

Article

OCMS-D: A Sparsity-Constrained Modal Orthogonality Method for Single-Frequency Source Depth Estimation with a Vertical Array

Wei Gao ^{1,*} , Yangjin Xu ² and Guocheng Gao ¹

¹ State Key Laboratory of Acoustics and Marine Information, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China; gaoguocheng@mail.ioa.ac.cn

² College of Information Science and Engineering, Ocean University of China, Qingdao 266100, China; xuyangjin@stu.ouc.edu.cn

* Correspondence: gaowei@mail.ioa.ac.cn

Abstract

To address the challenge of estimating the depth of a single-frequency (tonal) source using a vertical line array (VLA) without prior knowledge of seabed parameters, a method called OCMS-D (orthogonality-constrained modal search-based depth estimation) is presented. The method exploits the orthogonality of mode depth functions as a physical constraint and leverages the sparsity of propagating normal modes in the received field. A convex optimization model is formulated to jointly estimate modal wavenumbers, depth functions, and complex mode amplitudes. Furthermore, a depth-sign search (DSS) is introduced to compensate for mode phase signs, which effectively suppresses sidelobes in the depth ambiguity function and enables high-accuracy depth estimation. Numerical simulations analyze the influence of signal-to-noise ratio, array parameters and sound speed profile (SSP) uncertainty on the algorithm's performance. Comparisons with matched-mode processing (MMP) show that OCMS-D exhibits superior robustness to SSP uncertainty and requires a smaller array aperture. Experimental data from the SWellEx-96 campaign demonstrate that the method achieves depth estimation errors of less than 5.4 m for the shallows and less than 10.8 m for the deep spurce. Notably, it requires neither seabed prior information nor source motion, offering a robust array signal processing framework for tonal source localization.

Keywords: modal orthogonality; sparsity-constrained estimation; vertical line array; depth ambiguity function



Received: 24 July 2026

Revised: 8 September 2026

Accepted: 9 September 2026

Published: 11 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Underwater source localization using passive sensor arrays is a fundamental yet challenging task in maritime electronic surveillance and acoustic sensing systems [1,2]. Among the various localization parameters, source depth estimation plays a critical role in target identification and threat assessment, particularly for distinguishing submerged objects from surface interferers in shallow-water environments [3–6].

Conventional depth estimation methods heavily rely on either broadband signal features or precise environmental modeling. Methods based on interference structure, such as the Lloyd's mirror effect, can provide accurate depth estimates when both direct and reflected paths are strong; however, they are inherently limited to broadband sources and relatively short ranges [7,8]. In practice, distant vessels or underwater vehicles often radiate

predominantly line-spectrum (tonal) components, rendering such broadband approaches ineffective. Conventional matched field processing (CMFP) remains a classical solution for harmonic source localization [9–12]. Nevertheless, the performance of CMFP critically depends on the accuracy of the forward propagation model, especially the seabed geoacoustic parameters (sound speed, density, and attenuation). Since accurate seabed parameters are rarely available in operational scenarios, CMFP typically suffers from severe environmental mismatch, leading to significant depth estimation errors.

To overcome these limitations, several advanced signal processing strategies have been developed. Yang proposed a data-based matched mode processing (DMMP) method that synthesizes a horizontal aperture from source motion to estimate modal wavenumbers and amplitudes [13,14]. While DMMP alleviates the need for seabed prior knowledge, it requires the source to move over a sufficiently long track (at least one interference span) and assumes that the source's original frequency is known—conditions that are often impractical in passive electronic warfare scenarios where the target's emission parameters are entirely unknown. More recently, the seabed-independent depth estimation (SIDE) method was introduced to estimate depth without seabed parameters [15]. However, SIDE still demands source motion and at least two stable tonal components in the radiated spectrum, which is rarely satisfied by real-world targets radiating a single dominant line.

Alternatively, deep learning (DL) has emerged as a powerful data-driven tool for underwater acoustic signal processing, with several studies applying neural networks to source depth classification and regression [16–21]. While DL-based methods can achieve impressive performance in simulated or controlled environments, they suffer from limited generalization capability when the test environment differs from the training dataset [22]. Moreover, the black-box nature of deep networks hinders interpretability—a critical requirement for defense and security-related electronic systems—and the substantial training data required further constrains their practical deployment. Accordingly, MMP is selected as the sole baseline because it can be applied to the same single-frequency VLA data, whereas DMMP and SIDE require source motion (with SIDE additionally requiring at least two stable tonal components), and DL methods require dedicated training data.

In this paper, we propose OCMS-D, a sparsity-constrained modal orthogonality framework for single-frequency source depth estimation using a vertical line array (VLA). Unlike conventional approaches, OCMS-D combines the physical orthogonality of mode depth functions with the signal sparsity of propagating normal modes in the received field, formulating the joint estimation of modal wavenumbers, depth functions, and complex amplitudes as a convex optimization problem (solved via the CVX toolbox [23]). To address the inherent phase-sign ambiguity that plagues conventional depth ambiguity functions (DAFs), we introduce a depth-sign search (DSS) mechanism based on Kullback–Leibler (KL) divergence to compensate for mode signs, thereby suppressing sidelobes and achieving high-resolution depth estimation. The proposed framework offers three major advantages from an electronic system perspective: (i) it operates with a single tonal frequency without requiring broadband emission; (ii) it does not require source motion; and (iii) it is independent of seabed parameters, requiring only the water-column sound speed profile (SSP), which can be readily measured by onboard conductivity–temperature–depth (CTD) sensors. These features make OCMS-D suitable for offline post-processing of tonal-source data acquired by compact passive sonar electronics, autonomous underwater vehicle (AUV)-based sensing payloads, and distributed acoustic sensor networks. All numerical simulations and data-processing procedures were performed using MATLAB R2023b (MathWorks, Natick, MA, USA), with convex optimization problems solved using the CVX toolbox (version 2.2).

The remainder of this paper is organized as follows. Section 2 presents the theoretical foundation of OCMS-D, including modal estimation via orthogonality-constrained sparse

recovery and the DSS-based phase compensation strategy. Section 3 evaluates the method's performance through numerical simulations, examining the effects of SNR, VLA parameters, and SSP uncertainty. Section 4 validates the method using experimental data from the SWellEx-96 campaign, a widely recognized benchmark in underwater acoustics. Finally, Section 5 concludes the paper with a summary and discussion on potential electronic system implementations.

2. Principle and Algorithm of OCMS-D

Our proposed OCMS-D method accomplishes source depth estimation through two signal processing stages: (i) joint estimation of modal parameters (wavenumbers, depth functions, and complex amplitudes) from the VLA measurements, and (ii) depth extraction via a phase-sign-compensated ambiguity function. This section presents the mathematical foundations and algorithmic details of each stage, with an emphasis on the sparse recovery formulation and the sign search mechanism that distinguish OCMS-D from existing approaches.

2.1. Signal Model and Sparse Modal Representation

Consider a range-independent shallow-water waveguide, where the sound speed profile (SSP) $c(z)$ in the water column is known (e.g., measured by a CTD sensor onboard the electronic system). A vertical line array (VLA) consisting of N hydrophones is deployed at depths $\mathbf{z}_r = [z_1, z_2, \dots, z_N]$. A distant tonal source radiating a single frequency f (angular frequency $\omega = 2\pi f$) is located at depth z_s and horizontal range r from the VLA.

