA.
2. Bereketli A., Guldogan M.B., Kolcak T., Gudu T., Avsar A. (2015), Experimental results for direction of arrival estimation with a single acoustic vector sensor in shallow water, *Journal of Sensors*, **2015**(1): 1–9, https://doi.org/10.1155/2015/401353.
3. Cao J., Liu J., Wang J., Lai X. (2017), Acoustic vector sensor: Reviews and future perspectives, *IET Signal Processing*, **11**(1): 1–9, https://doi.org/10.1049/iet-spr.2016.0111.
4. Chen Y., Wang W., Wang Z., Xia B. (2018), A source counting method using acoustic vector sensor based on sparse modeling of DOA histogram, *IEEE Signal Processing Letters*, **26**(1): 69–73, https://doi.org/10.1109/LSP.2018.2879547.
5. Chen Y., Zhang G., Wang R., Rong H., Yang B. (2023), Acoustic vector sensor multi-source detection based on multimodal fusion, *Sensors*, **23**(3): 1301, https://doi.org/10.3390/s23031301.
6. Chung K.W., Li H., Sutin A. (2007), A frequency-domain multi-band matched-filter approach to passive diver detection, [in:] *2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*, pp. 1252–1256, https://doi.org/10.1109/ACSSC.2007.4487426.
7. Dong H., Suo J., Zhu Z., Li S. (2024), Improved underwater single-vector acoustic DOA estimation via vector convolution preprocessing, *Electronics*, **13**(9): 1796, https://doi.org/10.3390/electronics13091796.
8. Donskoy D.M., Sedunov N.A., Tsionskiy M.A. (2008), Variability of SCUBA diver's acoustic emission, [in:] Optics and Photonics in Global Homeland Security IV, pp. 272–282, https://doi.org/10.1117/12.783500.
9. Gorovoy S. *et al.* (2014), A possibility to use respiratory noises for diver detection and monitoring physiologic status, [in:] *Proceedings of Meetings on Acoustics*, **21**(1): 070007, https://doi.org/10.1121/1.4893767.
10. Gorovoy S. *et al.* (2015), Detecting respiratory noises of diver equipped with rebreather in water, [in:] *Proceedings of Meetings on Acoustics*, **24**(1): 070020, https://doi.org/10.1121/2.0000171.
11. Gray M., Rogers P.H., Zeddies D.G. (2016), Acoustic particle motion measurement for bioacousticians?: Principles and pitfalls, [in:] *Fourth International Conference on the Effects of Noise on Aquatic Life*, **27**(1): 010022, https://doi.org/10.1121/2.0000290.
12. Hari V.N., Chitre M., Too Y.M., Pallayil V. (2015), Robust passive diver detection in shallow ocean, [in:] *OCEANS 2015 – Genova*, pp. 1–6, https://doi.org/10.1109/OCEANS-Genova.2015.7271656.
13. Jin B., Xu G. (2020), A passive detection method of divers based on deep learning, [in:] *2020 IEEE 3rd International Conference on Electronics Technology (ICET)*, pp. 650–655, https://doi.org/10.1109/ICET49382.2020.9119556.
14. Johansson A.T., Lennartsson R.K., Nolander E., Petrović S. (2010), Improved passive acoustic detection of divers in harbor environments using prewhitening, [in:] *OCEANS 2010 MTS/IEEE SEATTLE*, pp. 1–6, https://doi.org/10.1109/OCEANS.2010.5664549.
15. Korenbaum V., Kostiv A., Gorovoy S., Dorozhko V., Shiryaev A. (2020), Underwater noises of open-circuit scuba diver, *Archives of Acoustics*, **45**(2): 349–357, https://doi.org/10.24425/aoa.2020.133155.
16. Korenbaum V.I. *et al.* (2016), The possibility of passive acoustic monitoring of a scuba diver, *Doklady Earth Sciences*, **466**(2): 187–190, https://doi.org/10.1134/S1028334X16020136.
17. Lennartsson R.K., Dalberg E., Persson L., Petrović S. (2009), Passive acoustic detection and classification of divers in harbor environments, [in:] *OCEANS 2009*, pp. 1–7, https://doi.org/10.23919/OCEANS.2009.5422407.
18. Levin D., Habets E.A.P., Gannot S. (2012), Maximum likelihood estimation of direction of arrival using an acoustic vector-sensor, *The Journal of the Acoustical Society of America*, **131**(2): 1240–1248, https://doi.org/10.1121/1.3676699.
19. Li S., Hu B., Zhou W., Zhao L. (2015), Experimental study for feature extraction of diver with atmospheric diving suit, [in:] *OCEANS 2015-MTS/IEEE Washington*, pp. 1–5, https://doi.org/10.23919/OCEANS.2015.7401820.
20. Liu A., Yang D., Shi S., Zhu Z., Li Y. (2019), Augmented subspace MUSIC method for DOA estimation using acoustic vector sensor array, *IET Radar, Sonar and Navigation,* **13**(6): 969–975, https://doi.org/10.1049/iet-rsn.2018.5440.
21. Ma L., Gulliver T.A., Zhao A., Ge C., Bi X. (2019), Underwater broadband source detection using an acoustic vector sensor with an adaptive passive matched filter, *Applied Acoustics*, **148**: 162–174, https://doi.org/10.1016/j.apacoust.2018.12.023.
22. Nagananda K.G., Anand G.V. (2017), Underwater target tracking with vector sensor array using acoustic field measurements, [in:] *OCEANS 2017 – Aberdeen*, pp. 1–10, https://doi.org/10.1109/OCEANSE.2017.8084844.
23. Nedelec S.L. *et al.* (2021), Best practice guide for underwater particle motion measurement for biological applications, Technical report by the University of Exeter for the IOGP Marine Sound and Life Joint Industry Programme, https://www.researchgate.net/publication/356911609.
24. Nehorai A., Paldi E. (1994), Acoustic vector-sensor array processing, [in:] *IEEE Transactions on Signal Processing*, **42**(9): 2481–2491, https://doi.org/10.1109/78.317869.