The complex acoustic pressure vector $\mathbf{p} \in \mathbb{C}^{N \times 1}$ received by the VLA can be expressed as a finite linear combination of the propagating normal modes:

$$\mathbf{p} = \mathbf{\Psi}(k_1)\mathbf{a}(k_1) \quad (1)$$

where

$$\mathbf{p} = [p(z_1), p(z_2), \dots, p(z_N)]^T, \quad (2)$$

$$\mathbf{\Psi}(k_1) = [\boldsymbol{\psi}_1(k_1), \boldsymbol{\psi}_2(k_2), \dots, \boldsymbol{\psi}_M(k_M)], \quad (3)$$

$$\boldsymbol{\psi}_m(k_m) = [\psi_m(z_1, k_m), \psi_m(z_2, k_m), \dots, \psi_m(z_N, k_m)]^T, \quad m = 1, 2, \dots, M, \quad (4)$$

$$\mathbf{a}(k_1) = [a_1(k_1), a_2(k_2), \dots, a_M(k_M)]^T. \quad (5)$$

Here, k_m denotes the m -th modal wavenumber, $\boldsymbol{\psi}_m(k_m)$ is the corresponding mode depth function, and $a_m(k_m)$ is the complex mode amplitude. The notation k_1 in $\mathbf{\Psi}(k_1)$ and $\mathbf{a}(k_1)$ serves as a label, with the property that $\mathbf{\Psi}(k_1) = \mathbf{\Psi}(k_2) = \dots = \mathbf{\Psi}(k_M)$ and $\mathbf{a}(k_1) = \mathbf{a}(k_2) = \dots = \mathbf{a}(k_M)$. The number of propagating modes is denoted by M .

The mode amplitude of the m -th order is explicitly given by the Hankel function solution:

$$a_m(k_m) = \frac{i}{4} S(\omega) \psi_m(z_s, k_m) H_0^{(1)}(k_m r) \quad (6)$$

where $S(\omega)$ is the source spectrum at the tonal frequency, and $H_0^{(1)}(\cdot)$ is the zeroth-order Hankel function of the first kind. The pressure field in Equation (1) is typically sparse in the modal domain, since the number of propagating modes M is often much smaller than the number of VLA elements N , especially for low-frequency tonal sources in shallow water. This sparsity property is the key enabler for the convex optimization formulation introduced in the next subsection.

2.2. Orthogonality-Constrained Modal Parameter Estimation

The key innovation of the OCMS-D framework lies in exploiting two fundamental properties simultaneously: the orthogonality of mode depth functions and the sparsity

of the modal coefficient vector. Under the ideal waveguide condition, the mode depth functions satisfy the orthogonality relation [11]:

$$\langle \psi_m(z, k_m), \psi_n(z, k_n) \rangle = \sum_{l=1}^L \psi_m(z_l, k_m) \psi_n^*(z_l, k_n) h' \approx \int_0^H \psi_m(z, k_m) \psi_n^*(z, k_n) dz \approx \delta_{mn} \quad (7)$$

where $z_l = lh'$ with $h' \ll 1$ being the discretization step, H is the water depth, δ_{mn} is the Kronecker delta, and the superscript $*$ denotes complex conjugation. Equation (7) implies that mode depth functions corresponding to different modal wavenumbers are mutually orthogonal.

To perform modal estimation without seabed parameters, OCMS-D begins by computing the mode depth functions for all candidate wavenumbers within the predefined search interval $[\xi_{\min}, \xi_{\max}]$, where $\xi_{\min}$ is determined by the maximum possible number of propagating modes, and $\xi_{\max} = \omega / \min(c(z))$ corresponds to the highest achievable wavenumber. This is achieved via the following difference equation derived from the depth-separated Helmholtz equation [24]:

$$\begin{cases} \psi_m(z+h, \xi_m) + \psi_m(z-h, \xi_m) = 2 \cos\left(h\sqrt{k^2(z) - \xi_m^2}\right) \psi_m(z, \xi_m), \\ \psi_m(0, \xi_m) = 0, \\ \sum_{l=0}^L |\psi_m(lh, \xi_m)|^2 h = 1, \end{cases} \quad (8)$$

with $k(z) = \omega/c(z)$. Given any trial horizontal wavenumber ξ_m and the known SSP, Equation (8) permits the recursive computation of its corresponding mode depth function $\psi_m(\xi_m)$ starting from the surface boundary condition $\psi_m(0, \xi_m) = 0$, together with the normalization constraint. This step is performed once for all wavenumbers in the interval, yielding a comprehensive library of modal depth functions.

For each candidate wavenumber ξ , the algorithm then constructs the mode matrix $\Psi(\xi)$ by selecting, from the precomputed library, those modes whose depth functions are orthogonal to $\psi(\xi)$ under the orthogonality constraint given in Equation (7). This procedure effectively identifies a complete set of mutually orthogonal modes associated with the candidate ξ . With $\Psi(\xi)$ constructed, we solve the following sparsity-promoting convex optimization problem to obtain the corresponding coefficient vector $\mathbf{a}(\xi)$ [25]:

$$\min \|\mathbf{a}(\xi)\|_1 \quad (9)$$

subject to

$$\|\mathbf{p} - \Psi(\xi)\mathbf{a}(\xi)\|_2 < \epsilon_n, \quad (10)$$

where $\epsilon_n > 0$ is a tolerance parameter determined by the noise energy measured on the VLA. The ℓ_1 -norm minimization in Equation (9) encourages sparse solutions, while the constraint in Equation (10) ensures fidelity to the observed pressure field. This optimization problem is convex and can be efficiently solved using standard numerical solvers such as CVX [23]. Then, among all candidates, we select the one that yields the smallest ℓ_1 -norm of the coefficients:

$$\{\hat{k}_m\}_{m=1}^M = \arg \min_{\xi} \|\mathbf{a}(\xi)\|_1 \quad (11)$$

The estimated set of wavenumbers is denoted as $\{\hat{k}_m\}_{m=1}^M$. Simultaneously, the corresponding mode depth functions $\Psi(\hat{k}_m)$ and the complex mode amplitudes $\mathbf{a}(\hat{k}_m)$ are obtained as by-products of this joint estimation process.

2.3. Depth-Sign Search (DSS) and Source Depth Estimation

Once the modal depth functions $\psi_m(k_m)$ and the magnitude of mode amplitudes $|\mathbf{a}(k_m)|$ are estimated, the source depth can, in principle, be inferred from the DAF originally formulated for matched-mode processing [11]:

$$D(z) = \left| \sum_{m=1}^M \psi_m(z, k_m) |a_m(k_m)| \delta_m \right|^2 \quad (12)$$

where $\delta_m = \pm 1$ is a sign compensation factor. Ideally, substituting Equation (6) into Equation (12) yields a sharp peak at the true source depth z_s :

$$D(z) \sim \left| \sum_{m=1}^M \psi_m(z, k_m) \psi_m(z_s, k_m) \right|^2 \sim \delta(z - z_s) \quad (13)$$

However, the direct application of Equation (12) is hindered by the fact that the phase signs δ_m are unknown because $a_m(k_m)$ estimated from the sparse recovery process retains an arbitrary global phase. Incorrect sign compensation leads to destructive interference and severe sidelobes in the ambiguity function, rendering depth estimation unreliable.

To resolve this phase-sign ambiguity, we introduce a depth-sign search (DSS) mechanism based on the Kullback–Leibler (KL) divergence [26]. The fundamental observation is that, when correctly compensated, the ambiguity function approximates a Dirichlet kernel, which is a well-behaved function with a prominent main lobe and low sidelobes [26].

For a candidate depth index $q = 1, \dots, Q$ (where $Q = H/\Delta z$ and Δz is the depth sampling interval), the sign of the m -th mode is defined as

$$\delta_m(q) = \text{sign}[\psi_m(z_q, k_m)] \quad (14)$$

The resulting DAF under the q -th sign combination is

$$D(z, q) = \left| \sum_{m=1}^M \psi_m(z, k_m) |a_m(k_m)| \delta_m(q) \right|^2 \quad (15)$$

The corresponding theoretical (ideal) ambiguity function is modeled as a Dirichlet kernel:

$$Ds(z, q) = \frac{\sin[(M+1)\pi(z - z_q)/H]}{\sin[\pi(z - z_q)/(2H)]} + \epsilon \quad (16)$$

where ϵ is a small regularization constant (e.g., 10^{-6}) to avoid singularities. The KL divergence is then employed as a similarity metric between the estimated ambiguity function and the ideal kernel:

$$KL(q) = \int_0^H D(z, q) \ln \frac{D(z, q)}{|Ds(z, q)|} dz \quad (17)$$

The optimal sign combination is selected as the one minimizing the KL divergence:

$$\delta(q_0) = [\delta_1(q_0), \delta_2(q_0), \dots, \delta_M(q_0)], \quad q_0 = \arg \min_q KL(q) \quad (18)$$

Finally, the source depth is estimated by locating the global maximum of the compensated ambiguity function:

$$z_s = \arg \max_z D(z, q_0) \quad (19)$$

It is worth noting that the sign combination found by DSS inherently corresponds to a specific depth z_{q_0} , but due to possible sign repetitions across multiple depths (especially

when M is small), we do not directly use z_{q_0} as the final estimate. Instead, the peak search in Equation (19) provides a more robust and accurate depth estimate. A detailed discussion on this ambiguity issue is available in the SIDE study [15]. Accordingly, Algorithm 1 presents the pseudocode for the OCMS-D method.

Algorithm 1 OCMS-D Method

Require: $\mathbf{p}$, $c(z)$, f , $\mathbf{z}_r$, $[\xi_{\min}, \xi_{\max}]$, ϵ_n , Δz

Ensure: $\hat{z}_s$, $\{\hat{k}_m\}_{m=1}^M$, $\Psi(\hat{k}_m)$, $\mathbf{a}(\hat{k}_m)$

- 1: Define depth grid $z_q = q\Delta z$, $q = 1, \dots, Q$.
 - 2: Compute mode depth functions corresponding to all wavenumbers in the interval $[\xi_{\min}, \xi_{\max}]$ using Equation (8).
 - 3: **for** each candidate $\xi \in [\xi_{\min}, \xi_{\max}]$ **do**
 - 4: Use Equation (7) on the obtained modal depth functions to find those orthogonal to $\psi(\xi)$ and construct the mode matrix $\Psi(\xi)$.
 - 5: Obtain $\mathbf{a}(\xi)$ by solving $\min \|\mathbf{a}\|_1$ s.t. $\|\mathbf{p} - \Psi(\xi)\mathbf{a}(\xi)\|_2 < \epsilon_n$.
 - 6: **end for**
 - 7: Select $\{\hat{k}_m\}_{m=1}^M = \arg \min_{\xi} \|\mathbf{a}(\xi)\|_1$.
 - 8: Obtain $\Psi(\hat{k}_m)$, $\mathbf{a}(\hat{k}_m)$.
 - 9: **for** each $q = 1$ to Q **do**
 - 10: Compute signs $\delta_m(q)$ via Equation (14).
 - 11: Build $D(z, q)$ using Equation (15).
 - 12: Compute KL divergence $K(q)$ from Equations (16) and (17).
 - 13: **end for**
 - 14: Select $q_0 = \arg \min_q K(q)$, set optimal signs $\delta(q_0)$.
 - 15: Construct $D(z, q_0)$ via Equations (15) and (18).
 - 16: Compute $\hat{z}_s = \arg \max_z D(z, q_0)$ as per Equation (19).
 - 17: **return** $\hat{z}_s$, $\{\hat{k}_m\}$, $\Psi(\hat{k}_m)$, $\mathbf{a}(\hat{k}_m)$
-

3. Numerical Simulations

This section evaluates the performance of OCMS-D through numerical simulations. The SSP used in the simulation was collected in the Yellow Sea during spring, as shown in Figure 1. It contains a thermocline, which linearly decreases from 1496 m/s at 8 m to 1485 m/s at 10 m, and remains nearly constant down to the sea bottom at 31 m. The sediment type is set to sandy silt, and its geoacoustic parameters are presented in Table 1. And in Table 1, λ denotes the wavelength.

Table 1. Geoacoustic parameters of sandy silt, sand–silt–clay, and coarse sediment types.

Sediment Type	Label	Density ρ_b (g/cm ³)	Sound Speed c_b (m/s)	Attenuation Coefficient α_b (dB/ λ)
Sandy silt	True	1.77	1652	0.2
Sand–silt–clay	Mismatch 1	1.60	1579	0.3
Coarse	Mismatch 2	2.03	1836	0.06

As shown in Figure 1, a VLA comprising 30 elements with an element spacing of 1 m is deployed in the water column. Unless otherwise specified, the first element is fixed at a depth of 1 m below the surface in the following simulations. A tonal source at depth $z_s = 20$ m radiates a single-frequency signal at 596 Hz, and the range r between the source and the VLA is 5 km. Under these conditions, the signal received by the VLA predominantly contains 10 modes. In data processing, all parameters except the SSP are assumed unknown, and the wavenumber search interval is set to $\xi_{\min} = 2.0851 \text{ m}^{-1}$ and $\xi_{\max} = 2.5217 \text{ m}^{-1}$.

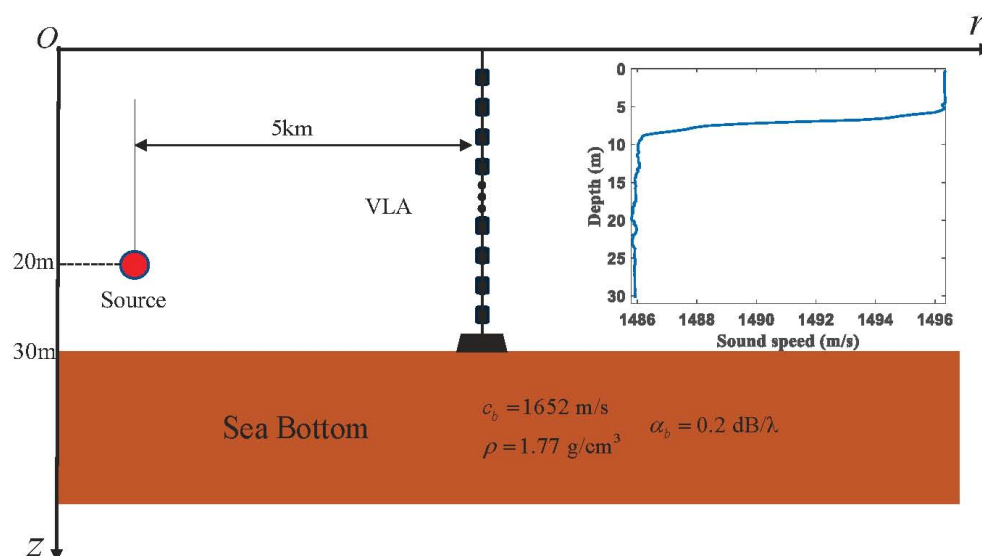


Figure 1. Sound speed profile and VLA geometry used in the numerical simulations.

The pressure received by the VLA is simulated using the KRAKEN program [27], with 30 dB Gaussian white noise added. The SNR is defined as

$$\text{SNR} = 20 \lg \frac{\|\mathbf{p}\|_2}{\|\mathbf{n}\|_2} \quad (20)$$

where $\mathbf{p} \in \mathbb{C}^{30 \times 1}$ and $\mathbf{n} \in \mathbb{C}^{30 \times 1}$ represent the pressure vector and the noise vector, respectively, and $\mathbb{C}$ denotes the complex number domain.