25. Paulraj A., Roy R., Kailath T. (1985), Estimation of signal parameters via rotational invariance techniques – Esprit, *Nineteeth Asilomar Conference on Circuits, Systems and Computers*, pp. 83–89, https://doi.org/10.1109/ACSSC.1985.671426.

26. Prodanov E.M. (2021), Classification of the real roots of the quartic equation and their pythagorean tunes, *International Journal of Applied and Computational Mathematics*, **7**(6): 1–14, https://doi.org/10.1007/s40819-021-01152-w.

27. Roh T. *et al.* (2022), Fabrication and underwater testing of a vector hydrophone comprising a triaxial piezoelectric accelerometer and spherical hydrophone, *Sensors*, **22**(24): 9796, https://doi.org/10.3390/s22249796.

28. Sun G., Yang D., Zhang L., Shi S. (2003), Maximum likelihood ratio detection and maximum likelihood DOA estimation based on the vector hydrophone [in Chinese], *Acta Acustica*, **28**(1): 66–72, https://doi.org/10.15949/j.cnki.0371-0025.2003.01.013.

29. Sutin A., Salloum H., DeLorme M., Sedunov N., Sedunov A., Tsionskiy M. (2013), Stevens passive acoustic system for surface and underwater threat detection, [in:] *2013 IEEE International Conference on Technologies for Homeland Security (HST)*, pp. 195–200, https://doi.org/10.1109/THS.2013.6698999.

30. Tichavsky P., Wong K.T., Zoltowski M.D. (2001), Near-field/far-field azimuth and elevation angle estimation using a single vector hydrophone, [in:] *IEEE Transactions on Signal Processing*, **49**(11): 2498–2510, https://doi.org/10.1109/78.960397.

31. Tu Q., Yuan F., Yang W., Cheng E. (2020), An approach for diver passive detection based on the established model of breathing sound emission, *Journal of Marine Science and Engineering*, **8**(1): 44, https://doi.org/10.3390/JMSE8010044.

32. Van Trees H.L. (2002), *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*, pp. 428–709, John Wiley & Sons, https://doi.org/10.1002/0471221104.

33. Wang X., Chen J., Han J., Jiao Y. (2014), Optimization for the direction of arrival estimation based on single acoustic pressure gradient vector sensor, *International Journal of Naval Architecture and Ocean Engineering*, **6**(1): 74–86, https://doi.org/10.2478/ijnaoe-2013-0164.

34. Yuan M., Wang C., Da L., Li Q. (2022), Signal detection method using a single vector hydrophone in ocean acoustics, *The Journal of the Acoustical Society of America*, **152**(2): 789–798, https://doi.org/10.1121/10.0013219.

35. Zhao A., Ma L., Hui J., Zeng C., Bi X. (2018), Open-lake experimental investigation of azimuth angle estimation using a single acoustic vector sensor, *Journal of Sensors*, **2018**(1), https://doi.org/10.1155/2018/4324902.