3.1. Results of Mode Estimation and Depth Estimation

Using the OCMS-D algorithm, the estimated modes are obtained as shown in Figure 2. On the computing platform (CPU: i7-12700, 32 GB RAM, CVX toolbox version 2.2), the average time for a single CVX optimization is 0.2950 s. With a wavenumber step size of $5 \times 10^{-4} \text{ m}^{-1}$, the total number of candidate wavenumbers is 874, and the overall average runtime of the entire OCMS-D algorithm is 258.4039 s. This runtime indicates that the current MATLAB/CVX implementation is computationally intensive. Specifically, processing one frequency-domain snapshot requires 874 CVX optimizations and approximately 258.4 s on the stated computing platform. Therefore, the present implementation of OCMS-D is intended for offline processing rather than real-time or near-real-time source localization. Figure 2a–c compare the estimated horizontal wavenumbers (blue asterisks), mode amplitudes (blue line with dots), and mode depth functions (blue line) with their reference values (red markers), respectively. The results indicate that the estimated results align well with the reference values.

The signs of the estimated mode depth functions are then determined using Equation (14). In this paper, the sign is sampled at an interval of 0.1 m, i.e., $Q = H/0.1$. These values are substituted into Equations (15)–(17) to estimate the mode signs. Subsequently, the source depth estimation result is derived using Equation (19). Figure 3 shows the DAF compensated by the mode sign combination corresponding to the minimum value of the loss function (blue line). The estimated depth is $\hat{z}_s = 20 \text{ m}$, which aligns exactly with the reference depth (green line).

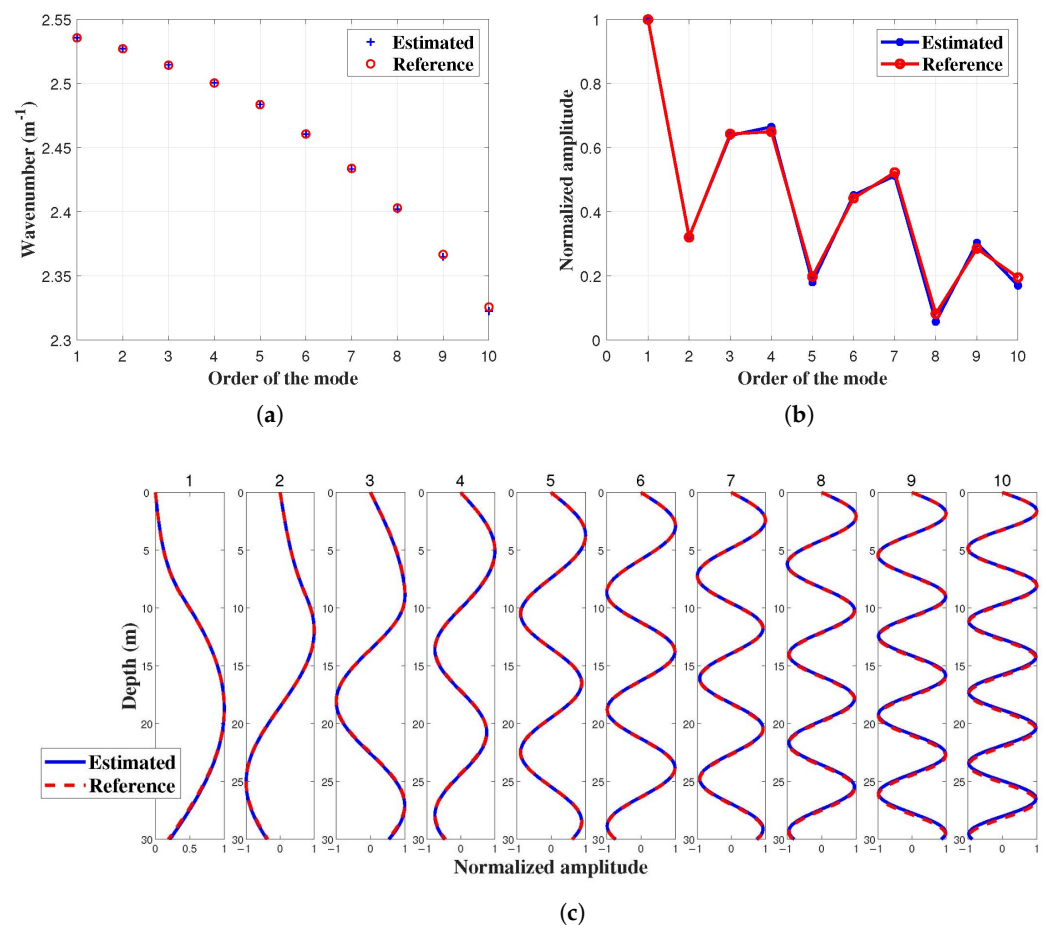


Figure 2. Comparison of estimated (a) horizontal wavenumbers, (b) mode amplitudes, and (c) mode depth functions with their reference values. Red markers indicate reference values, while blue markers represent estimated values.

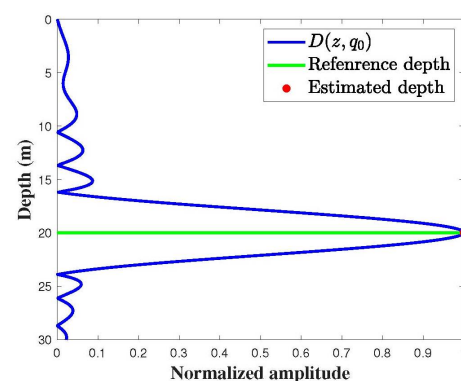


Figure 3. DAF compensated by the optimal mode sign combination (blue line). The green line indicates the reference depth.

Furthermore, in the following subsections, the depth estimation performance of OCMS-D is analyzed by varying the SNR, array aperture, number of array elements, and number of modes. For comparison, the estimation results of the MMP under these parameters are also presented, where the modal amplitudes are computed using the least square fit method [11]. Since the source distance is unknown, the DAF is constructed by enforcing positive contributions from all modal amplitudes. To highlight the impact of seabed parameter mismatch on MMP, the simulation further presents MMP estimation results under two typical mismatch scenarios, where the sediment type is incorrectly assumed

to be sand–silt–clay and coarse, respectively (while the true sediment type is sandy silt). The corresponding geoacoustic parameters are listed in Table 1.

3.2. Performance Under Different SNR Conditions

This subsection analyzes the performance of OCMS-D under different SNR conditions and compares its results with those of MMP. The SNR varies from -20 dB to 40 dB with increments of 2.5 dB. For each SNR condition, $J = 50$ independent Monte Carlo trials are conducted to evaluate the mean absolute error (MAE) and standard deviation of the depth estimation results. The absolute error of the j th Monte Carlo trial and the MAE are defined by

$$\begin{cases} \text{AE}_j &= |z_s - \hat{z}_{sj}| \\ \text{MAE} &= \frac{1}{J} \sum_{j=1}^J \text{AE}_j \end{cases} \quad (21)$$

where $\hat{z}_{sj}$ represents the estimated value of z_s at the j th Monte Carlo trial. The VLA configuration is kept consistent with that in the previous simulation. Results are shown in Figure 4. Figure 4a shows the MAE and standard deviation of the source depth estimated by OCMS-D and MMP under different SNR conditions. It can be observed from Figure 4a that under the given simulation parameters and without seabed parameter mismatch, the MAE of MMP (red curve) is lower than that of OCMS-D (blue curve) at low SNRs. However, when seabed parameter mismatch is present (black and magenta curves), the depth estimation performance of MMP degrades under low-SNR conditions, and this degradation becomes more severe as the mismatch increases. For example, when the sediment type is coarse, the depth estimation performance of MMP (magenta curve) is significantly worse than that of OCMS-D.

Figure 4b–d present the results of the modal parameters under different SNRs. The blue lines indicate the mean estimated values, and the blue shaded regions represent the standard deviation intervals. Since the estimated wavenumber error is small (maximum magnitude of 1×10^{-2}), the estimation results under different SNRs are directly presented in Figure 4b. Figure 4b also shows the wavenumbers calculated by the Kraken program as reference. Figure 4c presents the estimated errors for each mode depth function estimated by OCMS-D under different SNRs. To more clearly illustrate the impact of different orders of mode depth function errors estimated by OCMS-D as SNR varies, the errors are added to their respective order of the mode, i.e.,

$$\text{MAE}_{\text{E}}(m) = \frac{1}{JH} \sum_{j=1}^J \int_0^H \left| \psi_m^j(z, \hat{k}_m) - \phi_m(z, k_m) \right| dz + m, \quad (22)$$

where $\psi_m^j(z, \hat{k}_m)$ denotes the estimated value of mode depth function $\phi_m(z, k_m)$ at the j th Monte Carlo trial. The red dashed lines mark the zero-error line in Figure 4c. Similarly, Figure 4d shows the estimated errors of each modal amplitude vector, which is defined by

$$\text{MAE}_a = \frac{1}{J} \sum_{j=1}^J \|\mathbf{a}_j(\hat{k}_m) - \mathbf{a}(k_m)\|_2, \quad (23)$$

where $\mathbf{a}_j(\hat{k}_m)$ denotes the estimated value of $\mathbf{a}(k_m)$ at the j th Monte Carlo trial, and $\|\mathbf{a}(k_m)\|_2 = 1$. As observed in Figure 4b–d, the estimation errors of normal mode wavenumbers, mode depth functions, and $\mathbf{a}(k_m)$ all increase as the SNR decreases, which consequently leads to a growth in the target depth estimation error with declining SNR. However, it can be noted that the estimation errors of the normal mode parameters rise

only gradually as SNR decreases. This slow growth explains why the depth estimation error remains relatively small even under low-SNR conditions.

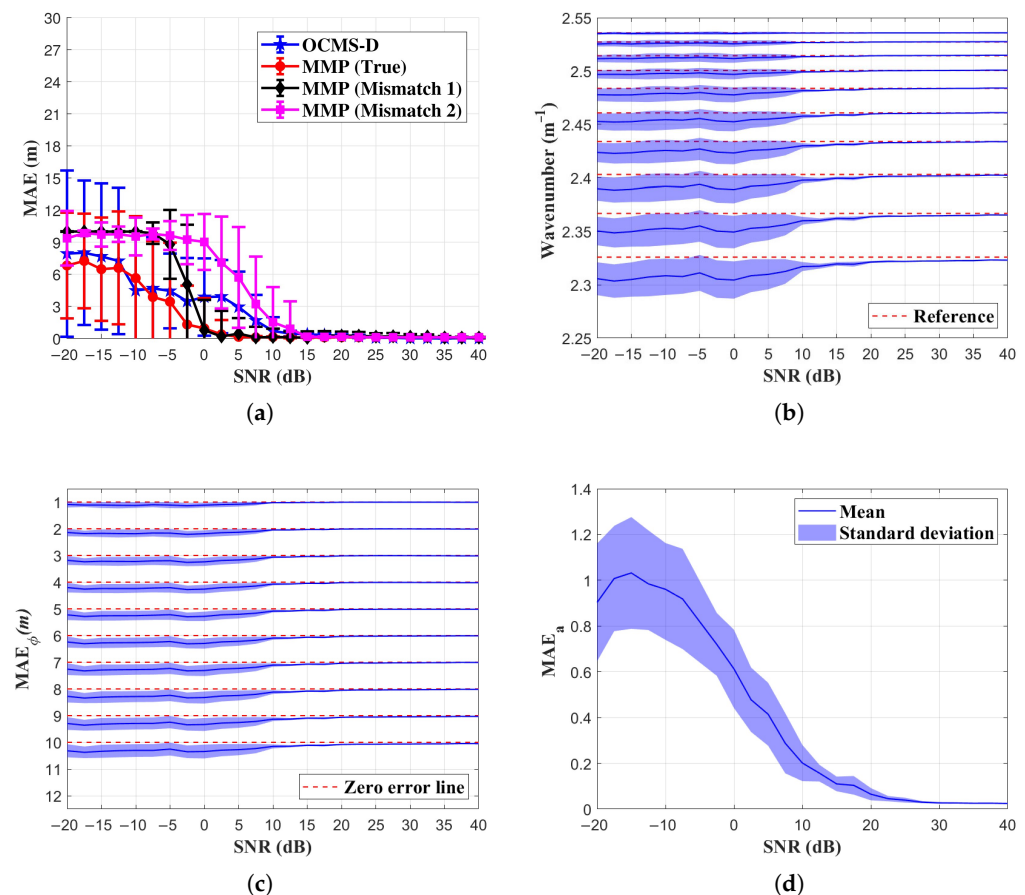


Figure 4. (a) Estimation results of source depth, along with their standard deviations, obtained by OCMS-D and MMP under different SNRs. (b–d) are estimation results of the wavenumbers, the mode depth functions, and the modal amplitudes, along with their standard deviations, obtained by OCMS-D under different SNRs.

3.3. Performance Under Different VLA Parameters

It is well known that the parameters of a VLA significantly influence array-based depth estimation methods. Therefore, this subsection evaluates the performance of OCMS-D under different VLA parameters with a fixed SNR of 30 dB. We vary the array aperture from 4 m to 29 m in steps of 1 m while keeping the number of elements fixed at 30, and also vary the number of elements from 5 to 30 with a fixed aperture of 29 m. The first element is always deployed at a depth of 1 m below the surface.

For the aperture variation, Figure 5a presents the estimation results of OCMS-D (solid blue line) and MMP (solid red, black and magenta lines corresponding to the true sediment type and two mismatch cases, respectively) under different array apertures. Notably, even when the seabed geoacoustic parameters are perfectly known, the aperture required by MMP is significantly larger than that required by OCMS-D under the considered simulation conditions. This result further demonstrates that OCMS-D does not require the VLA to span the entire water column.

Moreover, we also compare the depth estimation performance of OCMS-D and MMP when the VLA is deployed on the sea bottom versus at the sea surface. The simulation environment parameters are consistent with those specified earlier, with SNR set to 30 dB. The deepest element of the VLA is fixed at the bottom, and the array aperture is varied from 29 m down to 4 m in steps of 1 m. The corresponding results are shown as dashed lines

in Figure 5a. Comparing the surface and bottom deployments, one can find that placing the VLA on the bottom reduces the aperture required for accurate depth estimation for both methods: the required aperture for OCMS-D decreases by approximately 9 m, while that for MMP decreases by only about 2 m. A primary reason is that surface VLAs are not advantageous for receiving low-order modes. Taking the simulated source at 20 m depth as an example, the strongest excited mode is the first-order mode. As seen from Figure 2b,c, the first-order mode depth function is concentrated closer to the seabed; thus, in this case, deploying the VLA on the seabed enables effective sampling of this mode with a shorter aperture.

Furthermore, with the array aperture fixed at 29 m, we increase the number of VLA elements from 5 to 30 in steps of 1. Figure 5b presents the depth estimation results of OCMS-D and MMP at different numbers of elements when SNR = 30 dB. It can be observed that, under the given simulation conditions, OCMS-D achieves satisfactory depth estimation when the number of elements exceeds 9. With the true sediment type, MMP also yields satisfactory results when the number of elements exceeds 6. However, in the presence of mismatch, the performance of MMP degrades significantly, and the required number of elements becomes larger than that needed by OCMS-D.

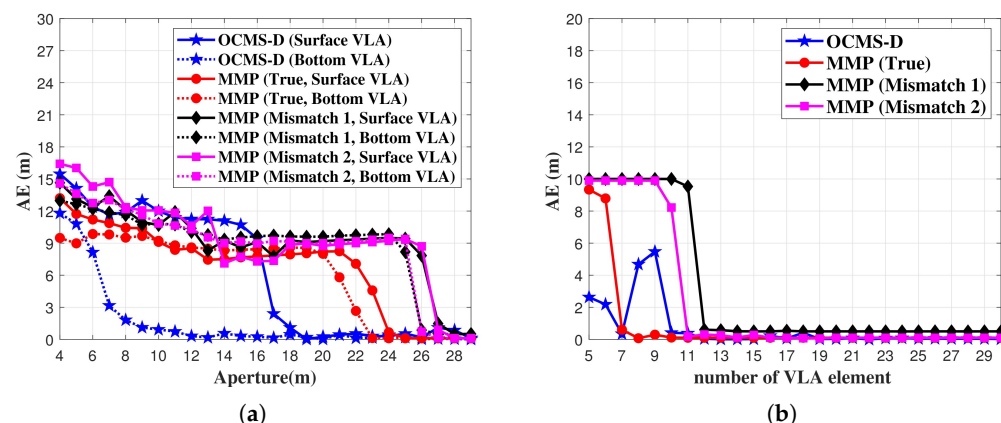


Figure 5. (a) Depth estimation results for OCMS-D and MMP at different array apertures when the SNR = 30 dB. Solid and dashed lines correspond to the VLA deployed at the surface and on the bottom, respectively. (b) Depth estimation results for OCMS-D and MMP at different numbers of VLA elements when the SNR = 30 dB.

3.4. Performance Under Different SSP Uncertainty

In the application scenario of OCMS-D, the SSP information is obtained from a sound speed profiler deployed with the VLA. However, in practical measurements, uncertainties are inevitably introduced into the SSP due to instrumental errors and dynamic oceanographic processes, including sensor calibration drift, spatiotemporal variability of water masses, and disturbances caused by internal waves. This poses an additional practical challenge to OCMS-D. Therefore, it is necessary to analyze the impact of SSP uncertainty on the OCMS-D method. Typically, the uncertainty in sound speed can be modeled as

$$c(z) = c_0(z) + \eta(z) \cdot \beta, \quad (24)$$

where $c_0(z)$ is the background SSP, $\eta(z)$ is a uniformly distributed random number in $[-1, 1]$, and its values are independent and identically distributed across different values of z . β is the sound speed variation ranging from 0.1 m/s to 5 m/s in steps of 0.25 m/s. In the simulations, the ocean environmental parameters and source parameters used to compute the acoustic field received by the VLA are the same as those in Section 3.1. The VLA covers the entire water column and consists of 16 elements. The SNR is set to 20 dB, and $J = 50$

independent Monte Carlo simulations are performed for each SSP uncertainty to evaluate the MAE of the results. OCMS-D performs depth estimation using the SSP given by Equation (24). The results are shown in Figure 6. For comparison, the estimation results of MMP under correct seabed parameters and two mismatch scenarios are also presented. Figure 6 shows that the results of OCMS-D are more robust to SSP uncertainty than those of MMP. Furthermore, the mean absolute error in the depth estimated by OCMS-D is less than 1 m when the uncertainty in the sound speed profile measurement is less than 2.5 m/s.

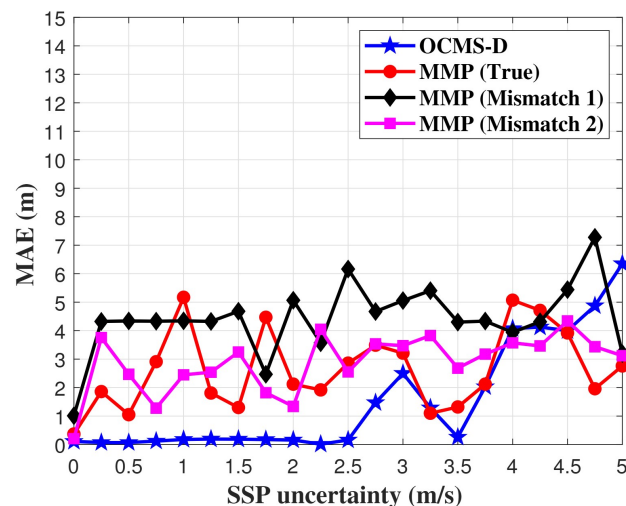


Figure 6. The MAE of depth estimation results under different SSP uncertainties.

4. Experimental Validation with SWellEx-96 Data

The validation of OCMS-D was performed using S5 event data from the SWellEx-96 experiment in this section. During this event, the source ship towed a deep source (at a depth of 54 m underwater) and a shallow source (at a depth of 9 m underwater) from southwest to northeast at five knots. The source trajectory is shown in Figure 7a, with most of the water depths along the trajectory ranging from 180 m to 220 m. The deep source emitted a series of tones between 49 Hz and 400 Hz, and the shallow source emitted nine frequencies between 109 Hz and 385 Hz. In addition, a 64-array VLA (data for 21 receivers of the VLA were published on the SWellEx-96 website) was deployed during the experiment at the location shown in Figure 7b, at a water depth of 216.5 m, spanning the bottom half of the water column. The lower-frequency signals from both the deep source and the shallow source were processed. Specifically, we chose 112 Hz and 130 Hz frequencies for the deep source, and 109 Hz and 127 Hz frequencies for the shallow source in data processing. To obtain clearer line spectra data for all four frequencies, 5 min data segments were processed as an example from the SWellEx-96 experiment, specifically from 10 min to 15 min after the start of the experiment, as indicated by the black box in Figure 7b. During data processing, a 16,384-point Fourier transform was applied to each non-overlapping 10 s segment of the recordings received by each VLA element to obtain the corresponding frequency-domain pressures with sufficient SNR. The frequency-domain data from all array elements within the same time segment were combined into a pressure vector $\mathbf{p}$, yielding a total of 31 frequency-domain pressure vectors.

Regarding source-range variation, the absolute values of the modal amplitudes were taken, as shown in Equation (12), to eliminate the influence of source range in the phase terms. The effect of range-dependent attenuation on the modal amplitude distribution was considered small and was therefore neglected in the data processing.

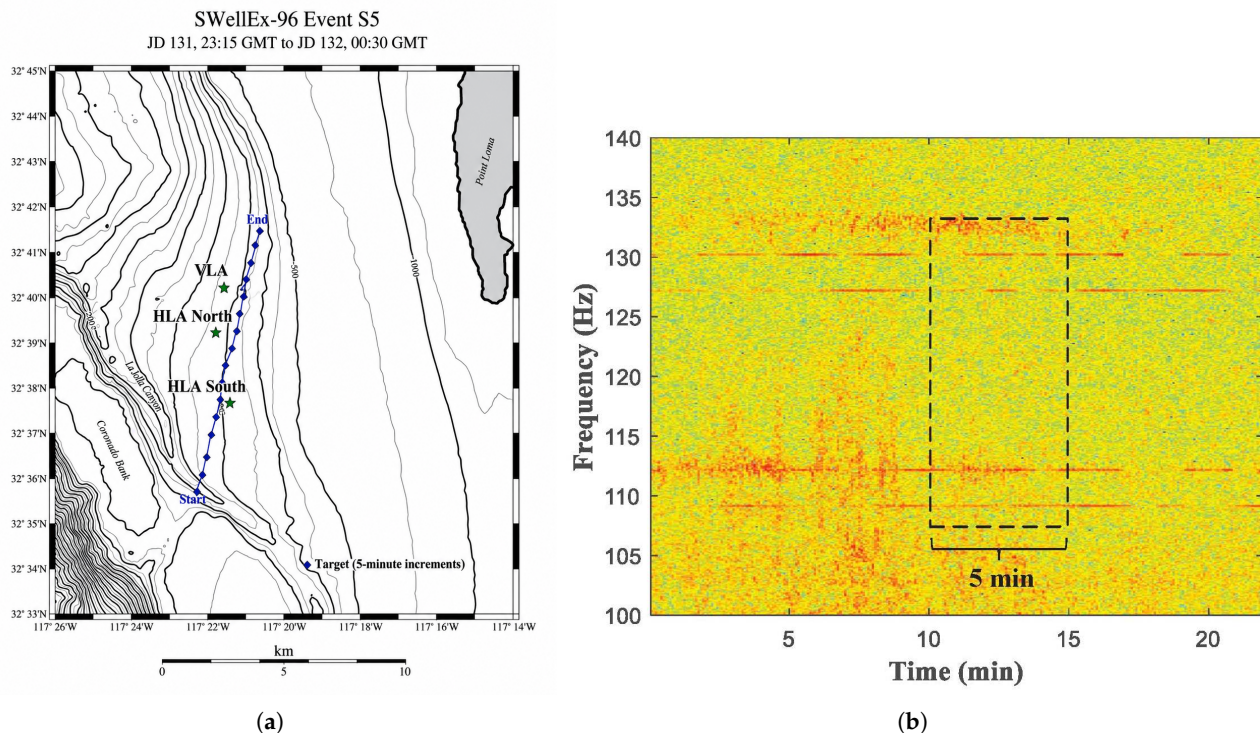


Figure 7. (a) Trajectory of the source movement for the S5 event and the locations of hydrophone arrays. (b) Time–frequency diagram of the first 22 min of data received by the hydrophone closest to the water surface. The data processed in this subsection are highlighted in the black box.

In this section, we chose $\zeta_{\min} = 2\pi f/1822 \text{ m}^{-1}$ and $\zeta_{\max} = 2\pi f/1488 \text{ m}^{-1}$ for different frequencies, and the sign sampling interval in the depth direction was fixed at 0.1 m. Figure 8 illustrates the DAFs for the deep source and shallow source estimated by the OCMS-D and MMP when the sources were approximately 6.89 km from the VLA. The red and blue curves denote the results of OCMS-D, while the yellow and purple curves denote those of MMP. The cyan line indicates the reference depth. Note that although the VLA covers only about half the water column, the OCMS-D method is still able to estimate the source depth accurately. In contrast, the DAFs obtained by the MMP method (for both shallow and deep sources) exhibit sidelobe peaks exceeding the main lobe, resulting in significant depth estimation errors.

Furthermore, the results of all 31 data segments for the deep source and shallow source are shown in Figure 9. Similar to Figure 8, blue and red markers denote the results of OCMS-D, yellow and purple markers denote the results of MMP, and cyan markers indicate the reference depth. For the deep source, the depth estimates obtained by OCMS-D at two frequencies all fall within the range of 46.4 m to 64.8 m, with a maximum AE of 10.8 m. For the shallow source, the estimates lie between 7.7 m and 14.4 m, with a maximum AE of 5.4 m. In contrast, the results of MMP for the deep source are generally distributed between 12.8 m and 43.5 m, but a number of segment results approach 200 m, resulting in significantly larger estimation errors than those of OCMS-D. Similar conclusions can be drawn for the shallow source results of MMP. Overall, OCMS-D demonstrates considerably better performance than MMP on the experimental data.

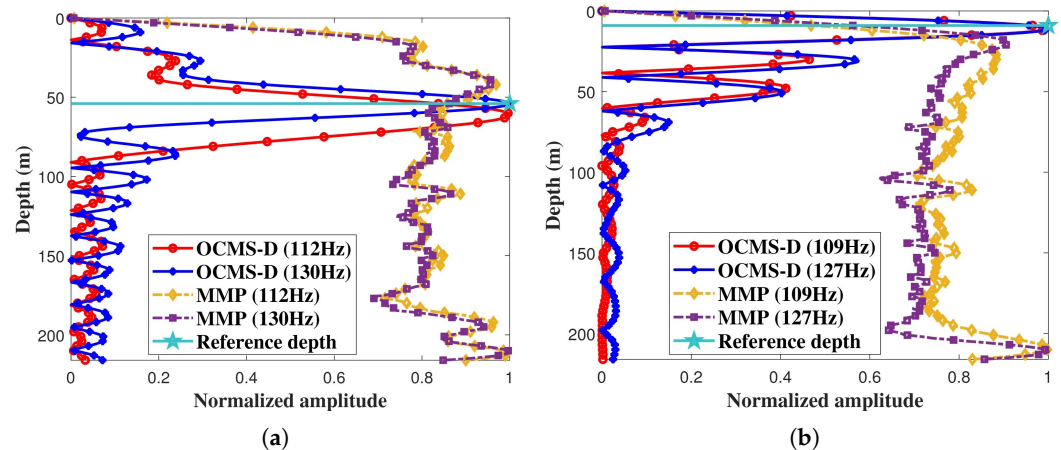


Figure 8. The DAFs at the four frequencies estimated by the OCMS-D when the sources were approximately 6.89 km from the VLA. (a) The deep source's DAFs at 112 Hz (red line) and 130 Hz (blue line). (b) The shallow source's DAFs at 109 Hz (red line) and 127 Hz (blue line). The green line indicates the reference depth.

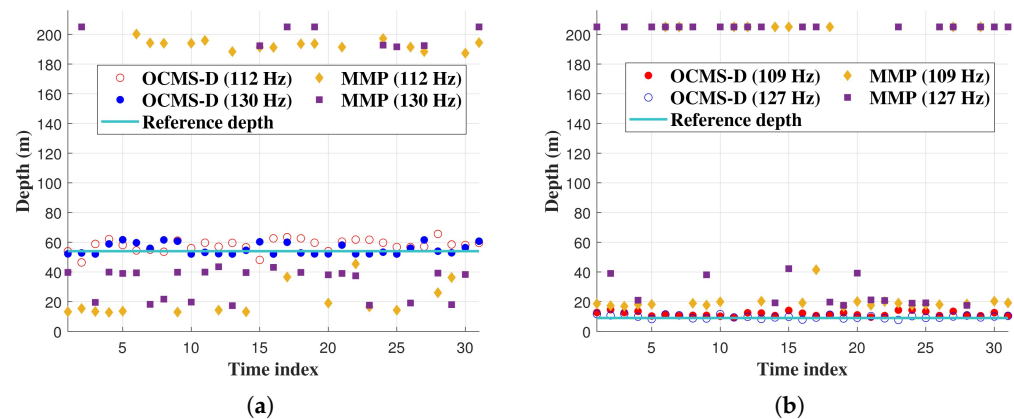


Figure 9. Results of the estimated depths. (a) The shallow source. (b) The deep source using OCMS-D and MMP for all 31 data segments.

5. Conclusions

This paper proposes a single-frequency source depth estimation method (OCMS-D) based on the orthogonality of mode depth functions and the sparsity of propagating normal modes in the received acoustic field. The method requires only the known water-column sound speed profile as the environmental input. First, the candidate modal space is constructed by combining the difference equation and the orthogonality normalization condition. Second, a sparse solver based on convex optimization is employed to jointly estimate the modal wavenumbers, depth functions, and complex amplitudes. Third, to address the inherent phase-sign ambiguity in the DAF, a depth-sign search (DSS) mechanism based on Kullback–Leibler (KL) divergence is introduced. By selecting the optimal sign combination, the sidelobes of the ambiguity function are effectively suppressed, thereby significantly improving the accuracy of the final depth estimation.

Numerical simulations demonstrate that the method maintains excellent estimation performance when the SNR exceeds 7.5 dB and the array aperture is larger than 16 m, verifying its robustness under limited array apertures. Comparisons with MMP show that, unlike MMP, whose performance degrades significantly under seabed parameter mismatches, OCMS-D does not require any seabed prior information and thus avoids such mismatch-induced errors. The method also shows stronger tolerance when the SSP

measurement uncertainty is less than 2.5 m/s. Experimental results from the SWellEx-96 campaign validate the effectiveness of OCMS-D in real ocean environments. For the shallow source (approximately 9 m), the depth estimation errors are better than 5.4 m, with a maximum relative error of about 2.49%; for the deep source (approximately 54 m), the errors are better than 10.8 m, with a maximum relative error of about 4.99%. The estimated depths at different frequencies are in good agreement with the reference depths, confirming that OCMS-D exhibits good adaptability to various source depths and frequencies.

Compared with existing approaches, the core advantages of OCMS-D are threefold: (1) it requires no prior knowledge of seabed geoacoustic parameters, only the water-column sound speed profile; (2) it does not require source motion, making it suitable for single-snapshot processing at fixed ranges; and (3) it does not require broadband signals and can handle single-frequency tonal sources. These features make OCMS-D a promising candidate for passive sonar detection in unknown seabed environments. Future work will investigate the robustness of the method under time-varying ocean conditions and explore the feasibility of extending it to moving source scenarios with motion compensation strategies.

Author Contributions: Conceptualization, W.G.; methodology, W.G.; software, Y.X.; validation, W.G., Y.X. and G.G.; formal analysis, W.G.; investigation, W.G.; resources, W.G.; data curation, Y.X.; writing—original draft preparation, Y.X.; writing—review and editing, W.G.; visualization, G.G.; supervision, W.G.; project administration, W.G.; funding acquisition, W.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Basic and Interdisciplinary Frontier Scientific Research Pilot Project (Category B) of the Chinese Academy of Sciences, grant number XDB1110102.

Data Availability Statement: The SWellEx-96 experimental data analyzed in this study are publicly available on the SWellEx-96 experiment website at <https://swellex96.ucsd.edu/> (accessed on 22 July 2026) [28].

Acknowledgments: The authors gratefully acknowledge the SWellEx-96 experiment team for making the data publicly available, which was essential for validating the proposed method. This work was supported by the Chinese Academy of Sciences. During the preparation of this manuscript, the authors did not use any generative AI tools for text generation, data analysis, or figure creation. The authors have reviewed and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

OCMS-D	Orthogonality-Constrained Modal Search-based Depth estimation
MFP	Matched Field Processing
MMP	Matched Mode Processing
DSS	Depth-Sign Search
VLA	Vertical Line Array
SNR	Signal-to-Noise Ratio
SSP	Sound Speed Profile
DL	Deep Learning

References

1. Premus, V.E. Mode scintillation and surface/submerged discrimination in SWELLEX-96. *J. Acoust. Soc. Am.* **1998**, *104*, 1837. [CrossRef]
2. Zhang, Y.; Yang, Q.; Yang, K. A source depth discrimination method based on wavenumber measured by bottom-mounted hla. *Ocean. Eng.* **2023**, *286*, 115502. [CrossRef]

3. Li, J.; Chen, Z.; Liu, S. Source depth discrimination using wavenumber domain feature with a horizontal array. *Appl. Acoust.* **2020**, *164*, 107287. [\[CrossRef\]](#)
4. Conan, E.; Bonnel, J.; Chonavel, T. Using the trapped energy ratio for source depth discrimination with a horizontal line array: Theory and experimental results. *J. Acoust. Soc. Am.* **2017**, *142*, 2776–2786. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Huaigang, Z.; Ren, Q.; Wang, Z. The discrimination of source depth regions in shallow water based on mode phase relation. *Acta Acust.* **2020**, *45*, 801–810.
6. Li, Z.; Guo, L.; Yan, C. Source depth discrimination in shallow water based on relation formula warping transform. *Chin. J. Acoust.* **2019**, *44*, 28–38.
7. Yao, Y.; Sun, C.; Liu, X.; Duan, R.; Xie, L.; Lu, Y. Multipath-based broadband source depth estimation using a single hydrophone. In Proceedings of the 2018 OCEANS—MTS/IEEE Kobe Techno-Oceans (OTO), Kobe, Japan, 28–31 May 2018; pp. 1–5.
8. Zhou, L.; Zheng, G.; Yang, T. Target depth estimation by frequency interference matching for a deep vertical array. *Appl. Acoust.* **2022**, *186*, 108493. [\[CrossRef\]](#)
9. Baggeroer, A.; Kuperman, W.; Mikhalevsky, P. An overview of matched field methods in ocean acoustics. *IEEE J. Ocean. Eng.* **1993**, *18*, 401–424. [\[CrossRef\]](#)
10. Bogart, C.W.; Yang, T.C. Source localization with horizontal arrays in shallow water: Spatial sampling and effective aperture. *J. Acoust. Soc. Am.* **1994**, *96*, 1677–1686. [\[CrossRef\]](#)
11. Yang, T.C. A method of range and depth estimation by modal decomposition. *J. Acoust. Soc. Am.* **1987**, *82*, 1736–1745. [\[CrossRef\]](#)
12. Shang, E.C. Source depth estimation in waveguides. *J. Acoust. Soc. Am.* **1985**, *77*, 1413–1418. [\[CrossRef\]](#)
13. Yang, T.C. Data-based matched-mode source localization for a moving source. *J. Acoust. Soc. Am.* **2015**, *137*, 2986. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Yang, T.C. Source depth estimation based on synthetic aperture beamforming for a moving source. *J. Acoust. Soc. Am.* **2015**, *138*, 1678–1686. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Li, X.; Xu, Y.; Gao, W.; Wang, H.; Wang, L. Harmonic source depth estimation by a single hydrophone under unknown seabed geoacoustic property. *Remote Sens.* **2024**, *16*, 2227. [\[CrossRef\]](#)
16. Li, X.; Wang, H.; Gao, W. Modal wavenumber estimation by combining physical informed neural network. *J. Acoust. Soc. Am.* **2023**, *153*, 2637–2640. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Zhu, S.; Gao, W.; Li, X. Ssanet: Normal-mode interference spectrum extraction via ssa algorithm-unrolled neural network. *Front. Mar. Sci.* **2023**, *10*, 1342090. [\[CrossRef\]](#)
18. Gao, W.; Li, X.; Wang, H. Recovering reverberation interference striations by a conditional generative adversarial network. *JASA Express Lett.* **2021**, *1*, 056001. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Niu, H. Evaluation of data-driven neural operators in ocean acoustic propagation modeling. *J. Acoust. Soc. Am.* **2024**, *155*, A44. [\[CrossRef\]](#)
20. Huang, Z.; Xu, J.; Gong, Z. Source localization using deep neural networks in a shallow water environment. *J. Acoust. Soc. Am.* **2018**, *143*, 2922–2932. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Wang, W.; Wang, Z.; Su, L.; Hu, T.; Ren, Q.; Gerstoft, P.; Ma, L. Source depth estimation using spectral transformations and convolutional neural network in a deep-sea environment. *J. Acoust. Soc. Am.* **2020**, *148*, 3633–3644. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Chi, J.; Li, X.; Wang, H.; Gao, D.; Gerstoft, P. Sound source ranging using a feed-forward neural network trained with fitting-based early stopping. *J. Acoust. Soc. Am.* **2019**, *146*, EL258–EL264. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Grant, M.; Boyd, S. CVX: Matlab Software for Disciplined Convex Programming, Version 2.1. 2014. Available online: <http://cvxr.com/cvx> (accessed on 22 July 2026).
24. Li, X.; Wang, H. Modal wavenumber extraction by finite difference vertical linear array data. *JASA Express Lett.* **2021**, *1*, 126002. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Simon Foucart, H. *A Mathematical Introduction to Compressive Sensing*; Birkhäuser: New York, NY, USA, 2013.
26. Cheney, M.; Kirsteins, I. Resolution of matched field processing for a single hydrophone in a rigid waveguide. *J. Acoust. Soc. Am.* **2022**, *152*, 3186–3197. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Porter, M.B. *The KRAKEN Normal Mode Program*; Naval Research Laboratory Report; SACLANT Undersea Research Centre: La Spezia, Italy, 1992.
28. Murray, J.; Ensberg, D. The SWellEx-96 Experiment. 2018. Available online: <http://swellex96.ucsd.edu> (accessed on 22 July 2026).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.